

Guardian AI

Elijah Cantu
Ellen Chlachidze
Mira Nauli Panjaitan
Somin Park
Benjamin Yee

Abstract

The U.S. foster care system operates within a complex legal framework of constitutional, federal, and state laws, making it difficult for professionals and families to navigate. Caseworkers and legal professionals are often overburdened, leading to errors that negatively impact children and families. This project proposes the development of GuardianAI, an AI-powered legal assistance tool designed to provide accurate and accessible legal guidance. The system will aggregate legal sources, procedural guidelines, and public resources to answer complex legal questions for professionals and offer plain-language explanations for families. Through data-driven AI development and usability testing, the tool aims to improve decision-making, reduce legal errors, and enhance accessibility. The project's success will be evaluated based on stakeholder feedback, error rate tracking, accessibility metrics, and overall impact on reducing delays and improving outcomes within the foster care system.

1 Introduction

The U.S. foster care system plays a crucial role in ensuring the well-being of vulnerable children, yet it remains one of the most legally complex and challenging areas of social services. Governed by a combination of constitutional, federal, and state laws, the system requires caseworkers, legal professionals, and families to navigate intricate legal procedures, often without adequate support. Overburdened professionals face high caseloads and limited resources, leading to delays, errors, and inequitable outcomes that disproportionately affect children and families.

To address these challenges, this project proposes GuardianAI, an AI-powered legal assistance tool designed to provide accurate, timely, and accessible legal guidance. By leveraging natural language processing (NLP) and curated legal data, GuardianAI will assist professionals in answering

complex legal and procedural questions while offering plain-language explanations to families. This tool aims to streamline decision-making, reduce errors, and enhance accessibility, ultimately improving outcomes within the foster care system. Through comprehensive data integration, AI model development, and rigorous evaluation, this project seeks to bridge the gap in legal accessibility and foster a more transparent, efficient, and equitable child welfare system.

2 Problem

The U.S. foster care system operates under a complex legal framework of constitutional, federal, state, and administrative laws, creating challenges for professionals and families. Overburdened caseworkers, legal professionals, and judges face heavy workloads, limited resources, and evolving legal requirements, leading to errors, delays, and inconsistent decisions that harm children and families. Families often lack legal literacy or resources to navigate the system, resulting in uninformed decisions, prolonged cases, and inequitable outcomes. The absence of accessible legal guidance fuels mistrust in the system.

To address these issues, GuardianAI proposes an AI-driven legal support tool that provides real-time, plain-language guidance. By integrating legal data and AI decision support, the tool aims to reduce errors, improve efficiency, and empower stakeholders to make informed, legally sound decisions, ultimately enhancing outcomes for children and families.

3 Impact

GuardianAI aims to improve legal accessibility, reduce errors, and empower both professionals and families in the foster care system. By providing AI-driven legal support, it will help caseworkers and legal professionals interpret complex laws efficiently,

reducing workload and minimizing procedural delays. For families, it will offer clear, plain-language guidance, enabling informed decision-making and reducing reliance on costly legal aid.

System-wide, GuardianAI will enhance consistency in legal processes, improve case outcomes, and build trust in the foster care system. By tracking legal inquiries and user feedback, it can also inform policy improvements. Through improved efficiency and accessibility, GuardianAI has the potential to create lasting, positive change in child welfare.

4 Proposed Solution

To address the challenges in the foster care system, GuardianAI will provide an AI-powered legal assistance tool designed to deliver accurate, accessible, and timely legal guidance. This tool will serve both professionals—including caseworkers, legal advocates, and judges—and families navigating the foster care system.

GuardianAI will integrate federal and state statutes, case law, administrative regulations, and procedural guidelines into a centralized AI-driven platform. Using natural language processing (NLP), it will interpret complex legal questions and provide clear, actionable answers. For professionals, the tool will offer precise legal references to aid decision-making, while for families, it will generate simplified explanations of rights and procedures in plain language.

The system will be deployed as a web-based tool to ensure accessibility for diverse users. Additionally, usability testing and stakeholder feedback will refine its clarity, functionality, and accuracy. By streamlining access to legal information, GuardianAI will reduce errors, improve efficiency, and empower all stakeholders in the foster care system.

5 Related Work

The application of large language models (LLMs) in the legal domain has gained significant attention, particularly in providing legal assistance while addressing ethical and technical challenges. Prior research has explored the **responsibility of LLMs in legal advising**, the **collaborative potential of AI in legal consultations**, and the **importance of citation-based verification in legal AI systems**.

Cheong et al. (Cheong et al., 2024) investigate the *challenges of using LLMs for legal advice*, em-

phasizing concerns such as the unauthorized practice of law, liability for incorrect guidance, and the need for AI to help users ask the right questions rather than provide definitive legal answers. Through workshops with legal experts, they propose a four-dimension framework analyzing user attributes, query nature, AI capabilities, and social impacts. Their findings highlight the necessity of responsible LLM deployment, which aligns with our project’s goal of ensuring **ethical and explainable AI** in the foster care legal system.

Sun et al. (Sun et al., 2024) address *multi-agent AI systems for legal consultation*, presenting *LawLuo*, a model designed to **simulate real-world legal interactions**. Unlike single-agent legal consultation systems, *LawLuo* employs multiple specialized AI agents, such as receptionists, lawyers, and secretaries, to improve **multi-turn interactions and legal compliance**. Their work demonstrates that structured AI collaboration enhances response personalization, ambiguity resolution, and adherence to legal procedures. Our proposed foster care legal assistance system could benefit from similar multi-agent architectures to provide **context-aware and role-specific legal guidance**.

Zhang et al. (Zhang et al., 2024) focus on *legal citation accuracy*, introducing *CitaLaw*, a benchmark designed to assess LLMs’ ability to generate legally sound responses with appropriate references. Their study finds that integrating **retrieved legal documents and precedent cases** into LLM-generated responses significantly improves response quality. Additionally, they introduce a **syllogism-based evaluation method** to verify legal alignment between retrieved references and AI-generated answers. This approach is particularly relevant to our project, as ensuring that foster care professionals and families receive **legally grounded, citation-supported guidance** is critical for trust and reliability.

Collectively, these studies provide valuable insights into the development of AI-driven legal assistance tools. Our work builds on these foundations by focusing on **U.S. foster care legal guidance**, ensuring responsible AI deployment, integrating **multi-agent collaboration**, and incorporating **citation-backed responses** to enhance accuracy and usability.

6 Data

6.1 Data Collection Method

To develop our AI-powered legal assistance tool for the U.S. foster care system, we collect data from multiple authoritative sources. These include federal and state statutes relevant to child welfare, administrative policies governing foster care procedures, and landmark constitutional case law impacting parental rights and child custody. Additionally, we aggregate procedural guidelines used by case-workers and judges, as well as publicly available educational materials designed for families navigating the system.

Legal texts are sourced from government repositories such as the U.S. Code, Code of Federal Regulations (CFR), state legislative websites, and judicial opinions available through publicly accessible legal databases. Procedural guidelines are obtained from agencies such as the U.S. Department of Health and Human Services (HHS), the Administration for Children and Families (ACF), and state child welfare departments. Educational materials, FAQs, and advocacy resources are gathered from non-profit organizations and legal aid groups working within the foster care system.

6.2 Data Annotation Method

In developing the AI-powered legal assistance tool, the annotation process is closely tied to key concepts from information retrieval (IR), particularly text processing, vector space models, and text classification. Given the complexity of legal language, the first step involves text preprocessing, which includes tasks such as tokenization, stemming, and lemmatization. These text processing techniques ensure that the legal data is properly prepared for subsequent stages of analysis and retrieval.

The legal texts are then categorized according to specific topics such as parental rights, adoption procedures, and foster care eligibility. These categories are used to build a structured index, which will support efficient querying and retrieval within the vector space model. Each document is represented as a vector, and the terms within the documents are weighted based on frequency and importance, enabling effective retrieval of relevant information.

We annotate key entities, such as statutory citations, case law references, and procedural steps, which are essential for understanding the legal context. This annotation process is aligned with best

practices in text classification, where entities are tagged based on their relevance to specific query types. To ensure accuracy and reliability, we employ manual verification, collaborating with legal experts to refine and validate annotations.

This annotated dataset is used to train information retrieval models, allowing us to apply techniques such as relevance feedback and query expansion to further improve the system's performance. Evaluation of the system is conducted through standard IR metrics such as precision, recall, and F1 score, allowing us to assess the effectiveness of the tool in delivering relevant legal information to users.

7 Methodology

7.1 Implementation

The implementation phase of GuardianAI is focused on developing a lightweight legal assistance prototype designed to provide accessible and reliable legal guidance. The system will integrate a pre-trained model via an API, utilizing advanced options such as OpenAI's GPT-4, Anthropic's Claude, or Google's Gemini. This approach enables rapid development while ensuring that GuardianAI benefits from state-of-the-art natural language processing capabilities without the need for extensive machine learning infrastructure. By leveraging an existing LLM, the system can focus on refining its ability to provide legally relevant responses while maintaining efficiency and scalability.

To ensure that GuardianAI remains relevant and precise, its scope will be limited to addressing a specific legal issue within the foster care system, such as adoption rights, caseworker responsibilities, or the rights of foster children. A curated set of legal documents, frequently asked questions (FAQs), and other verified resources will be used to guide response generation, ensuring that users receive accurate and contextually appropriate legal information. GuardianAI will employ a keyword-based retrieval mechanism to surface relevant legal information from predefined sources. This approach balances accuracy and feasibility by enabling structured and context-aware AI-generated responses while keeping the system manageable within the project's constraints.

The system will be deployed as a simple web-based application, designed for ease of use and accessibility. Technologies such as Flask or Streamlit will be used to create an intuitive, lightweight

interface that allows users to interact with the system seamlessly. The user interface will prioritize clarity, ensuring that both legal professionals and individuals unfamiliar with legal terminology can easily navigate and understand the information provided. Responses will be structured hierarchically, beginning with a direct legal answer, followed by a simplified explanation in layman's terms, and concluding with links to authoritative legal resources for further reading. This structured approach ensures that users receive clear, actionable information while allowing for deeper exploration of legal topics when needed.

Security and ethical considerations are central to GuardianAI's development. AI-generated legal guidance carries inherent risks, as incorrect or misleading responses could have serious consequences for users. To mitigate these risks, every AI-generated response will include a disclaimer stating that the information provided is for informational purposes only and does not constitute legally binding advice. Additionally, a manual oversight mechanism will allow users to flag responses that appear unclear, incomplete, or potentially inaccurate, ensuring continuous refinement of the system. Ethical considerations will remain a guiding principle throughout development, with a strong emphasis on transparency, accountability, and user trust.

7.2 Evaluation

The evaluation phase will focus on assessing GuardianAI's accuracy, usability, and ethical integrity. Given the constraints of the project timeline, this assessment will rely on manual validation, small-scale usability testing, and performance monitoring rather than large-scale automated verification. Legal accuracy will be a top priority, with AI-generated responses being manually reviewed against authoritative legal sources, including legal codes, government guidelines, and verified legal databases. Responses will be evaluated for factual correctness, clarity, and consistency with established legal interpretations. Any incorrect or misleading answers will be flagged for refinement, ensuring that the system continues to improve over time.

Usability testing will be conducted with a small group of participants, including students, caseworkers, and legal professionals, to evaluate how effectively the system meets user needs. Participants will interact with GuardianAI by submitting real-

istic legal inquiries and providing feedback on the clarity, ease of use, and perceived reliability of the responses. A short survey will be used to gather structured feedback, while qualitative insights will help identify areas for improvement. Ensuring that GuardianAI is user-friendly and accessible is critical to its success, as legal information must be both accurate and easy to understand for a wide range of users.

Performance monitoring will further support GuardianAI's evaluation by tracking key system metrics, including response time, error handling, and overall stability. The system's responsiveness will be assessed to ensure that users receive timely answers, while error handling mechanisms will be tested to verify system reliability under different conditions. A user feedback mechanism will also be implemented, allowing participants to flag problematic responses and contribute to iterative improvements. These evaluations will help ensure that GuardianAI remains functional, efficient, and ethically responsible as it evolves.

By implementing a structured approach to development and evaluation, GuardianAI aims to provide a practical and ethical legal assistance tool tailored to the foster care system. Future iterations could build upon this foundation, incorporating more advanced retrieval techniques, expanding to additional legal domains, and enhancing the system's ability to deliver reliable and meaningful legal guidance. While the current focus remains on establishing a strong prototype, these evaluations and safeguards will lay the groundwork for future improvements and potential real-world applications.

References

- Inyoung Cheong, King Xia, K. J. Kevin Feng, Quan Ze Chen, and Amy X. Zhang. 2024. [I am not a lawyer, but...: Engaging legal experts towards responsible llm policies for legal advice](#). In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, pages 2454–2469. ACM.
- Jingyun Sun, Chengxiao Dai, Zhongze Luo, Yangbo Chang, and Yang Li. 2024. [Lawluo: A multi-agent collaborative framework for multi-round chinese legal consultation](#). *arXiv preprint*, 2407.16252.
- Kepu Zhang, Weijie Yu, Sunhao Dai, and Jun Xu. 2024. [Citalaw: Enhancing llm with citations in legal domain](#). *arXiv preprint*, 2412.14556.