



Data Article

Dataset for legal question answering system in the Indian judiciary context



Veningston K*, Apratim Mishra

Department of Computer Science and Engineering, National Institute of Technology Srinagar, Jammu and Kashmir
190006, India

ARTICLE INFO

Article history:

Received 10 February 2025

Revised 21 April 2025

Accepted 5 May 2025

Available online 12 May 2025

Dataset link: [IndicLegalQA Dataset \(Original data\)](#)

Keywords:

Legal question answering

Indian judiciary

Supreme Court of India

Dataset generation

Metadata extraction

Information retrieval

ABSTRACT

Legal documents, such as court judgments and statutes, are vital for understanding judicial decisions, legal principles, and procedural details. However, these documents are often dense, complex, and abundant, making it challenging for lawyers, researchers, and citizens to quickly and easily locate and retrieve relevant information. The Legal Question Answering (LQA) [1] task involves developing systems that can automatically answer legal questions based on relevant legal documents centred on the *constitution* and *law*, preferably from delivered judgments that are considered public property of the nation. The need for specialized datasets in LQA is particularly pressing in countries like India, where legal texts follow distinct judicial structures, specialized terminologies, and procedural intricacies [2]. Due to the lack of a relevant dataset for an LQA system [3], this paper presents a comprehensive dataset designed for LQA in the Indian judiciary context that facilitates efficient legal information retrieval. The dataset comprises 10,000 question-answer pairs derived from 1256 Indian Supreme Court judgments across various legal domains, including 538 criminal and 718 civil cases available on Mendeley Data [4]. Each QA pair is derived from detailed legal judgments from the Apex court (i.e. Supreme Court of India), with the questions framed to capture essential legal issues, principles, or facts, and answers extracted directly from the text. The dataset covers a balanced mix of legal topics in criminal and civil law, such as constitutional mat-

* Corresponding author.

E-mail addresses: veningstonk@nitsri.ac.in (V. K), apratim_2022bcse061@nitsri.ac.in (A. Mishra).Social media: [@veningston](#) (V. K), [@Apratim07](#) (A. Mishra)

ters, property disputes, criminal offences, procedural matters, family disputes, employment matters, financial and taxation issues, and public welfare concerns. Additionally, it includes metadata such as case name and judgement date. This dataset supports the development of AI-driven LQA systems to enhance access to precise legal information and aid legal professionals/common citizens about India's complex legal system. To evaluate its effectiveness for legal question-answering tasks, the IndicLegalQA Dataset is fine-tuned on the "meta-llama/Llama-2-7b-hf" model [5] using Parameter-Efficient Fine-Tuning (PEFT), specifically the Low-Rank Adaptation (LoRA) technique [6–8]. The fine-tuned model is evaluated using Sentence-BERT (SBERT) [9], with the "paraphrase-MiniLM-L6-v2" model embedding. Cosine similarity measures how well the model captures the nuances of legal language between actual and generated answers. This ensures the dataset is well-suited for real-world legal applications, making it a valuable resource for improving AI-driven legal information retrieval systems.

© 2025 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Specifications Table

Subject	Computer Sciences
Specific subject area	Legal question-answering in the Indian judiciary Context
Type of data	Data Types: Table, Figure Data Format: JSON file
Data collection	The IndicLegalQA dataset collects Indian Supreme Court judgments from a publicly available Kaggle repository. We use ChatGPT-4o with structured prompts to generate supervised question-answer pairs covering key legal aspects of each case. Metadata such as case name and judgement date are also extracted. The final dataset is structured in JSON format, ensuring organized and accessible legal QA data in the Indian Judiciary context.
Data source location	Institution: National Institute of Technology (NIT) Srinagar City/Town/Region: Srinagar, Jammu and Kashmir Country: India Primary Data Source: Indian Supreme Court Judgments (Publicly Available Judgement PDFs) Dataset Accessed From: https://www.kaggle.com/datasets/vangap/indian-supreme-court-judgments
Data accessibility	Repository name: IndicLegalQA Dataset Data identification number: 10.17632/gf8n8cnmvc.2 Direct URL to data: https://data.mendeley.com/datasets/gf8n8cnmvc/2
Related research article	None

1. Value of the Data

- This dataset provides a structured collection of 10,000 question-answer pairs, case names, and judgement dates extracted from 1256 Indian Supreme Court judgments, covering diverse legal domains in the Indian judiciary context.
- Researchers in legal informatics, law students, developers of legal AI tools, and general public can benefit from this dataset, particularly those working on enhancing access to legal information, developing AI-enabled legal assistants.
- It enables fine-tuning legal question-answering models and supports a range of downstream tasks, including legal language modeling, case outcome prediction, and legal information extraction.

- The IndicLegalQA dataset is accessible via Mendeley Data under a Creative Commons Attribution 4.0 (CC BY 4.0) International licence.
- The IndicLegalQA dataset can effectively train and fine-tune Large Language Models (LLMs) for legal question-answering tasks in the Indian judicial context. This dataset enables the development of domain-adapted LLMs that understand legal terminology, reasoning patterns, and procedural context specific to Indian law. These fine-tuned models can be employed in various applications, including AI-based legal chatbots and judgment summarization systems. Additionally, the dataset supports indexing and retrieval tasks in the Retrieval Augmented Generation (RAG) pipeline, where a retriever fetches relevant context based on a query before the generator produces an answer.

2. Background

There are many existing datasets focused on various legal systems globally. However, resources specifically tailored to the Indian judiciary remain limited. To bridge this gap, we compile IndicLegalQA, a dataset derived from Indian Supreme Court judgments.

Globally, several notable datasets have been developed. PrivacyQA [10] and PolicyQA [11] address privacy and data protection in the United States, while CUAD [12] focuses on legal contract review. JEC-QA [13] contributes to the Chinese legal domain, covering fundamental legal principles and case analysis questions from China's National Judicial Examination. In Europe, LLeQA [14] provides insights into statutory law using questions posed by Belgian citizens and statutory articles collected from the Belgian legislation, while BSARD [15] is a French native corpus for studying statutory article retrieval.

IndicLegalQA Dataset [4] is unique in its focus on Indian legal practices, offering 10,000 descriptive QA pairs sourced from 1,256 Supreme Court judgments (538 criminal and 718 civil cases) along with metadata (case name and judgement date). It covers various domains, including constitutional matters, property disputes, criminal offenses, procedural issues, and family disputes.

We emphasize that while several global LQA datasets exist as shown in Table 1, they are either focused on non-Indian jurisdictions or do not reflect the structural and procedural complexity of Indian legal texts. IndicLegalQA is the first large-scale, supervised Q&A dataset derived from Indian Supreme Court judgments to capture domain-specific legal language, judicial reasoning styles, and procedural content relevant to India's legal system.

3. Data Description

3.1. Statistics on the original judgment documents

Table 2 presents statistical information about the length of judgment documents in terms of words and pages. It includes three key metrics - maximum length, minimum length, and average length. The longest document contains 126,546 words spread over 538 pages, while the shortest document has 193 words over just 2 pages. On average, the judgment documents have 5,485 words and span 23 pages. This data highlights the variability in the length of legal judgments, ranging from brief decisions to extensive, detailed rulings.

3.2. Statistics on the QA pairs

Table 3 provides statistical details on the lengths of questions and answers in the dataset. It shows that the maximum length for a question is 39 words, while the longest answer contains 104 words. The shortest question and answer consist of 4 and 5 words, respectively. On average,

Table 1
Existing LQA datasets and comparison with the proposed IndicLegalQA.

Datasets	Source	Country	Legal Domain	Dataset Size
PrivacyQA [10]	Privacy policies of mobile applications on the Google Play Store	United States of America	Privacy and data protection policies.	1750 questions with 3500 expert annotations of relevant answers
JEC-QA [13]	National Judicial Examination of China	China	Fundamental law of the nation	9, 761 Knowledge-Driven Questions and 16, 604 Case-Analysis Questions (MCQs)
LLeQA [14]	Legal statutes from the Belgian legislation	Belgium	Statutory Law	1868 legal questions posed by Belgian citizens and labeled by experienced jurists based on 27,941 statutory articles collected from the Belgian legislation
Lawyer GPT India [19]	Indian Constitution (fundamental rights), judiciary, legislative, and various socio-political issues in India	India	Indian polity	150 QA pairs
BSARD [15]	Statutory articles from the Belgian legislation	Belgium	Statutes and legal IR	1100+ French native legal questions labelled by experienced jurists with relevant articles from a corpus of 22,600+ Belgian law articles.
PolicyQA [11]	Privacy policies from 115 websites	United States of America	Privacy and data protection	25,017 question-passage-answer triples.
CUAD [12]	Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system,maintained by the U.S. SEC	United States of America	Legal contract review	More than 500 contracts with over 13,000 expert annotations across 41 label categories.
COLIEE (Task 2) [20]	Japanese Bar Exam Legal Documents, Civil Case Law from the Supreme Court of Japan	Japan	Civil and Statute Law	~300 QA pairs of Yes/No Questions from Statute Law QA
IndicLegalQA (proposed dataset)	Indian Supreme Court Judgements	India	Constitutional matters, property disputes, criminal offenses, procedural matters, family disputes etc.	10,000 descriptive QA pairs from 1256 judgments (538 criminal and 718 civil cases) along with metadata (case name & judgement date).

Table 2
Word and page counts for judgment documents.

Statistic	Words	Pages
Max length	126,546	538
Min length	193	2
Average length	5485	23

Table 3
Distribution of question-and-answer lengths in words.

Statistic	Questions	Answers
Max length	39	104
Min length	4	5
Average length	15	34

Table 4
Sample data instances from IndicLegalQA dataset.

S. No.	Fields	Instance
1	case_name	Union of India vs. Maj. Gen. Manomoy Ganguly
	judgement_date	1st August 2018
	question	What decision did the Armed Forces Tribunal (AFT) make regarding Maj. Gen. Manomoy Ganguly's promotion?
	answer	The AFT directed the appellants to post Maj. Gen. Manomoy Ganguly as DGMS (Army) as expeditiously as possible and within one month from the date of the judgment.
2	case_name	Union of India vs. Board of Control for Cricket in India & Ors.
	judgement_date	22nd August 2017
	question	What did the High Court of Delhi rule regarding the sharing of live broadcasting signals in the case Union of India vs. Board of Control for Cricket in India & Ors.?
	answer	The High Court of Delhi ruled that the signals received by Prasar Bharati from BCCI should not be retransmitted by cable operators and that this retransmission was not covered under Section 8 of the Cable Act, 1995.
3	case_name	Shakti Bhog Food Industries Ltd. vs. The Central Bank of India & Anr.
	judgement_date	5th June 2020
	question	What was the primary issue in the case Shakti Bhog Food Industries Ltd. vs. The Central Bank of India & Anr.?
	answer	The primary issue was whether the suit filed by Shakti Bhog Food Industries Ltd. for rendition of accounts and recovery of excess amounts charged by the Central Bank of India was barred by the law of limitation.
4	case_name	Vikash Kumar vs. Union Public Service Commission & Ors.
	judgement_date	11th February 2021
	question	What did the Supreme Court say about the obligation of examining bodies under the Rights of Persons with Disabilities Act, 2016?
	answer	The Supreme Court stated that examining bodies have an obligation to provide reasonable accommodations, such as the assistance of a scribe, to candidates with disabilities to ensure their full and effective participation in examinations.
5	case_name	Apollo Institute of Medical Sciences & Research vs. Union of India & Anr.
	judgement_date	31st August 2017
	question	Why did the Central Government debar Apollo Institute of Medical Sciences & Research from admitting students for two years?
	answer	The Central Government debared the college based on a negative recommendation from the Medical Council of India (MCI), citing deficiencies in faculty and resident doctors during inspections.

questions are 15 words long, and answers are 34 words long. This data reflects the variability in the complexity and detail of questions and answers within the dataset.

3.3. Sample dataset instances

Each entry in the IndicLegalQA dataset is structured in JSON format and includes a question-answer pair along with key metadata fields such as the case name and judgement date. The example instances, shown in Table 4, illustrates records derived from a Supreme Court judgment.

3.4. Data visualizations

The visualizations emphasize the main themes within the IndicLegalQA dataset. These provide a comprehensive view of the dataset's focus on core legal elements.

The word clouds shown in Fig. 1 (a) & (b) illustrate recurring terms in questions and answers. In the question's word cloud, prominent terms like "Supreme Court," "decision," and "legal issue" emphasize judicial rulings and principles, while phrases like "vs State" and "Union India" underscore the involvement of specific parties. In the answers word cloud, similar terms appear

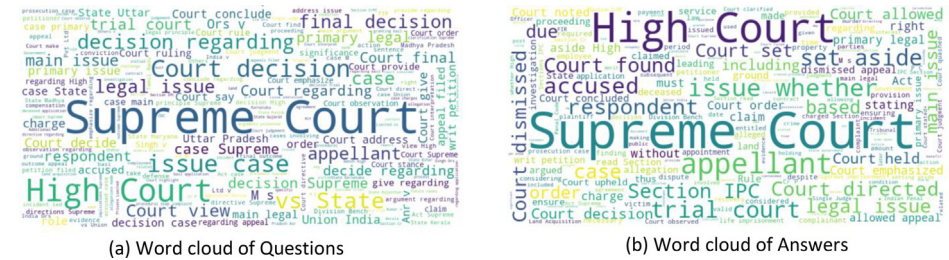


Fig. 1. (a) & (b). Word cloud representation of most frequent terms in questions and answers.

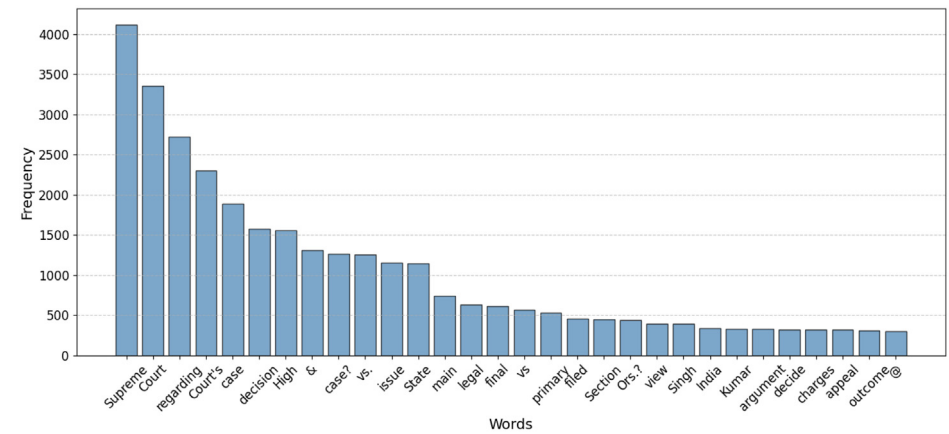


Fig. 2. Word frequency distribution of questions.

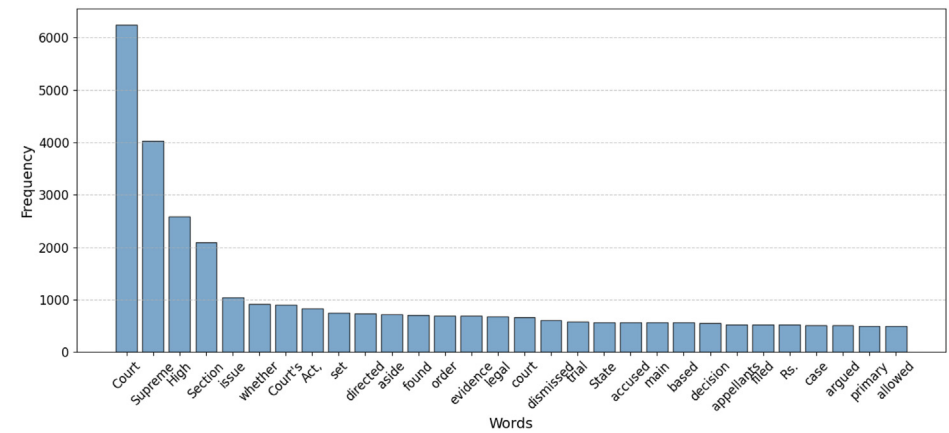


Fig. 3. Word frequency distribution of answers.

alongside words like “set aside” and “directed,” indicating a focus on judicial actions and case outcomes. Together, these word clouds establish a consistent emphasis on the judiciary context. The word frequency distributions shown in Figs. 2 and 3 quantify the focus on judicial roles and confirm that both questions and answers are oriented toward judicial decisions and legal principles. Fig. 2 emphasizes case-related terms (e.g., “case,” “decision,” “appeal”), while

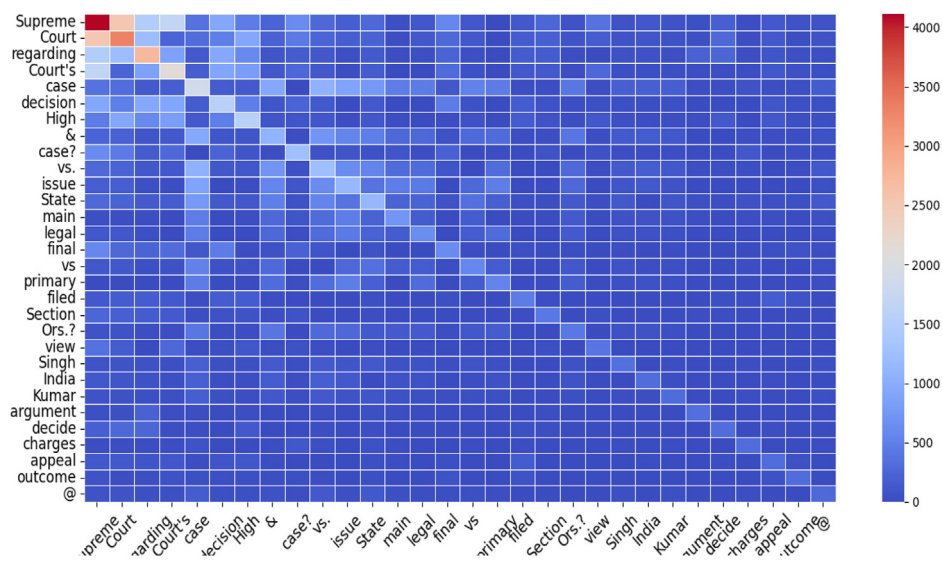


Fig. 4. Word heatmap of questions.

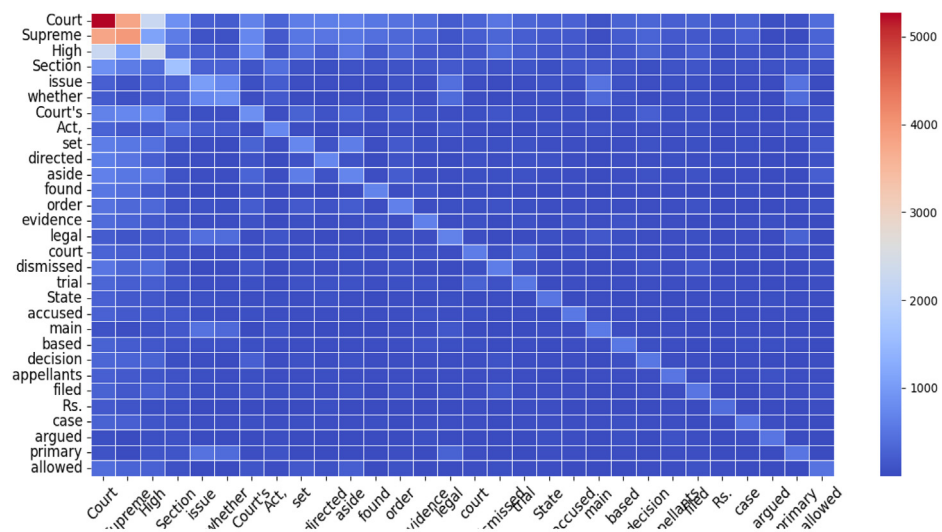


Fig. 5. Word heatmap of answers.

Fig. 3 captures a mix of statutory references (e.g., “Act,” “Section”) and procedural outcomes (e.g., “dismissed,” “argued”).

Co-occurrence heatmaps shown in Figs. 4 and 5 reveal frequent word pairings, such as “Supreme Court” with “decision” and “appellant” with “respondent,” which highlight structured legal phrasing and recurring patterns in judicial discourse.

Fig. 6 illustrates the distribution of word lengths for questions (in blue) and answers (in orange) in the dataset. The histogram reveals that questions are generally shorter, with their lengths peaking between 10 and 20 words, reflecting concise and precise phrasing. In contrast, answers tend to be longer, with a higher frequency observed around 30 to 40 words, high-

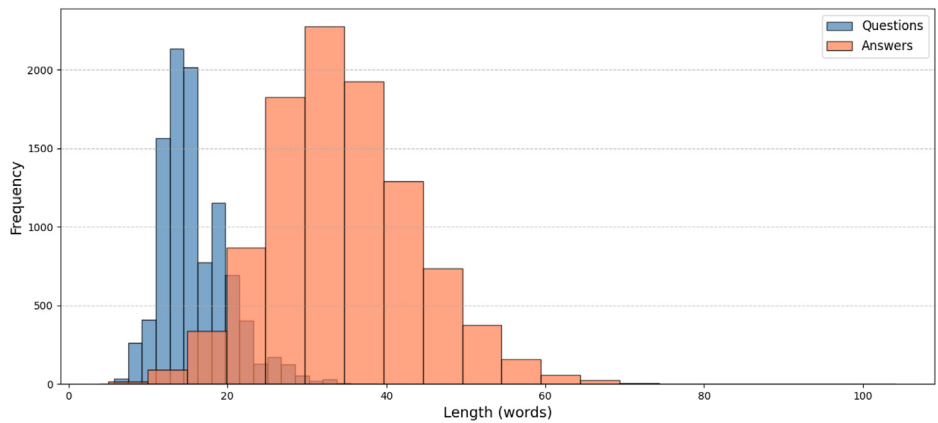


Fig. 6. Relationship between length of questions and answers.

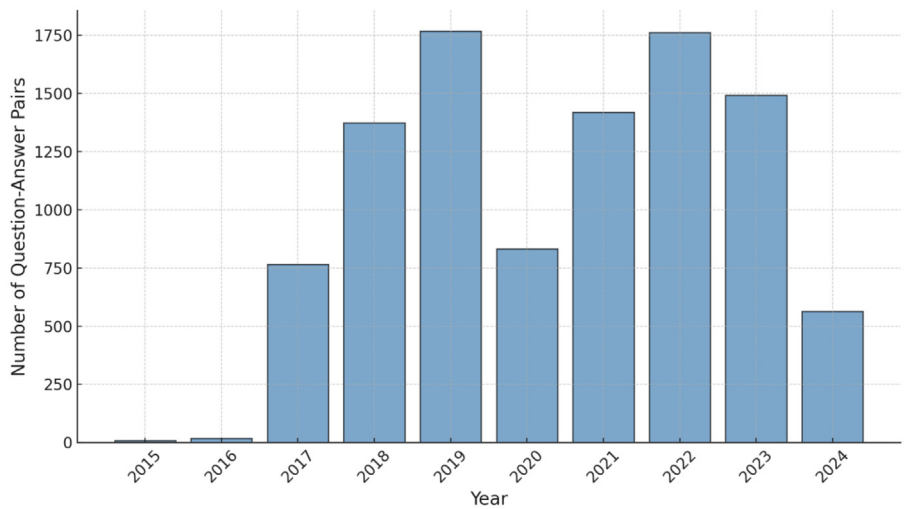


Fig. 7. Number of question-answer pairs prepared by year

lighting their explanatory nature. The overlap between the two distributions indicates occasional similarity in length, but the overall spread for answers is broader, extending up to around 100 words. This distinction underscores the dataset’s structure, where questions are designed to be succinct, while answers provide detailed and context-rich responses. This contrast aligns with the purpose of answers in providing detailed clarity.

Fig. 7 illustrates the number of question-answer pairs prepared, categorized by the year of the judgments they are derived from. This distribution indicates a strong emphasis on judgments from 2017 to 2022, reflecting the dataset’s focus on contemporary legal issues and the availability of comprehensive data for these years.

3.5. Types of questions

The questions in the IndicLegalQA dataset span a range of inquiry types relevant to legal contexts. The question types mentioned in the dataset as listed below are not mutually exclusive

and may overlap in practice. A single question can exhibit characteristics of multiple categories depending on its phrasing and the legal context it addresses.

- **Factual:** Direct questions that ask for specific information.
- **Interpretive:** Questions seeking explanations or analysis.
- **Procedural:** Inquiries about steps, methods, or processes.
- **Contextual:** Questions requiring background understanding.
- **Predictive:** Questions about predictions or expected outcomes.

While each question typically aligns with a specific category, certain questions might fall into multiple categories depending on the individual's interpretation. This classification provides diverse types of legal inquiries rather than a strict distribution to support various legal inquiry types in a non-overlapping manner.

4. Experimental Design, Materials and Methods

The IndicLegalQA dataset is created in two stages: (i) collecting Supreme Court of India judgment documents and (ii) generating supervised question-answer pairs along with extracting metadata, such as case name and judgement date from these documents. This approach helps us create a clear and organized dataset for a legal question-answering system as shown in Fig. 8.

A comprehensive flowchart of the QA generation pipeline shown in Fig. 8 visually represent the sequence of steps involved in semi-automated dataset creation process.

4.1. Judgment documents collection

We begin by collecting Supreme Court of India judgment PDFs to ensure comprehensive legal coverage. We collected Indian Supreme Court Judgments from Indian Supreme Court Judgments available on Kaggle [16], which contains PDF documents of Indian Supreme Court judgments starting from the year 1950. The collection also contains a CSV file with the list of judgments extracted from the government website API. The CSV file contains metadata such as judgement date, bench, petitioner, respondent.

4.2. Uploading judgment PDFs to the LLM

Once the judgment documents are collected, we upload them into a LLM to generate supervised question-answer pairs and extract key metadata. For this purpose, we used pre-trained LLM ChatGPT 4o to ensure a consistent and efficient process for handling the legal documents. [17,18]

4.3. Prompt design

A structured prompt is designed to guide the LLM in generating accurate and meaningful question-answer pairs. The entered prompt ensures the following objectives.

- Each question-answer pair explicitly contains the case name and judgement date, enabling immediate identification of the case and its context without referencing the entire judgment.
- The questions are crafted to address the critical legal issues of each case. We have not explicitly enforced answer length thresholds, but typical answers range between 5 and 100 words.

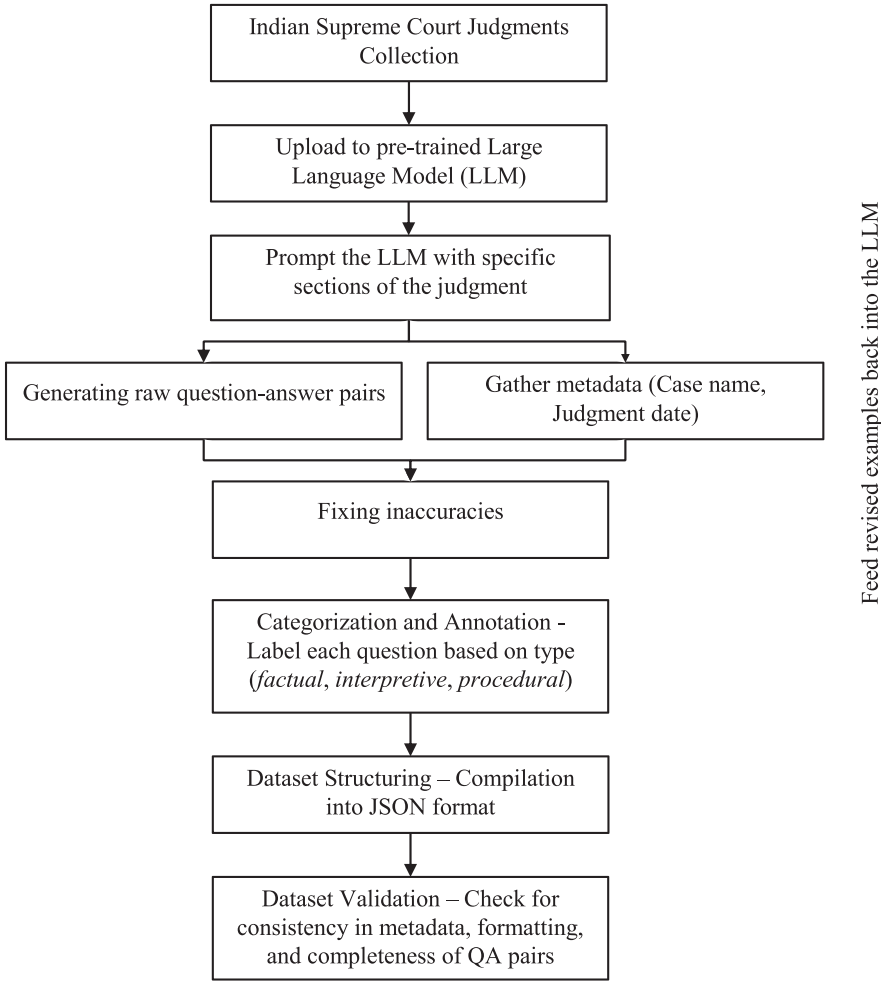


Fig. 8. Workflow for curating QA pairs.

Table 5
Specifications of equipment and software.

S. No.	Specifications	Tools
1	Language model used	ChatGPT-4o
2	Platform used	Web interface via OpenAI's ChatGPT
3	Document Processing	Judgement PDFs are processed via structured prompt injection
4	Annotation and structuring data instances	Python scripts for JSON compilation and validation of metadata

- For each judgment, 4 to 15 questions were generated, covering essential factual, interpretive, procedural, and clarifying details. There can be more question-answer pairs too for a case. We focused on sampling complete and legible Supreme Court judgments from 2010–2022, balanced across civil and criminal domains.
- We employed specifications shown in Table 5 to generate question-answer pairs due to their capability to handle various question types effectively, particularly factual and procedural inquiries. Research on LLM models as a question-answering system shows that it performs well

when provided with contextual input, enhancing both precision and recall in complex tasks by grounding responses in the surrounding text, a crucial factor for legal documents. Additionally, structured prompt designs reduce hallucinations by keeping responses within the given context, closely aligning with the factual accuracy requirements in legal QA tasks. This suitability makes ChatGPT-4o an effective choice for generating diverse question types—factual, interpretive, procedural, and clarifying types within the proposed dataset.

4.4. QA pair generation and metadata extraction

After uploading the documents and applying the structured prompt, the LLM generates the question-answer pairs. Simultaneously, we extract metadata, such as the *case name* and *judgement date*, from each judgment. This process ensures that the resulting dataset is well-organized, with each pair explicitly tied to its respective case. We acknowledge that minor variability may occur due to LLM stochasticity. To address this, we employed prompt templates and manual review of outputs for consistency across cases.

4.5. Structuring the final dataset in JSON format

The generated question-answer pairs, along with their associated metadata, are compiled into a structured JSON format file. Each JSON object includes the following.

- *Case Name* - The name of the case.
- *Judgement Date* - The date of the judgment, formatted as “25th January 2023.”
- *Question* - A query highlighting a key legal issue or significant fact from the case.
- *Answer* - A concise answer directly addressing the question.

The duration of data collection and QA generation was conducted over six weeks, with validation and formatting performed in subsequent stages. Metadata was cross-verified with the original judgment documents.

Limitations

- The dataset focuses exclusively on Supreme Court judgments. While the dataset includes a balanced mix of civil and criminal cases from the Supreme Court, it currently does not include cases from High Courts, tribunals, and district courts, which may affect its applicability for region-specific legal models. This has been highlighted as an avenue for future dataset extensions.
- Although QA pairs are extracted with structural validation and accuracy checks, rare instances of minor errors may exist due to semi-automated extraction.
- The dataset currently does not support multilingual queries beyond English.

Ethics Statement

The authors confirm that the work complies with ethical publication guidelines. This dataset does not involve human subjects, animal experiments or any data collected from social media platforms. There are no restrictions on the data shared, since all documents are derived from publicly available Supreme Court judgments. No proprietary, confidential, or privacy-sensitive information is included.

CRedit Author Statement

Venningston K: Conceptualization, Investigation, Resources, Writing - Review & Editing, Supervision. **Apratim Mishra:** Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft, Visualization.

Data Availability

[IndicLegalQA Dataset \(Original data\)](#) (Mendeley Data)

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Martinez-Gil, A survey on legal question-answering systems, *Comput. Sci. Rev.* 48 (2023), doi:[10.1016/j.cosrev.2023.100552](#).
- [2] S.K. Nigam, S.K. Mishra, A.K. Mishra, N. Shallum, A. Bhattacharya, Legal question-answering in the indian context: efficacy, challenges, and potential of modern AI models, *arXiv preprint arXiv:2309.14735*, 2023 (2023).[https://doi.org/10.48550/arXiv.2309.14735](#).
- [3] A. Abdallah, B. Piryani, A. Jatowt, Exploring the state of the art in legal QA systems, *J. Big Data* (2023), doi:[10.1186/s40537-023-00802-8](#).
- [4] K. Venningston, A. Mishra, IndicLegalQA dataset, Mendeley Data, V2 (2024).[https://doi.org/10.17632/gf8n8cnmnc.2](#).
- [5] Meta AI, meta-llama/llama-2-7b-hf, hugging face model hub. [https://huggingface.co/meta-llama/llama-2-7b-hf](#). (accessed 14 October 2024)
- [6] Martin Weyssow, Xin Zhou, Kisub Kim, David Lo, and Houari Sahraoui. 2025. Exploring Parameter-Efficient Fine-Tuning Techniques for Code Generation with Large Language Models. *ACM Trans. Softw. Eng. Methodol.*, Jan 2025. [https://doi.org/10.1145/3714461](#).
- [7] Luping Wang, Sheng Chen, Linnan Jiang, Shu Pan, Runze Cai, Sen Yang, Parameter-efficient fine-tuning in large language models: a survey of methodologies, *Artif Intell Rev* 58 (2025) 227 [https://doi.org/10.1007/s10462-025-11236-4](#).
- [8] E.J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, LoRA: low-rank adaptation of large language models, *arXiv preprint arXiv:2106.09685*, 2021 (2021).[https://doi.org/10.48550/arXiv.2106.09685](#).
- [9] N. Reimers, I. Gurevych, Sentence-BERT: sentence embeddings using siamese BERT-networks, in: *Proceedings of the 2019 Conference Empirical Methods Natural Language Processing (EMNLP-IJCNLP)*, 2019, 2019, pp. 3982–3992. [https://aclanthology.org/D19-1410/](#).
- [10] Abhilasha Ravichander, Alan W Black, Shomir Wilson, Thomas Norton, and Norman Sadeh. 2019. Question Answering for Privacy Policies: Combining Computational and Legal Perspectives. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4947–4958, Hong Kong, China. Association for Computational Linguistics. [https://doi.org/10.18653/v1/D19-1500](#).
- [11] Wasi Ahmad, Jianfeng Chi, Yuan Tian, and Kai-Wei Chang. 2020. PolicyQA: A Reading Comprehension Dataset for Privacy Policies. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 743–749, Online. Association for Computational Linguistics. [https://doi.org/10.18653/v1/2020.findings-emnlp.66](#).
- [12] D. Hendrycks, C. Burns, A. Chen, S. Ball, CUAD: an expert-annotated NLP dataset for legal contract review, *arXiv preprint arXiv:2103.06268*, 2021 (2021).[https://doi.org/10.48550/arXiv.2103.06268](#).
- [13] H. Zhong, C. Xiao, C. Tu, T. Zhang, Z. Liu, M. Sun, JEC-QA: a legal-domain question answering dataset, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, 38, 2024, pp. 22266–22275, doi:[10.1609/aaai.v34i05.6519](#). (2020).
- [14] A. Louis, G. van Dijk, G. Spanakis, Interpretable long-form legal question answering with retrieval-augmented large language models, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, 38, 2024, pp. 22266–22275, doi:[10.1609/aaai.v38i20.30232](#).

- [15] A. Louis, G. Spanakis, A Statutory Article Retrieval Dataset in French, arXiv preprint arXiv:2108.11792, 2022 (2022).<https://doi.org/10.48550/arXiv.2108.11792>.
- [16] Indian Supreme Court Judgments Collection. <https://www.kaggle.com/datasets/vangap/indian-supreme-court-judgments>. (accessed 14 October 2024).
- [17] Alberto Olmo, Sarath Sreedharan, Subbarao Kambhampati. GPT3-to-plan: Extracting plans from text using GPT-3. 2021, International Conference on Automated Planning and Scheduling (ICAPS) — Knowledge Engineering for Planning and Scheduling. https://icaps21.icaps-conference.org/workshops/KEPS/Papers/KEPS_2021_paper_8.pdf
- [18] H. Bahak, F. Taheri, Z. Zojaji, A. Kazemi, Evaluating ChatGPT as a Question Answering System: a Comprehensive Analysis and Comparison with Existing Models, *Journal of Computing and Security* (2024), doi:[10.22108/jcs.2024.140680.1141](https://doi.org/10.22108/jcs.2024.140680.1141).
- [19] Lawyer GPT India https://huggingface.co/datasets/nisaar/Lawyer_GPT_India [Accessed on 14 October 2024]
- [20] J. Rabelo, M.Y. Kim, R. Goebel, M. Yoshioka, Y. Kano, & K. Satoh, (2022). Overview and discussion of the competition on legal information extraction/entailment (COLIEE) 2021. Pages 111-133. [10.1007/s12626-022-00105-z](https://doi.org/10.1007/s12626-022-00105-z).