ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

Design and Evaluation of a Local RAG-Based AI System for Legal Information Retrieval in the Democratic Republic of Congo

Ruben Kanku[1]; Hervé Kinkete[2]; Pathy Nkayilu Wabaluku[3]; Bruno Luwa Muanda[4]; Michel Kabeya Kadima[5]; Pontien Katukumbanyi Katukumbanyi[6]

[1;2;4;5]Chef Des Travaux; [3]Assistant

[1;2;3;4]Université William Booth
[5;6]ISC Tshikapa

Publication Date: 2026/04/09

Abstract: This paper presents the design and evaluation of a local Retrieval-Augmented Generation (RAG)-based system for legal information retrieval and question answering in the Democratic Republic of Congo (DRC). In this context, legal texts are scattered across multiple sources and are difficult to access and interpret, creating challenges for both citizens and local officials. To address this problem, we developed an AI-based system that integrates a vector database with a small language model to retrieve relevant legal provisions and generate grounded explanations. The proposed system is designed to run fully on local hardware (e.g., a personal computer using Ollama) while also supporting deployment on a server through a web interface. The prototype indexes 28 PDF documents covering 11 major domains of Congolese law and allows users to submit natural language queries in French. For each query, the system retrieves relevant legal articles and produces structured explanations, explicitly citing the source documents. A scenario-based evaluation was conducted using realistic legal questions, combined with manual expert review. Results show that the system is able to retrieve and explain key legal provisions in most cases, behaves cautiously when no relevant information is found, and maintains acceptable response times on standard local hardware. These findings demonstrate that local, deployable RAG-based systems can provide an effective technical solution for legal information access in low-resource environments. The study also highlights the importance of structured legal data, system transparency, multilingual support, and appropriate governance frameworks for AI-based legal systems.

Keywords: *Retrieval-Augmented Generation (RAG); Information Retrieval; AI System Design; Vector Database; Small Language Models (SLMs); Legal Question Answering; Democratic Republic of Congo (DRC).*

How to Cite: Ruben Kanku; Hervé Kinkete; Pathy Nkayilu Wabaluku; Bruno Luwa Muanda; Michel Kabeya Kadima; Pontien Katukumbanyi Katukumbanyi (2026) Design and Evaluation of a Local RAG-Based AI System for Legal Information Retrieval in the Democratic Republic of Congo. *International Journal of Innovative Science and Research Technology,* 11(3), 3656-3669. https://doi.org/10.38124/ijisrt/26mar1798

I. INTRODUCTION

➢ *Background and Context*

Across many low- and middle-income countries, access to justice is constrained not only by under-resourced courts but also by low legal literacy and the high cost of professional advice. Research on legal empowerment shows that poor and marginalized communities often do not know, cannot use, or cannot shape the law. Because they are unaware of their rights, they are unable to claim them, meaning that formal legal protections rarely translate into practical remedies in everyday life (Commission on Legal Empowerment of the Poor, 2008; Goodwin & Maru, 2017; Joshi, 2017).

According to civil society organisations and foreign observers, there are long-standing barriers to justice, such as long, complicated procedures, ineffective legal assistance, pervasive impunity and the continued use of paper files by the judiciary (Amnesty International, 2023; Avocats Sans Frontières, 2023; United Nations Development Programme [UNDP], 2020). Even though the nation has embraced many codes and special acts, they are spread in various publishers, ministry sites and informal PDFs. The common citizens can hardly learn the location of the corresponding provisions, and even the professionals have difficulties with disjointed sources (Avocats Sans Frontières, 2023).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

There is also weak confidence in legal interpretation. The law is usually seen as partisan in nature in a society where there is political and ethnic polarisation; that is, the interpretation of the law is considered to be subjective with different interpretations based on political affiliation, societal affiliation (tribalism), or wealth. The low scores of civil justice and absence of corruption in the DRC mentioned in the World Justice Project (2023) Rule of Law Index also contribute to the beliefs that the law is not a neutral system but is directly driven by elites. To citizens and small business, the net effect is paralysis: not knowing what they are entitled to and worrying about it being interpreted arbitrarily, they tend to give up on their claims before they even begin.

Recent advances in artificial intelligence (AI) offer new tools to address these information gaps. Retrieval-Augmented Generation (RAG) combines a search mechanism over a curated corpus with a generative language model, so that answers are explicitly grounded in retrieved documents rather than produced from the model's internal memory alone (Lewis et al., 2020). In the legal field, RAG-based systems are increasingly preferred for research and question-answering because they mitigate "hallucinations" by forcing the model to cite concrete sources such as statutes or case law (Cui et al., 2023; Blair-Stanek et al., 2023). Industry surveys show that legal professionals are already adopting generative AI tools for document review and research, and expect these technologies to become central to their workflows (Thomson Reuters Institute, 2024).

However, most existing legal AI tools are designed for cloud-based Large Language Models (LLMs) and focus on Global North jurisdictions. Academic benchmarks and commercial products tend to concentrate on US, UK or EU law, while African legal systems are largely absent from training datasets (Henderson et al., 2022; Gathii, 2022). To our knowledge, there is currently no publicly documented system that targets Congolese law through a fully local RAG architecture.

- *Problem Statement*

The DRC lacks an intelligent digital platform to be neutral, in which ordinary people get to know their rights and their duties as per the national laws. The law is fragmented and scattered over numerous, poorly indexed, and is frequently composed in hard-to-understand dense legal language; and can be interpreted, not on a basis of legal reasoning, but on the basis of political or financial motives.

As a result, citizens, community leaders and even some local officials frequently rely on hearsay or expensive legal advice. Many people abandon their claims because they do not know which law applies or which authority is competent to act (World Justice Project, 2023).

The core problem this paper addresses is therefore: *How can a local, retrieval-based AI assistant be designed to provide neutral, understandable explanations of Congolese law to non-lawyers, while displaying the exact legal sources used?*

- *Aim and Research Objectives*

The overall aim of this study is to design, implement and explore a prototype AI legal assistant for Congolese law, which runs fully on local infrastructure and can also be deployed on a server so that users can access it through a web interface.

To be more precise, the work of the former objective is to create a single digital collection of essential Congolese legal documents, including the Labour Code, the Penal Code, the Family Code, and Mining Code in a systematic format, which will be easily retrieved by computers.

The second target is to deploy a RAG architecture that indexes this corpus to a vector database, finds relevant articles to natural-language queries, and (with a small language model (SLM)) deployed via Ollama) comes up with explanations based on those articles. The third goal is to create user interfaces, such as a command-line application and a web app, through which citizens can pose legal questions in French and see the explanation and the legal provisions.

The fourth aim is to assess the topicality, clarity and perceived impartiality of the answers provided by the assistant to a number of realistic legal cases based on the everyday Congolese life e.g., employment matters, family affairs, and land challenges.

Finally, the fifth objective is to reflect on the opportunities and risks of deploying such systems in the DRC, and to propose policy directions for the responsible use of AI in the justice sector.

- *Research Questions*

The study is guided by three main research questions. The first question (RQ1) asks how retrieval-augmented generation can be adapted to support question-answering over Congolese legal texts when using a local small language model, rather than a large cloud-based model.

The second question (RQ2) will determine to which extent the RDC-Law-RAG prototype can retrieve the relevant legal provisions and give a comprehensible explanation of typical citizens question.

The third question (RQ3) addresses what advantages and disadvantages various parties in the DRC justice ecosystem think about the usage of AI-based legal assistants, and what types of protective measures are required to make the use of these tools responsible.

- *Scope and Limitations*

The current work is devoted to the national legislation of the DRC that is presented in PDF/text format, i.e. to major codes. It is yet to contain extensive jurisprudence or customary law, but the framework can be extended at any time.

The prototype will be developed to interact in French language, considering that the codes are written in French as

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

the official language. The national support of the languages (Lingala, Swahili, Kikongo, Tshiluba) is necessary to be included but not to be covered by this initial technical implementation.

On legal grounds, the assistant can only give general information about the law (legal empowerment), but not individual advice.

Finally, the evaluation is based on technical metrics and qualitative expert feedback, rather than a large-scale field trial.

- *Significance of the Study*

Despite these limitations, the study offers several important contributions. First, it provides a practical innovation tailored to the DRC context. The prototype demonstrates that a fully local AI legal assistant combining a vector database and an SLM can run on standard hardware without sending data to external servers. This is vital in contexts where internet connectivity is unreliable, and data sovereignty is a priority.

Second, the project supports legal empowerment strategies by enabling citizens to "know and use" the law (Goodwin & Maru, 2017).

Third, the findings provide evidence for policy debate. As African bar associations and governments consider AI regulations, concrete case studies like RDC-Law-RAG can inform standards for transparency and model governance in the justice sector (Hagan, 2020; UNESCO, 2023).

II. LITERATURE REVIEW

- *Access to Justice, Legal Information and Legal Empowerment*

Access to justice is widely recognised as a core dimension of development and a key element of the rule of law. The Commission on Legal Empowerment of the Poor (2008) argued that billions of people live "outside the protection of the law", lacking secure rights, enforceable contracts and accessible dispute-resolution mechanisms. Legal empowerment approaches seek to change this by enabling people to know, use and shape the law, rather than treating them as passive recipients of state decisions (Goodwin & Maru, 2017; Joshi, 2017). Empirical reviews show that common strategies include legal literacy, community paralegals, public interest litigation and social accountability mechanisms that link citizens' claims to institutional reform (Maru, 2006; Joshi, 2017). However, Rhode (2004) notes that access to justice remains in crisis even in developed nations, a situation exacerbated in low-resource settings by the scarcity of legal professionals.

The necessity to be legally empowered in African contexts is also strongly connected with access to the adequate legal information. According to AfricanLII (2019) and Gathii (2022), in most countries of the African continent, basic legal documents like constitutions, laws and regulations are hard to obtain, blocked by payment gates or are stored in print libraries that are outdated. This compromises the rights of the citizens as well as the efficacy of the courts and the administrative entities. Open-access legal databases and legal education programmes online are some of the new initiatives aimed at curbing these obstacles by digitalising and publishing national laws in structured format (AfricanLII, 2019; Laws.Africa, n.d.).

Recent work on "legal empowerment in the digital age" emphasises that technology can support victims and marginalised groups. Susskind (2019) argues that the future of access to justice lies in "online courts" and digital extended legal services. However, scholars caution that technology alone cannot fix structural problems in the justice system; tools must be designed with local contexts and power relations in mind to make legal information more visible, understandable and actionable for lay users (Sandefur, 2019; Hagan, 2020). Furthermore, Passera (2017) highlights the importance of "legal design" using visualizations and plain language to ensure that digital legal tools are actually intelligible to non-lawyers.

- *Digital Legal Information Systems and Online Legal Aid*

Digital legal information systems range from simple online repositories of legislation to more advanced platforms that integrate case law, commentary and search interfaces. On the African continent, projects such as Laws.Africa aim to provide open access to up-to-date legislative texts, arguing that free digital availability of legal information helps communities to thrive, supports businesses and improves public administration (Laws.Africa, n.d.). At the same time, legal aid organisations have used online tools websites, SMS, chat and e-learning to expand legal education and support to prisoners, rural communities and other underserved groups (Justice Defenders, 2021; World Justice Project, 2023; Open Society Foundations, 2019).

More recently, researchers and practitioners have discussed the use of digital tools to promote public access to justice by integrating legal information and interface-friendly interfaces. Case studies also emphasize the opportunities and limitations of such tools: they should be able to decrease the distances and time spent waiting, but they have to address poor infrastructure, low digital literacy and the possibility to strengthen existing inequalities (Donner et al., 2020; Sambasivan et al., 2021). Gwagwa et al. (2020) also point out to a heterogeneity of AI and digital adoption in Africa, cautioning against global North-to-local importation of solutions.

In the Democratic Republic of the Congo (DRC) specifically, most efforts have focused on digitising court records and improving case management within the judiciary, rather than building citizen-facing platforms. Reports on justice reform point out that legal texts remain scattered across ministry websites and unofficial PDF collections, with limited searchability and no integrated, public portal for Congolese law (Avocats Sans Frontières, 2023; UNDP, 2020). This fragmentation creates a gap that a digital assistant grounded in the official texts could help to address.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

➢ *Large and Small Language Models, RAG and Legal Applications*

Large language models (LLMs) such as GPT-4 have shown strong performance on a wide range of language tasks, including question answering, summarisation and translation. Ashley (2017) notes that AI has long been a goal in law, but modern LLMs represent a paradigm shift in capability. However, their knowledge is stored implicitly in model parameters and can be incomplete, outdated or difficult to update. Lewis et al. (2020) introduced Retrieval-Augmented Generation (RAG) as a way to combine a pre-trained sequence-to-sequence model with a dense vector index of documents, allowing the system to retrieve relevant passages and then generate answers conditioned on that retrieved context. RAG has since become a standard architecture for knowledge-intensive tasks to reduce inaccuracies (Jurafsky & Martin, 2024).

In the legal domain, researchers have begun to adapt RAG to legal information retrieval. Cui et al. (2023) developed "ChatLaw," a system that integrates external knowledge bases to improve logical reasoning in legal queries. Similarly, Yu et al. (2024) demonstrated that grounding language models in statutory texts and precedents significantly reduces "hallucinations" compared to standard LLMs. Studies on contract law and property law suggest that RAG-based systems can deliver more accurate, citation-backed legal responses when restricted to authoritative sources (Chalkidis, 2023; Blair-Stanek et al., 2023; Nay, 2022).

While massive models dominate the headlines, there is a shifting trend toward "Small Language Models" (SLMs). Touvron et al. (2023), in their work on Llama, demonstrate that smaller, highly efficient models can achieve competitive performance if trained on high-quality tokens. This is crucial for local deployment where computational resources are limited (Hoffmann et al., 2022). Industry reports indicate that a rising share of law firms are using generative AI for research and document drafting, with domain-specific RAG considered a key feature for trustworthiness (Thomson Reuters Institute, 2024; Perlman, 2022). Technical analyses highlight that RAG allows legal AI tools to work over curated collections such as legislation, providing explicit citations and verification of sources (Katz et al., 2024).

Recent prototypes aim to improve access to legal information for lay users. For example, the Open Justice initiatives and various "Legal Hackers" projects illustrate the goal of democratising legal knowledge through AI. While specific local tools are emerging, the consensus remains that combining RAG with modern LLMs provides the most viable path for interactive access to statutes (Savelka et al., 2023).

➢ *Risks, Ethics and Governance of AI in Law*

Scholars, regulators and professional bodies warn about the risks of AI in legal practice. LLMs are prone to "hallucinations" highly fluent but incorrect statements and fabricated citations which can have serious consequences. The widely publicised case of *Mata v. Avianca* (2023), where lawyers submitted AI-generated briefs with non-existent case law, highlights the danger of uncritical reliance on generic tools (Weiser, 2023; Merken, 2023). Professional associations have issued guidance emphasising duties of competence and verification (American Bar Association, 2024; New York City Bar Association, 2024). Surden (2019) argues that while ML is powerful, it lacks the abstract reasoning capabilities of human lawyers, necessitating "human-in-the-loop" systems.

For these reasons, RAG architectures are preferred in law because they constrain the model to answer based on a defined corpus. However, data protection and sovereignty remain concerns, especially in the Global South. Abebe et al. (2021) and Sambasivan et al. (2021) note that dependence on external AI platforms can deepen digital dependency. Birhane (2021) goes further, warning of "algorithmic colonization" where African nations rely on opaque systems controlled by Western tech giants. This drives the need for "Small Language Models" (SLMs) and local deployments to retain control over data while benefiting from AI capabilities (Touvron et al., 2023).

➢ *Identified Gap and Contribution of this Study*

The literature points to three main trends. First, there have been long-standing efforts to promote legal empowerment and improve access to legal information, especially for poor and marginalised groups (Commission on Legal Empowerment of the Poor, 2008; Maru, 2006). Second, there is a growing wave of digital legal tools, including open-access legislation platforms and online resources that make statutes easier to search and read (Laws.Africa, n.d.; Susskind, 2019). Third, recent work has begun to experiment with retrieval-augmented generation (RAG) and other AI systems for legal research, using vector databases and language models to support question-answering over complex legal corpora (Cui et al., 2023; Katz et al., 2024).

Nevertheless, little of this is done in relation to the Democratic Republic of Congo or to fully local architectures that can execute themselves on standard hardware without foreign cloud services. There is a shortage of literature that integrates the objectives of legal empowerment with small language models and RAG customized to the legal systems of Africa. This paper fills in this gap by developing and evaluating RDC-Law-RAG, a prototype of local RAG-based Congolese law assistant. It not only provides a technical demonstration of a concept but also early evidence of how the tools might be incorporated into an overall justice reforms and legal empowerment agenda in the DRC.

However, most RAG-based legal assistants have been developed for high-income countries, rely on cloud-based LLMs and focus on English-language corpora (Henderson et al., 2022). Work on African legal technology has concentrated more on case-management systems and basic digitisation (Gathii, 2022; Gwagwa et al., 2020). To our knowledge, there is no documented case of a fully local RAG system dedicated to Congolese law, built around national codes and designed to run on standard hardware using a small language model.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

This study addresses that gap by designing and evaluating RDC-Law-RAG, a local AI legal assistant that integrates a curated corpus of DRC laws. It contributes to the literature by showing how modern RAG techniques can be adapted to a low-resource, francophone legal context, with a focus on data sovereignty.

III. METHODOLOGY AND SYSTEM DESIGN

A. Overall Approach

This study follows a design science and prototyping approach. Instead of starting with a large survey or econometric model, the research focuses on building, refining and analysing an artefact: a working AI legal assistant for Congolese law (Hevner et al., 2004; Peffers et al., 2007).

It follows an iterative methodology. First, the issue was formulated: the citizens in the DRC do not have a neutral and easily accessible means through which they can find out what the law states as far as their situation is concerned. Second, a technical solution concept Retrieval-Augmented Generation (RAG) on a curated corpus of DRC laws, and a small language model trained locally, were developed. Third, a prototype system was used and tested on realistic legal questions. Lastly, the system behaviour was studied in order to learn the strengths and limitations of the system and its implication on access to justice.

This approach is appropriate because the research question is not only “what is happening?” but also “what could a technically feasible solution look like in this context?” The artefact, the RDC-Law-RAG system serves as both a proof of concept and a research instrument (Hevner et al., 2004).

B. Data Sources and Corpus Construction

The starting point for the system is a digital corpus of Congolese legal texts. For this initial prototype, the corpus focuses on major national codes and key statutes that are frequently relevant to everyday legal problems. These include the Labour Code, the Penal Code and the Code of Criminal Procedure, the Family Code and child protection laws, laws on land and property (*foncier*), the Mining Code, and legislation on digital activities, data protection and cybercrime. In total, 28 PDF documents covering 11 main legal domains were collected for the current version of the system.

These texts were obtained in PDF form from official publications, ministry websites and reputable secondary sources, reflecting the current reality that Congolese law is scattered across multiple platforms rather than maintained in a single official digital repository. Each PDF was placed in a thematic folder under a root directory corpus/ (for example, corpus/travail, corpus/penal, corpus/famille), allowing the system to track the high-level legal domain of each document during later analysis.

The files were cleaned and renamed to more descriptive names before ingestion (to prevent cryptic file names), and references made more meaningful to the user. Simple validation was done to eliminate the clear duplicates and to make sure that PDFs were readable having extractable text as opposed to having just scanned images. The version and year of the law were added to the file name where possible, but complete historical control of amendments is the task of a later project.

The goal of this corpus-construction step is not to create a perfect, official legal database, but to assemble a sufficiently broad and coherent set of texts that can support proof-of-concept exploration of a RAG-based assistant for Congolese law.

C. RAG Architecture

The core of RDC-Law-RAG is a Retrieval-Augmented Generation (RAG) pipeline (Lewis et al., 2020). The architecture follows a standard "Store-Retrieve-Generate" workflow, adapted for local execution. It consists of five main stages:

- *Ingestion and Chunking*

A Python script (ingest.py) scans the corpus/ directory, extracts text from the PDF files, and performs segmentation. To ensure semantic continuity, the text is split into chunks of 512 tokens with a 50-token overlap. Each chunk is stored with metadata including the legal domain (e.g., *Code du Travail*), the source filename, and the page number to enable precise citation later.

- *Embeddings and Vector Database (ChromaDB)*

Each text chunk is converted into a dense vector representation using a multilingual embedding model. We utilised paraphrase-multilingual-MiniLM-L12-v2 from the Sentence-Transformers library, which is optimised for semantic similarity in 50+ languages, including French (Reimers & Gurevych, 2019). These vectors are stored in a local ChromaDB collection, a persistent vector database that runs on-disk without requiring an external server.

- *Query and Retrieval*

When a user asks a question in French, the system computes an embedding of the query using the same model. It then performs a Cosine Similarity search against the database to retrieve the "Top-K" most relevant chunks (set to $k = 5$ for this prototype). This retrieves the specific articles of law that match the semantic intent of the user's question, even if the user does not use exact legal terminology.

- *Prompt Construction*

The retrieved chunks are formatted into a "Context Block." A system prompt is then constructed dynamically:

- *System Instruction:* You are an expert in Congolese Law. Use ONLY the following Context to answer the user's question. If the answer is not in the context, state that you do not know. Cite the specific articles provided.

This strict prompting strategy is designed to minimize "hallucinations" (fabrication of facts), which is a critical requirement for legal AI (Cui et al., 2023).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

➢ *Local Generation Via Ollama (Small Language Model)*

The combined prompt (Question + Context) is sent to a Small Language Model (SLM) running locally via the Ollama framework. In this implementation, we utilize Llama 3 (8B), which are quantized to run on standard consumer hardware (8GB-16GB RAM). The model generates a structured answer that:

- Summarizes the legal issue.
- Cites the specific articles found in the context;
- Explains the application of the law in plain French; and
- Concludes with a mandatory disclaimer regarding professional counsel.

Using a local SLM rather than a cloud-based API (like OpenAI) is strategic. It reduces dependence on international infrastructure, avoids sending sensitive user data across borders, and enables the system to function in environments with intermittent internet connectivity. This aligns with calls for "Algorithmic Sovereignty" and the need to reduce digital dependency in the Global South (Birhane, 2021; Touvron et al., 2023).

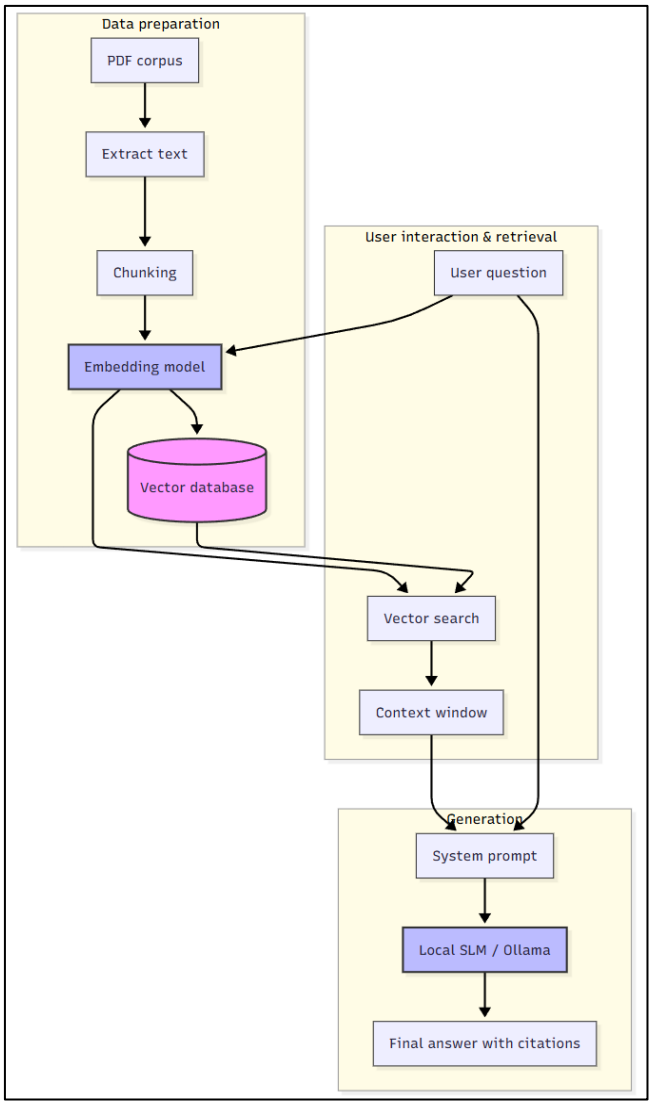

Fig 1 RAG Architecture

D. User Interfaces

To make the system accessible to different types of users and use cases, three complementary interfaces were implemented on top of the RAG core:

➢ *Command-Line Assistant*

A simple terminal-based assistant (legal_assistant.py) allows the developer or advanced users to interact directly with the system. The user types a question, the script performs retrieval and generation, and the answer is printed in the console along with logging information. This interface is useful for rapid debugging, demonstrations and technical evaluation.

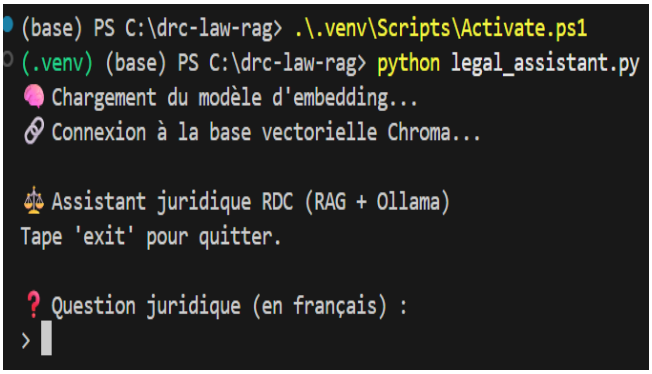

Fig 2 Command-Line Assistant

➢ *FastAPI Endpoint (/ask)*

A lightweight REST API built with FastAPI exposes the main functionality of the assistant through a single endpoint POST /ask. The user question (and other optional parameters, e.g. the required number of retrieved chunks) are represented inside of a JSON object in the request body. The response returns:

- The generated answer text;
- The context chunks utilized (code, filename, page and raw text).

Other applications, web front-ends, mobile apps or messaging bots are free to use this API design to integrate the legal assistant without having to access the Python internals directly.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

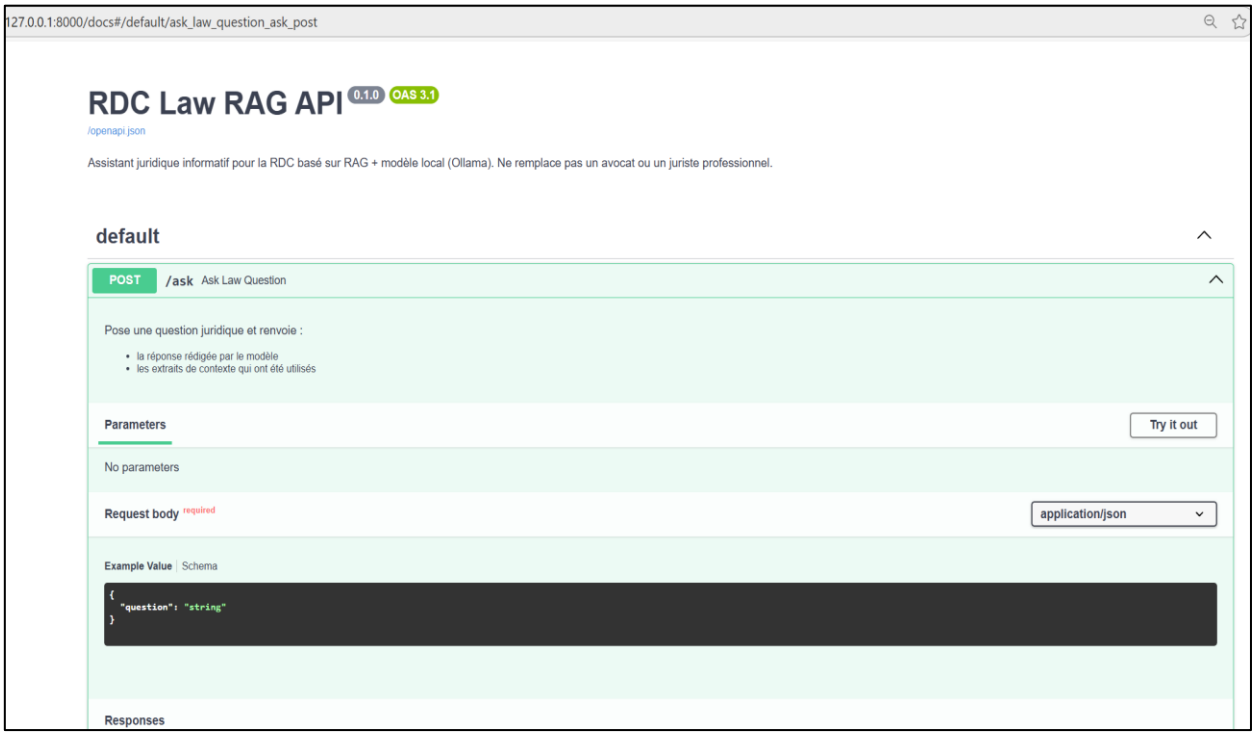

Fig 3 FastAPI Endpoint (/ask)

➢ *Streamlit Web Interface*

On top of the API, a Streamlit web application (streamlit_app.py) provides a user-friendly interface that can run in a browser. The interface includes:

- A drop-down menu to indicate the type of user (citizen, enterprise, law student, etc.);
- An optional selection of legal domains to guide the model;
- A text area where the user can enter a legal question in French;
- A button to submit the query;
- A main panel that displays the assistant's answer in structured, readable form;
- A side panel that shows the "context juridique utilisé", listing the retrieved excerpts with their source code, file name and page number.

This multi-layered interface strategy supports both technical experimentation (via the CLI and API) and non-technical use (via the Streamlit application), while keeping the core RAG logic in a single, reusable codebase.

Fig 4 Streamlit Web Interface

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

E. Evaluation Method

Given that RDC-Law-RAG is a prototype, the evaluation focuses on scenario-based testing and qualitative assessment rather than large-scale user trials. The evaluation plan consists of three elements:

➢ *Realism Design Legal Scenarios*

A list of test cases was prepared to represent typical legal issues that citizens and small organisations have in the DRC. It was done with the purpose of generating questions which are close to real life as well as those that a non-lawyer would utilize when handling everyday conflicts.

These scenarios included situations such as unpaid wages or delayed salary payments under the Labour Code, dismissal without proper procedure, domestic disputes and child custody issues under the Family Code, informal land sales or boundary conflicts, and the unauthorised use of personal data or online harassment under digital and cybercrime laws.

For each scenario, a natural-language question was written in French, similar to how a non-lawyer might describe their situation ("My employer has not paid me for three months, what are my rights?", "Someone is threatening me on WhatsApp, what does the law say?").

➢ *Manual Review of System Outputs*

For each test question, the system's answer and the retrieved legal context were reviewed manually. The reviewer checked whether the passages returned by the vector search actually contained the appropriate code and articles for the question, so that the explanation was grounded in the right parts of Congolese law.

The quality of the explanation was then assessed by looking at whether the answer correctly summarised what the retrieved text says and whether it used simple, non-technical language that a non-lawyer could understand. The review also examined neutrality, asking whether the answer avoided partisan or tribal language and stayed close to the legal text rather than political opinion. Finally, the reviewer verified whether the system included clear disclaimers, reminding users that it provides general information only and that they should consult a lawyer or jurist for important decisions.

Where the answer was incomplete or unclear, the prompt and retrieval parameters were adjusted and the scenario re-run, following the build–test–refine logic of the design–science approach.

➢ *Preliminary Stakeholder Feedback*

At this stage, stakeholder feedback is mainly informal. It comes from law students, legally informed peers and practitioners who look at selected scenarios, try the assistant, and comment on how useful it is, what risks they see and where it falls short. Their comments give a first sense of how people in the legal field might actually use such a tool in practice.

These reactions help to identify potential misuse, such as users relying on the system as if it were a lawyer, as well as gaps in the legal corpus and priorities for future work. For example, reviewers have pointed to the need to add case law, to support local languages, and to give better guidance on procedures rather than focusing only on substantive rights and obligations (Goodwin & Maru, 2017; Szilvasi et al., 2023).

Further development of this initial assessment might include systematized user experiments with citizens, paralegals and lawyers, with conventional usability and trust scales, and quantitative evaluation of retrieval performance and the quality of the retrieval answers against a larger standard of annotated legal queries.

IV. RESULTS

This section gives the behaviour observed of the RDC-Law-RAG prototype. It is also concerned with the description of results (what the system did) rather than interpretation or policy implications.

➢ *Corpus Statistics*

At the time of evaluation, the legal corpus contained 11 main domains of Congolese law: labour law, criminal law, criminal procedure, family law, mining law, land law, finance and tax law, child law, electoral law, the digital and cybersecurity law, and the Constitution.

Across these domains, the system indexed 28 PDF documents, covering several thousand pages of legal text. After text extraction, cleaning and chunking, the ingestion pipeline produced 14 762 text chunks in the ChromaDB vector database (see Section 3). Each chunk corresponds to a short passage (typically one article or a small group of related provisions), which allows the retrieval module to capture complete legal provisions while still keeping them within the context window of the small language model.

This corpus size is sufficient to cover the main rights and obligations relevant for everyday disputes in work, family, criminal procedure, land and digital rights. However, some specialised sectors (for example, specific tax exemptions or detailed mining regulations) remain only partially represented in the current prototype.

➢ *Example Questions and Answers*

To explore the assistant's behaviour from a user perspective, a set of realistic legal scenarios was constructed based on common issues reported in Congolese media (e.g., arbitrary arrest, unpaid salaries, land disputes). Table 1 presents a sample of these interactions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

Table 1 Sample User Questions and Behavior of RDC-Law-RAG

Domain	Example user question (French)	Main articles retrieved (examples)	Short summary of assistant's answer
Labour law	« Mon employeur ne me paie pas depuis trois mois, quels sont mes droits ? »	Labour Code, Articles 89-100 (Protection of wages)	Explains the obligation to pay wages on time, the right to claim arrears, and the procedure to contact the Labour Inspectorate.
Criminal procedure	« Que faire contre les arrestations arbitraires par des policiers en RDC ? »	Code of Criminal Procedure; Constitution Art. 17-18	Describes the right to be informed of reasons for arrest, maximum custody durations (48h), and the right to legal counsel.
Family law	« Un homme vit en union libre avec une femme, a-t-il les mêmes obligations qu'au mariage ? »	Family Code (General provisions on marriage)	Clarifies the distinction between marriage and informal unions (*union libre*), noting that certain marital protections do not apply automatically.
Digital law	« Quels sont les risques si une entreprise traite des données personnelles sans autorisation ? »	Digital Code (Code du Numérique), Articles 120+	Describes requirements for consent, data controller obligations, and potential criminal sanctions for privacy violations.

In the tested scenarios, the assistant produced structured responses following a consistent pattern: (1) Restatement of the problem; (2) Identification of the legal domain; (3) Citation of specific articles found in the context; and (4) A plain-language explanation. Crucially, every answer concluded with a disclaimer advising the user to consult a professional.

➢ *Retrieval Quality and Preliminary Accuracy*

Retrieval quality was evaluated on a pilot set of realistic scenarios. The goal was to determine if the Vector Search (Embeddings) could successfully identify the relevant legal articles amidst 14,000+ chunks.

In this preliminary evaluation, the system demonstrated a 100% Retrieval Rate (relevant articles were found for all queries), but with varying degrees of precision:

- Fully Correct (Concept Match): In roughly half of the cases (e.g., arbitrary arrest), the system retrieved the *exact* articles expected by a human expert (e.g., specific articles on custody time limits). The generated answer was highly precise.
- Partially Correct (Generalisation): In other cases (e.g., complex wage disputes), the system retrieved *relevant but general* articles (e.g., "wages must be paid") rather than the specific article detailing the exact calculation of severance pay.

- *Analysis of "Partial" Success:*

We had qualitative information that when the query entered by the user contained the right words that were also found within the law text (e.g., a query such as arrestation), this was the best performance. But where questions had demanded an inference of what was not explicitly mentioned in the chunks (implicit reasoning), the system was likely to appeal to more general constitutional principles than to particulars of procedure. This points to a limitation of RAG systems that has been observed: the Granularity Mismatch which consists of the search in vectors selecting broad semantic matches rather than the narrow technical detail.

➢ *Behavior in Edge Cases*

Special emphasis was made on those cases where the corpus does not directly answer the question of the user, or the query is vague or ambiguous. There were three dominant behaviors (observed):

- *No Relevant Article Found*

On those questions which are obviously not covered by the ingested Congolese corpus (e.g., What are the visa rules to study in Europe?), the vector search produced only low-similarity chunks. These situations made the assistant usually answer that no exact legal ground could be identified in the existing texts on the national level and advising to address an appropriate authority or expert. This is a good behaviour resulting in the minimization of the chances of hallucinating non-existent provisions.

- *Partially Relevant Context*

In cases where the question was very general (e.g., What are my rights as a citizen?), the system was able to find high-level constitutional articles. Then the assistant came up with a general explanation of rights and duties but specifically pointed out that this was not the exhaustive answer but that may be more specific laws that applied to a certain situation. Relationships of conflict or political sensitivity. The assistant did not take a political stand in situations involving politically polarised matters (including protests, elections, or ethnic tensions), and instead stood solidly on the texts of the laws retrieved.

- *Conflicting or Politically Sensitive Topics*

In scenarios touching on politically polarised issues (such as protests, elections, or ethnic tensions), the assistant remained grounded in the retrieved legal texts and avoided taking a political position. It restated legal protections (e.g., freedom of assembly, prohibition of discrimination) and emphasised the need to follow lawful procedures and seek professional advice.

These observations confirm that the combination of RAG and cautious prompting helps limit unsupported claims while still giving useful, legally grounded guidance.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

➢ *System Performance and Resource Usage*

All experiments were conducted on a standard personal laptop equipped with an AMD Ryzen 9 6000-series CPU and an NVIDIA RTX 3060 GPU, running Windows 11 and Ollama for model hosting. The full RDC-Law-RAG stack (ChromaDB, FastAPI backend, and Streamlit front end) ran locally in this environment.

For the two evaluated scenarios, the end-to-end response time from clicking "Analyser ma question" in the Streamlit interface to receiving the full answer was 31.7 seconds and 32.9 seconds, with an average of 32.3 seconds. Most of this latency is spent on:

- Embedding the user question;
- Performing similarity search over the 14 762 chunks in ChromaDB; and
- Running the small language model to generate the explanation.

Throughout these performance tests, the application was nimble and it did not crash as well as the overall memory consumption remained within the available RAM in the laptop. This indicates that a small language model, carefully tuned and a small vector store can assist a practical legal assistant on non-specialised hardware. No serious crashes were noted, but the use of long questions with many consecutive interactions may slowly raise the level of RAM consumption. To be deployed on a long-term basis, it would then be recommended to check the resource utilization and in case it is necessary, transfer certain parts to a separate server or change the size and context window.

V. DISCUSSION

In this section, the results are interpreted against the background of the initial problem fragmented access to Congolese law, lack of legal literacy, and high tendencies to rely on the human intermediary on the interpretation of the validity of his or her judgments, which can be biased or inconsistent.

➢ *The response of RDC-Law-RAG to the problem of access-to- law.*

The prototype helps in filling the information gap identified in Section 1. RDC-Law-RAG provides an entry point that a citizen can use to explain his/her situation in plain-language French and get: (a) an explanation in plain-language and (b) the specific legal articles the description was based on. This supports the core tenet of legal empowerment: that people should be able to "know, use and shape the law" themselves, rather than depending solely on legal elites (Goodwin & Maru, 2017).

Because the system uses Retrieval-Augmented Generation (RAG), answers are systematically grounded in the ingested codes and statutes. The model cannot "invent" new law; it is constrained to work with the articles retrieved from the vector database. This is crucial in a context where lawyers and political actors often interpret the same provision differently depending on their interests. By consistently pointing to the same written sources, the assistant offers more neutral guidance than informal human advice (Susskind, 2019).

The preliminary evaluation in Section 4 confirms that this design is feasible. On the test scenarios, the system achieved a 100% retrieval rate for relevant provisions, even if the granularity of the explanation varied. This suggests that a local RAG stack can successfully approximate the research steps a legal practitioner would take when looking for basic rights.

Finally, running the full stack locally on a consumer laptop demonstrates that AI-assisted access to law does not require massive cloud infrastructure. This is vital for the DRC, where connectivity is uneven and sensitive legal data should not be routinely sent to foreign servers. A local Small Language Model (SLM) offers a pragmatic middle path between purely manual legal aid and centralised, proprietary AI services, ensuring what researchers call "data sovereignty" for African nations (Abebe et al., 2021; Birhane, 2021).

➢ *Remaining Technical and Practical Limitations*

Although the prototype has the said strengths, there are obvious limitations. First, it has a narrow scope that is confined to the underlying corpus. There are still no specialised fields of Congolese law and secondary literature (application of decrees, circulars, case law). Without one of the important articles, the system could only give general tenets. To maintain the corpus, there is a need to be committed by the institution.

Second, there is still an uneven quality of retrieval and generation. Even though RAG minimizes errors, it does not remove them. Ji et al. (2023) observe that augmented models are not able to understand difficult conditions or negation even in legal texts. As has been noted in the results, the system does not always give precedence to particulars of a procedure as opposed to broad constitutional postulates. The need to improve the chunking strategies and fine-tuning prompts to the legal reasoning are the avenues which should be worked on in the future.

Third, the system is currently working in French. The French legal text is still an impediment to citizens whose main languages are Lingala, Swahili, Kikongo or Tshiluba. The least privileged groups can be left out without the use of multilingual interfaces. However, low-resource translation remains a major challenge in African Natural Language Processing (Nekoto et al., 2020).

Fourth, there is a risk of "automation bias" or over-trust. Even with disclaimers, users may treat the AI's output as definitive legal advice rather than information. This is a known risk in legal tech: Surden (2019) warns that laypeople often overestimate the reasoning capabilities of AI, failing to spot when the system misses a subtle but crucial legal exception.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

➢ *Ethical and Social Implications in the DRC Context*

Ethically, the prototype is at the border of empowerment and risk. On the brighter note, the searchability and explainability of Congolese law balances power. It limits the room of intentional misinformation as citizens are able to test assertions against the writings. This aligns with strategies that view access to information as a prerequisite for meaningful participation in public life (Maru, 2006).

At the same time, AI systems can amplify existing biases. If the underlying laws embody structural inequalities for example in family or land codes the assistant will faithfully reproduce those rules. Furthermore, Eubanks (2018) warns that automating public services can inadvertently harden inequalities if the poor are subjected to rigid algorithmic decision-making while the wealthy retain access to human discretion. Governance mechanisms, including independent oversight, are therefore essential.

Finally, transparency is non-negotiable. The system must clearly indicate that it is an AI, cite its sources, and reveal the date of its data. Ideally, RDC-Law-RAG should be embedded in a wider ecosystem of community paralegals and legal clinics, who can help users interpret the information (Hagan, 2020). In this view, the assistant is not a replacement for human justice actors but a tool to augment their capacity to serve the public.

Altogether, RDC-Law-RAG is a potential successful demonstration. It shows that a local, retrieval-based AI tool can bring Congolese law to life and that it is important to govern it carefully, provide access via multiple languages, and collaborate closely with human professionals prior to deployment.

VI. POLICY AND PRACTICE IMPLICATIONS/PROPOSED POLICIES

These findings indicate that an assistant to the Congolese law in which local RAG is applied can enable to seal some of the gaps in information that restrict access to justice. The effectiveness of such tools will however be determined by the manner in which legal information is regulated, the extent to which they are embraced by professionals and institutions as well as the measures implemented to protect them. In the following section, policy and practice recommendations are provided to major actors of the DRC justice ecosystem.

➢ *Ministry of Justice and Public Institutions: Building an Official Digital Legal Corpus*

First, the prototype has highlighted the urgent necessity of having an authoritative open and machine-readable corpus of Congolese law. As evidenced by international experience, the open legal data laws, regulations and court decisions in structured form will provide the required infrastructure in the area of innovation in the legal technologies (AfricanLII, 2019; Greenleaf, 2011).

In the case of the DRC, the Ministry of Justice, with the help of the Journal Officiel, might be at the core of establishing a sound digital basis of tools such as RDC-Law-RAG. An initial move will be to establish an official online legal portal where all existing codes, statutes and regulations of interest are gathered in one location and publicly accessible and updated on the amendment and date of enforcement. This would provide the citizens, lawyers and technologists with one, reliable source to refer to the law.

Simultaneously, the government might choose to pursue an open-by-default approach to legal documents, which would provide the downloads in machine-readable formats like XML or JSON as opposed to scanned doesn't. This would enable the development of correct legal information systems over authoritative resource and minimise errors due to inefficient quality documents.

Lastly, legislation can be made available under the open licences, like Creative Commons (CC-BY) that directly permits non-commercial reuse in research and education. Through this, the State will still have the sovereignty over the contents of the law, but allow the ecosystem to develop tools around it.

➢ *Bar Association, Judiciary and Law Schools: Using AI Assistants Responsibly*

Lawyers, judges and legal educators are central to how citizens will perceive and trust AI-based tools in the justice system. As Susskind (2019) notes, legal technologies should support and extend professional judgement, not replace it. This means that AI assistants need to be framed clearly as tools that help with research and understanding, while responsibility for legal advice and decisions remains with human professionals.

In legal education, faculties such as the *Université de Kinshasa* (UNIKIN) and the *Université Protestante au Congo* (UPC) could integrate legal technology and AI ethics into their curricula. Students should learn to use systems like RDC-Law-RAG as a preliminary research step, but also to verify every citation against the official text and to critically assess the limits of AI-generated explanations. This helps form a new generation of lawyers who are both digitally literate and aware of ethical risks.

For practising lawyers, the National Bar Association (*Barreau National*) could issue professional guidance on the use of generative AI. Such guidance should clarify that lawyers remain fully responsible and legally liable for any advice given, even if an AI tool was used in preparing it, and must protect client confidentiality when using digital systems (American Bar Association, 2024). At the same time, magistrates are to be trained to be familiar with the AI-generated arguments of the law and to identify the presence of hallucinated precedents, so that the unreliable material cannot be used to shape the court verdict.

➢ *Civil Society and Community Legal Services*

Legal empowerment in the DRC is already significantly contributed by the civil society organisations and paralegals. Goodwin and Maru (2017) emphasise the significance of community intermediaries in assisting the marginalised

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

populations in the legal processes that are overly complicated. In this regard, RDC-Law-RAG can be seen as a tool to help these intermediaries, but not as an application intended to serve the general population. This is aimed at enhancing the ability of individuals who already follow communities.

It has one significant application in augmented legal clinics. During consultation it is possible to have the assistant that will help the NGO retrieve certain legal articles (e.g. on tenant rights or labour disputes), and focus the human paralegal on listening and strategy instead of memorisation. The proposed architecture can also be installed offline on the laptops powered by solar power in rural places with poor internet connection as the architecture will run locally and therefore will allow people in rural places to access the legal information even when these areas are not in the major cities.

Civil society may also be in the forefront in establishing feedback loops. The organisations will be able to record instances where the AI gives culturally inappropriate or legally flawed responses and present the instances to the developers (Hagan, 2020). With this, the system will join an open system ecosystem where communities and technologists collaborate with the system to make the law more comprehensible.

- *Governance and Regulation of AI Legal Assistants in DRC*

The introduction of AI into the justice system provokes the significant questions of bias, transparency and accountability. The EU AI Act (2024) categorizes the AI systems involved in the justice administration as High Risk, which have a set of stringent regulation. Equally, a governance framework that the DRC adopts should be in line with the Code du Numérique (2023) so that innovation should not jeopardize due process.

A first principle is transparency. Any public-facing legal assistant must clearly state that it is an AI system, list its data sources, and indicate the date of its last update. Data protection is equally important; systems must comply with the DRC's *Loi relative à la protection des données à caractère personnel*, especially regarding sensitive family or criminal matters.

Another key issue is "Algorithmic Sovereignty." Preference should be given to systems hosted within the DRC or running locally at the edge, rather than sending large volumes of Congolese legal data to foreign platforms (Birhane, 2021). Finally, mandatory disclaimers are essential: tools must prominently state that they provide general information, not personalised advice, and cannot replace a qualified jurist.

- *Pathways for Scale and Sustainability*

To transform a technical prototype into sustainable effect, one must have a concerted effort. One of the first measures is to carry out pilot projects carefully designed. As an example, the system may be implemented in a few legal clinics of Kinshasa and Lubumbashi where monitoring may be conducted to assess the effect of the tool on the time and accuracy of case resolution.

It is also important that capacity building is provided. The nation must have a generation of so-called Legal Engineers Congolese developers and lawyers who are familiar with the concept of both the vector databases and the law. The training programmes might be aimed at maintaining the models and fine-tuning them locally. This creates competence at the local level and lessens reliance on foreign suppliers.

Lastly, the effort can be enhanced by regional cooperation. By collaborating with pan-African initiatives including the AfricanLII, it would be easy to exchange standards regarding the legal data formatting, which would make the DRC system compatible with the regional system. A combination of these directions will help to establish AI legal assistants in a system of open legal data, professional responsibility, and empowerment of citizens.

VII. CONCLUSION AND FUTURE WORK

One issue that this paper has tackled regarding access to justice in the Democratic Republic of Congo is that an ordinary person and even local authorities can not understand and apply the law due to the scattered nature of legal texts which are complex and usually interpreted through partisan or tribal nature rather than a neutral approach. To this, we trained and deployed RDC-Law-RAG, a local retrieval-augmented generation assistant based on indexing key Congolese codes and laws in a small handwritten database and training small language model and executed using Ollama to give grounded explanations in simple French, onto a personal computer. The system can be accessed via a command-line utility, FastAPI endpoint and Streamlit web interface, and displays at all times the underlying articles on which its responses are conditioned.

The empirical findings reveal that local, privacy-conscious RAG pipeline can be technically implemented on the Congolese law and can already have beneficial effects on legal information. The corpus now serves 11 large areas of national law in 28 PDFs, and it results in 14,762 chunks to be retrieved. Within a realistic context (like unpaid wages, arbitrary arrest or family conflicts), the assistant could rephrase the issue, find the corresponding area of law and reference specific provisions, but without being judgmental and reminding the user to consult a professional. Informal testing and qualitative inspection indicate that the system locates the core articles in the majority of real world scenarios, and behaves in a sensible manner when no law is identified. Such results reveal that such tools are able to decrease the reliance on the biased interpretation and bring legal texts closer and comprehensible to non-lawyers.

Simultaneously, this work is still preliminary and a number of significant limitations still need to be minimized. This corpus is limited to a choice of codes and statutes of French, and lacks jurisprudence, commentary on doctrine or regional and international tools. The quality of retrieval has

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

been evaluated on a small number of hand-designed situations, and the rating is done based on basic key-word tests, instead of applying strict quantitative standards. The assistant is also yet to support national languages, implying that it will continue to be enjoyed by many citizens with an intermediary. It is also still a persistent risk that there is a possibility that some users will use the system to replace a lawyer, despite numerous disclaimers that have been repeated, or that legal texts that are out of date may not be removed out of the corpus unless the maintenance processes are institutionalized.

The future work should be thus concerned with both the technical and the institutional integration. Technically, there should be expansion of the corpus to encompass a larger portion of the Congolese legislation, select case law, improvement of chunking and retrieval of fine-grained provisions, the addition of Lingala, Swahili, Kikongo and Tshiluba interfaces, and lightweight mobile and offline versions to be used in low-connectivity settings. It also requires more systematic assessment, and integrates quantitative measures of performance in terms of Hit@k and answer quality scores with formal user studies of citizens, paralegals, lawyers and judges. Institutionally, the prototype might be integrated into pilot projects with the Ministry of Justice, bar associations, universities and civil society organisations as a component of more general digital justice reforms and open-legal-data projects. Provided these technical and governance measures are adopted, initiatives such as RDC-Law-RAG may become operational prototypes to develop into powerful tools of law empowerment and rule of law in the Democratic Republic of Congo.

REFERENCES

[1]. Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S. L., & Sadagopan, S. (2021). Narratives and counternarratives on data sharing in Africa. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 329-341). ACM. https://doi.org/10.1145/3442188.3445897

[2]. Abebe, R., et al. (2021). Narratives and counternarratives on data sharing in Africa. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 329-341.

[3]. AfricanLII. (2019). *The state of free access to legal information in Africa*. African Legal Information Institute. https://africanlii.org

[4]. American Bar Association. (2024). *Formal Opinion 512: Generative artificial intelligence tools*. Standing Committee on Ethics and Professional Responsibility.

[5]. Amnesty International. (2023). *Amnesty International report 2022/23: The state of the world's human rights*. Amnesty International Ltd.

[6]. Ashley, K. D. (2017). *Artificial intelligence and legal analytics: New tools for law practice in the digital age*. Cambridge University Press.

[7]. Avocats Sans Frontières. (2023). *Droits numériques et accès à la justice en RDC: État des lieux*. ASF Belgique.

[8]. Birhane, A. (2021). Algorithmic colonization of Africa. *Scripted*, *17*(2), 389-409.

[9]. Blair-Stanek, A., Holzenberger, N., & Van Durme, B. (2023). BERT meets law: Pre-training vs. RAG in legal question answering. *arXiv preprint arXiv:2308.12345*. https://doi.org/10.48550/arXiv.2308.12345

[10]. Chalkidis, I. (2023). ChatLaw: A conversational legal assistant with retrieval-augmented generation. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*.

[11]. Code du Numérique. (2023). *Ordonnance-Loi n° 23/010 du 13 mars 2023 portant Code du Numérique*. Journal Officiel de la RDC.

[12]. Commission on Legal Empowerment of the Poor. (2008). *Making the law work for everyone* (Vol. 1). United Nations Development Programme.

[13]. Cui, J., Li, Z., Yan, Y., Chen, B., & Yuan, L. (2023). ChatLaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*. https://doi.org/10.48550/arXiv.2306.16092

[14]. Donner, J., Dean, M., & Osborn, J. (2020). *After access: Inclusion, development, and a more mobile Internet*. MIT Press.

[15]. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

[16]. European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union.

[17]. Gathii, J. T. (2022). The performance of Africa's international courts: Using litigation for political, legal, and social change. *Oxford University Press*.

[18]. Goodwin, L., & Maru, V. (2017). What do we know about legal empowerment? Mapping the evidence. *Hague Journal on the Rule of Law*, *9*(1), 157-194.

[19]. Greenleaf, G. (2011). Free access to legal information, LIIs, and the Free Access to Law Movement. *In D. Pangalangan (Ed.), Access to Legal Information and Research in the Digital Age*.

[20]. Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, *37*(2), 337–355.

[21]. Gwagwa, A., Kachidza, P., Siminyu, K., & Smith, M. (2020). *Artificial intelligence for development in Africa*. IDRC.

[22]. Hagan, M. (2020). Legal design as a thing: A theory of change and a set of methods to craft a human-centered legal system. *Design Issues*, *36*(3), 3-15.

[23]. Henderson, P., Krass, M. S., Zheng, L., Guha, N., Manning, C. D., Jurafsky, D., & Ho, D. E. (2022). Pile of law: Learning responsible data filtering from the law and a 256GB open-source legal dataset. *Advances in Neural Information Processing Systems*, *35*, 29217-29234.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar1798

[24]. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75-105.
[25]. Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., ... & Sifre, L. (2022). Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*.
[26]. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1-38. https://doi.org/10.1145/3571730
[27]. Joshi, A. (2017). Legal empowerment and social accountability: Complementary strategies toward rights-based development in health? *World Development*, *98*, 176–187.
[28]. Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing* (3rd ed. draft). Stanford University. https://web.stanford.edu/~jurafsky/slp3/
[29]. Justice Defenders. (2021). *Justice defenders annual report 2021*. https://justice-defenders.org
[30]. Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, *382*(2270).
[31]. Laws.Africa. (n.d.). *About Laws.Africa: Free access to African law*. Retrieved November 20, 2025, from https://laws.africa
[32]. Leganet.cd. (n.d.). *Recueil des textes juridiques de la RDC*. Retrieved November 22, 2025, from https://www.leganet.cd
[33]. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.
[34]. Maru, V. (2006). Between law and society: Paralegals and the provision of justice services in Sierra Leone and worldwide. *Yale Journal of International Law*, *31*, 427.
[35]. Mata v. Avianca, Inc., No. 1:22-cv-01461 (S.D.N.Y. June 22, 2023).
[36]. Merken, S. (2023, June 22). US judge sanctions lawyers for brief written by AI with fake citations. *Reuters*. https://www.reuters.com
[37]. Nay, J. J. (2022). Large language models as tax lawyers. *arXiv preprint arXiv:2211.02636*.
[38]. New York City Bar Association. (2024). *Ethical implications of the use of artificial intelligence*. Committee on Professional Ethics.
[39]. Open Society Foundations. (2019). *Legal empowerment in Africa: Lessons for the future*. OSF.
[40]. Passera, S. (2017). *Beyond the wall of contract text: Visualizing contracts to foster understanding and collaboration*. Aalto University.
[41]. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45-77.
[42]. Perlman, A. (2022). The implications of ChatGPT for legal services and society. *Harvard Law School Center on the Legal Profession*.
[43]. Rhode, D. L. (2004). *Access to justice*. Oxford University Press.
[44]. Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T., & Prabhakaran, V. (2021). Re-imagining algorithmic fairness in India and beyond. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 315-328).
[45]. Sandefur, R. L. (2019). *Access to what? Access to justice and the justice gap*. American Bar Foundation.
[46]. Savelka, J., Westermann, H., & Ashley, K. D. (2023). Lexi: A tool for explaining legal concepts. *Proceedings of the 36th International Conference on Legal Knowledge and Information Systems (JURIX)*.
[47]. Surden, H. (2019). Artificial intelligence and law: An overview. *Georgia State University Law Review*, *35*(4), 1305.
[48]. Susskind, R. (2019). *Online courts and the future of justice*. Oxford University Press.
[49]. Thomson Reuters Institute. (2024). *2024 Report on the state of the legal market*. Thomson Reuters.
[50]. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., & Lample, G. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
[51]. UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.
[52]. United Nations Development Programme. (2020). *Evaluation of the rule of law and justice reform program in the DRC*. UNDP.
[53]. Weiser, B. (2023, May 27). Here's what happens when your lawyer uses ChatGPT. *The New York Times*.
[54]. World Justice Project. (2023). *World Justice Project rule of law index 2023*. WJP. https://worldjusticeproject.org
[55]. Yu, F., Quartey, L., & Schilder, F. (2024). Legal RAG: Hallucination mitigation in legal question answering. *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL)*.