

SvedbergOpen
DISSEMINATION OF KNOWLEDGEInternational Journal of Artificial
Intelligence and Machine LearningPublisher's Home Page: <https://www.svedbergopen.com/>

Research Paper

Open Access

Bridging The Vernacular Divide: An Empirical Study of AI-Assisted Legal Access in Jammu and Kashmir, India

Dr. Nitán Sharma¹, Dr. Ankita Yadav², Dr. Paramjeet Kour³, Bhavika Sharma⁴, Dr. Sushma Bala⁵, Dr. Ritika⁶, Nishu Sudan⁷, Dr. Alisha Sharma⁸¹Assistant Professor, School of Law, SMVDU-Katra, J&K. E-mail: nitan.sharma@smvdu.ac.in²Dean & Head, School of Law, SMVDU-Katra, J&K.³Assistant Professor, School of Teacher Education, Cluster University of Jammu, J&K.⁴Independent Researcher and Legal Practitioner.⁵Associate Professor in Education, GCE, Cluster University of Jammu, J&K.⁶Assistant Professor, School of Law, Model Institute of Engineering and Technology, Jammu, J&K.⁷Research Scholar, Department of Law, University of Jammu, J&K.⁸Assistant Professor, Dogra Law College (Autonomous), J&K.

Abstract

Access to justice entails not only availability of the court of laws and the avenues of justice but also that individuals can access and understand the legal information from which their rights and obligations arise. This becomes more important in linguistically diverse societies where understanding of the legal information becomes difficult. This could be also because of limited proficiency in legal English or due to the user other languages. This study examines the role of artificial intelligence (AI), including multilingual Large Language Models (LLMs) in improving access to legal information in the Union Territory of Jammu and Kashmir. The study adopts a descriptive, cross-sectional survey design involving 113 respondents - 31 legal educators and 82 law students. Though, the current study does not involve any benchmarking of AI technology; hence cannot be taken as means by which AI improves access to justice. Nevertheless, this study attempts to find out the conditions under which AI assisted legal accessibility may be pursued. The empirical findings are examined alongside constitutional provisions. The paper proposes an 'assistive bridge' model in which AI acts as an interface between individuals and legal materials rather than a source of the same or a substitute for human legal judgment.

Keywords: Artificial Intelligence; Large Language Models; Legal Translation; Legal Summarisation; Access to Justice; Multilingual AI; Legal NLP; Jammu and Kashmir; AI Governance; Legal Technology

1. Introduction

Access to justice has a 'democratic resonance'¹ that is intrinsically linked with fundamental rights and the rule of law. Earl Warren, Chief Justice of the Supreme Court of the United States, stated that "It is the spirit and not the form of law that keeps justice alive."² The mere existence of a legal right does not necessarily ensure meaningful access to justice. For justice to be truly accessible, the law through which such rights are interpreted and enforced must also be capable of being understood. The denial of meaningful access to justice is, in itself, an injustice. In this sense, access to justice cannot be limited to the physical availability of courts or the formal availability of legal remedies. It must also extend to meaningful access to the law, including the judgments through which courts interpret the law and determine the rights and obligations of individuals. If the law is only accessible in form, but it is difficult to comprehend or put it into practice the promise of access to justice then remains incomplete. Therefore it also has with itself an informational and linguistic dimension - where legal information is difficult to understand or comprehend due to linguistic barriers then meaningful access may be impaired.

This study has been undertaken in the Union Territory of Jammu and Kashmir wherein this linguistic dimension is particularly visible. The Jammu and Kashmir Official Languages Act, 2020³ recognises Kashmiri, Dogri, Urdu, Hindi and English as official languages of the Union Territory for the purposes

¹ Nick Robinson, "The Structure and Functioning of the Supreme Court of India", in Gerald Rosenberg, Sudhir Krishnaswamy & Shishir Bail (eds.), *A Qualified Hope: The Indian Supreme Court and Progressive Social Change* (Cambridge University Press, 2019), available at: SSRN: <https://ssrn.com/abstract=3483275>.

² Earl Warren, "The Law and the Future", in Henry M. Christman (ed.), *The Public Papers of Chief Justice Earl Warren* 221, 229 (Simon & Schuster, 1959).

³ Jammu and Kashmir Official Languages Act, 2020 (Act No. 23 of 2020).

specified by the Act. The constitutional framework governing the language of proceedings before the higher judiciary also gives English a central position. However, the linguistic diversity of Jammu and Kashmir extends beyond the five official languages. Gojri, the language of the Gujjar and Bakarwal communities, and Pahari, spoken by communities in the Pir Panjal region, are among the most widely spoken vernacular languages in the Union Territory.⁴ Although these languages do not presently enjoy the status of official languages under the 2020 Act, they remain the primary means of comprehension for a substantial segment of the population. The result is a potential gap between the language in which legal information is produced or accessed and the language in which an individual may find that information easiest to understand. This makes the vernacular divide not merely a matter of official language policy but a lived reality for communities whose linguistic heritage remains outside the formal legal-linguistic framework.

This issue significantly increases in the case of judicial decisions. This is because while summing up judgments, three main challenges emerge that are different from ordinary text summarisation. First, there is the issue of the length of the judgments - beginning from the short factual matrix of the case, to analyzing the law in India and elsewhere, and then applying the law to the facts of the case. This leads to a rather lengthy judgement and a meaningful summary cannot solely reduce length of the judgement; it must also preserve its character. Second, judgments are complex in nature, and sometimes it becomes hard to distinguish between what is said in terms of ratio decidendi or obiter dicta. Summaries must distinguish between them otherwise it will lead to an incorrect understanding or give undue importance to factors. Third, judgments are mostly written in complex English, thus adding another challenge for people who lack adequate proficiency in legal English. All these challenges make the process of summarizing legal judgments distinct from the ordinary text summarization since a useful summary of a judgment does not only require reduction of length but should also include legally important facts, issues, reasoning, applicable law and the final outcome of the decision.

In this paper, an attempt has been made to bridge the vernacular divide through the use of Artificial Intelligence (AI) and multilingual Large Language Models (LLMs). The object is to facilitate a more meaningful access to law and judicial decisions. It examines the perceptions of law students and legal educators regarding the use of artificial intelligence in understanding legal information and judicial decisions. More specifically, this study attempts to determine whether AI-powered multilingual tools could act as an instrument for rendering legal information available to individuals who might encounter difficulties in understanding judicial opinions due to language barriers.

2. RESEARCH QUESTIONS

The present study addresses three essential questions. These include:

1. First, what is the frequency with which AI-LLM powered technology is being employed by law students and legal academics in Jammu and Kashmir?
2. Second, how do the respondents view the linguistic accessibility, reliability, explainability, and verifiability of AI-enabled legal translation and summarisation?
3. Third, what are the requirements that must be met in order to ensure responsible usage of such technologies for legal information access?

This research will not strictly test or compare individual AI models. Instead, its value lies in being empirical and regulatory in the sense of assessing the perception of legal academics and students within the evolving Indian regulatory framework around AI usage.

3. Research Methodology

3.1 Study Design

This study adopts a mixed-method approach. It also employs a descriptive, cross-sectional survey design as part of the empirical analysis. The study was limited to the Union Territory of Jammu and Kashmir. Two separate structured questionnaires were used: one aimed at law educators, and another aimed at law students. The questions pertained to the use of AI in translation and summarisation. It also contained questions regarding data confidentiality and privacy, concerns of fabricated/ hypothetical cases, source verification, and preparedness for AI use.

3.2 Participants and Sampling

The final sample size included 113 respondents 31 law educators and 82 law students from premier legal institutions within the Union Territory of J&K by specifically utilizing a combination of purposive and convenience sampling techniques. Recognizing the distinct roles, perspectives, and patterns of AI usage

⁴ Zahoor Ahmad Bhat and Mahmood Ahmad Khan, "Exploring the Indigenous language of Gujjar and Bakerwal Communities" 26 *The Journal of Indian Arts History Congress* (2020-2021).

between these two groups, the data from educators and students were collected via two separate, parallel questionnaires and subsequently analysed independently. This approach allows for a nuanced comparative understanding of the phenomenon from both sides of the pedagogical relationship.

3.3 Data Analysis

The data analysis is descriptive. The responses received are reported primarily as frequencies and percentages. Because several questions permitted multiple selections, the percentages for those questions do not sum to 100 per cent. The collected data was then further analysed in order to find out the perception of both educators and students towards the usage of AI and LLMs in legal education. The findings are not presented as representative of all law students or legal educators of Jammu and Kashmir.

3.3 Ethical Considerations

Participation in the study was voluntary. All responses are kept confidential and used only for academic research purposes. Any personally identifiable information is not published. Participants were also encouraged to answer honestly.

4. Literature Review - Technical Benchmarks

In this regard, one of the most pertinent that have been conducted until now is the MILDSum study named "MILDSum: A Novel Benchmark Dataset for Multilingual Summarization of Indian Legal Case Judgments"⁵. MILDSum is unique as it was the first work to enable 'cross-lingual summarization of Indian case judgments'.⁶ Particularly, the study tries to solve the problem of summarization of Indian legal case judgments across languages. The researchers have built a dataset including 3,122 Indian court judgments that have been collected from the Supreme Court and multiple High Courts, with summaries sourced from LiveLaw.

Two main methods were considered in the experiments for multilingual legal summarisation. The first method considered the "summarise then translate" pipeline where an English summary was first created from the judgement, which was followed by translation to Hindi. The second one used the direct cross-lingual method, which created the Hindi summary directly from the English judgement. It can be seen that the former method proved to be more successful, since the best model scored 32.27 ROUGE-2 in English and 24.87 in Hindi as opposed to 21.76 in the latter case. One notable aspect about MILDSum is the presence of relatively high amounts of extractive nature in the reference summaries. According to the authors, the summaries included sentences or sections of text directly copied from the original judgements, hence the reason why extractive summarisation techniques scored well in the benchmark test. The highest ROUGE score was attained by the supervised extractive model known as SummaRuNNer but some abstractive models did well in BERTScore. In this case, the results cannot prove that extractive summarisation is always better; it only shows how performance is linked to the nature of the reference summaries being used for assessment.

MILDSum has significance in the technical evaluation of different summarisation models but it also reveals how AI can be applied to achieve multilingual summarisation of English judgments for readers whose command of legal English is inadequate. However, the study highlights the challenge associated with the cross-lingual legal summarisation and calls for research on more suitable models to deal with such texts. Notably, the assessment of MILDSum was largely done using computational measures like ROUGE and BERTScore. As a result, it serves to show how effective the computational process can be, but by no means does it guarantee that the summary produced by the AI is legally sufficient to replace the original decision or legal opinion. This point is especially relevant when it comes to the concept of access to justice, where linguistic intelligibility should be complemented by legal meaning.

In an earlier work undertaken by Shukla et al. in 2022,⁷ an effort was made to investigate the performance of legal case document summarization techniques using a series of experiments conducted using a range of extractive and abstractive summarization techniques. The research focused specifically on Indian Supreme Court judgments, exploring the performance of different summarization techniques on long and

⁵ Debtanu Datta et al., "MILDSum: A Novel Benchmark Dataset for Multilingual Summarization of Indian Legal Case Judgments," Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, 5291–5302 (2023) DOI: 10.18653/v1/2023.emnlp-main.321, available at: <https://aclanthology.org/2023.emnlp-main.321/>.

⁶ *Ibid.*

⁷ Abhay Shukla et al., "Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation," Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing, Vol. 1, 1048–1064 (2022) DOI: 10.18653/v1/2022.aacl-main.77, available at: <https://aclanthology.org/2022.aacl-main.77/>.

complicated legal texts. It was found out that the process of legal judgment summarization differs significantly from text summarization because the summary should preserve important legal data from various sections of the judgment, which include facts, issues, arguments, statutes, precedents, reasoning, and the decision itself. It should be noted that no single summarization technique performed equally well on all sections of a judgment.

Furthermore, the study also tested the quality of the summaries created by using the help of legal practitioners and realized that automated metrics for summarizing did not always correlate with the expert opinion of the lawyers. This finding was crucial since it demonstrated that a summary should not just be considered effective based on its length or any other metric. It is important that a summary should also include the legal significance of the judgement. Subsequently, this paper served as a foundation for further research into Indian legal judgment summarization. In particular, the MILDSum study carried out by Datta et al. in 2023 mentions the datasets created by Shukla et al., adding that these datasets included only the English case documents and summaries. The aim of the MILDSum was to expand the scope of the previous research through English and Hindi summaries of Indian legal judgments.

Another work of significance is that done by Joshi et al., which introduced IL-TUR (Indian Legal Text Understanding and Reasoning) benchmark.⁸ The benchmark comprises monolingual as well as multilingual tasks with Indian Legal texts in nine Indian languages as well as English and Hindi languages. The tasks are meant to cover various aspects of understanding Indian legal documents such as legal entity recognition, rhetorical role prediction, court judgment prediction, bail prediction, legal statute identification, prior case retrieval, summarization, and legal machine translation. This research is particularly pertinent because it acknowledges that legal documents are far longer and more complicated than normal texts and contain legal terms with specialised meaning. The benchmark was therefore developed to test different aspects of legal understanding and reasoning in the Indian legal context.

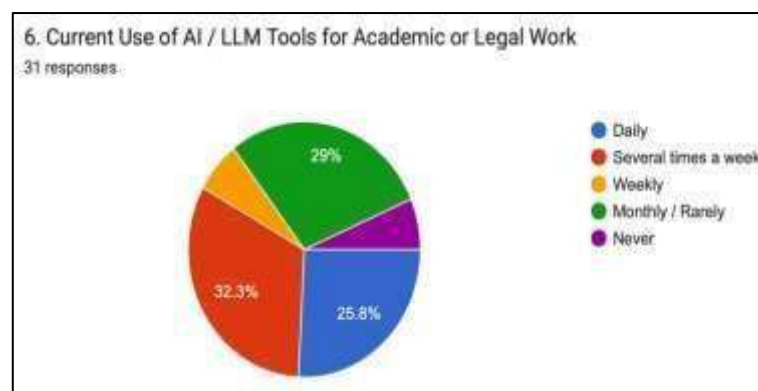
However, while the existing studies primarily examine the technical performance of legal language models and AI, there remains limited empirical research on the perception of these technologies in relation to their technical performance. This paper attempts to examine how these technologies are perceived by the actual users of legal information, especially in a multilingual region like Jammu and Kashmir. This study attempts to fill this void by analysing the perceptions of legal educators as well as students regarding the use of AI and multilingual tools in comprehending legal information and judgments.

5. EMPIRICAL FINDINGS

A. LAW EDUCATORS

1. Frequency and kinds of Tools used

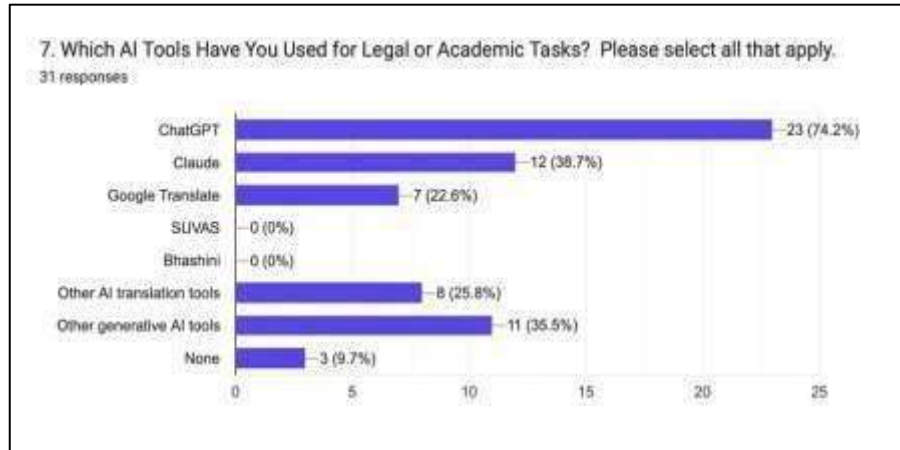
In the context of legal education, the responses show that AI/LLM tools are already being used by teachers and educators for legal or academic work. Out of the 31 respondents, 32.3% stated that they use such tools several times a week, while 25.8% reported using them almost daily. Another 29% stated that they use AI/LLM tools monthly or rarely. 6% of respondents stated that they use it weekly and another 6.5% respondents stated that they never use them. This indicates that largely, AI has already become a part of the academic and legal work of a considerable number of educators. Although, the frequency of its use differs among the respondents.



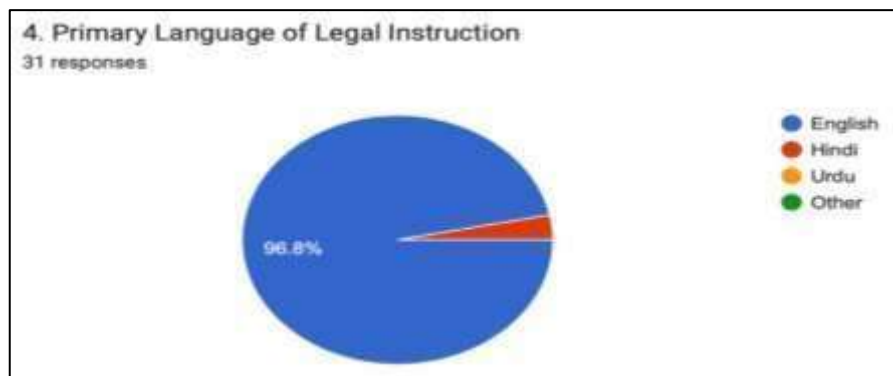
⁸ Abhinav Joshi, Shounak Paul, Akshat Sharma, Pawan Goyal, Saptarshi Ghosh & Ashutosh Modi, "IL-TUR: Benchmark for Indian Legal Text Understanding and Reasoning", Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) 11460–11499 (2024), DOI: 10.18653/v1/2024.acl-long.618, available at: <https://aclanthology.org/2024.acl-long.618>

Among the different AI tools used for legal or academic tasks, ChatGPT had the most usage of 74.2% of the respondents. The use of Claude was mentioned by 38.7% of the respondents, whereas the use of other generative AI tools was mentioned by 35.3%. Moreover, other AI tools for translations were used by 25.8% of the respondents, while 7% of the respondents used Google Translate. This shows that the generative AI tools are being used more than the translation-related tools.

1. Identification of Linguistic Barrier



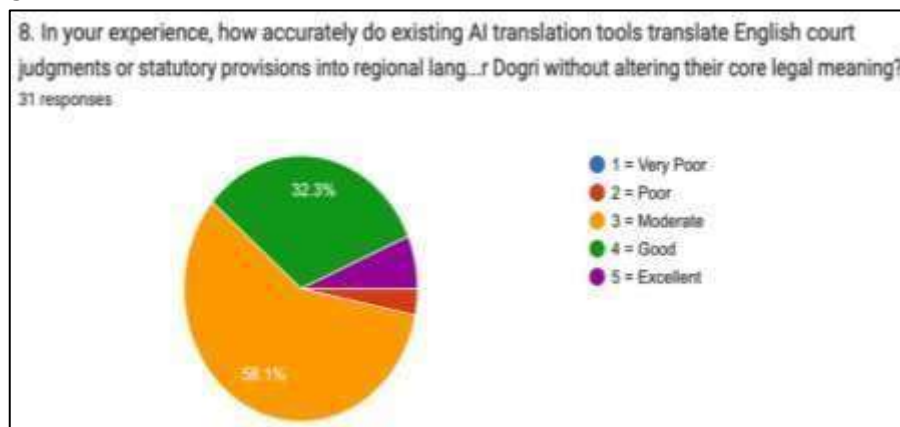
At the same time, the responses highlight the continuing linguistic barriers in accessing legal information. English was identified as the primary language of legal instruction by 96.8% of respondents and only 3.2% of Respondents identified Hindi.



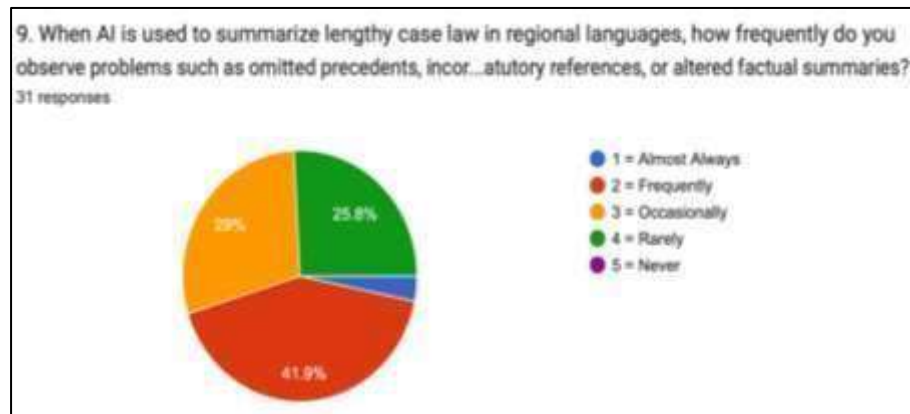
This becomes crucial when taking into consideration that over 45.2% considered reliance on English-language legal documentation to be a moderate barrier for vernacular-speaking litigants in the Jammu region. 25.8% of the respondents considered it a significant barrier and over 6.5% of respondents also classified it as a critical/high barrier.

2. Reliability of AI tools

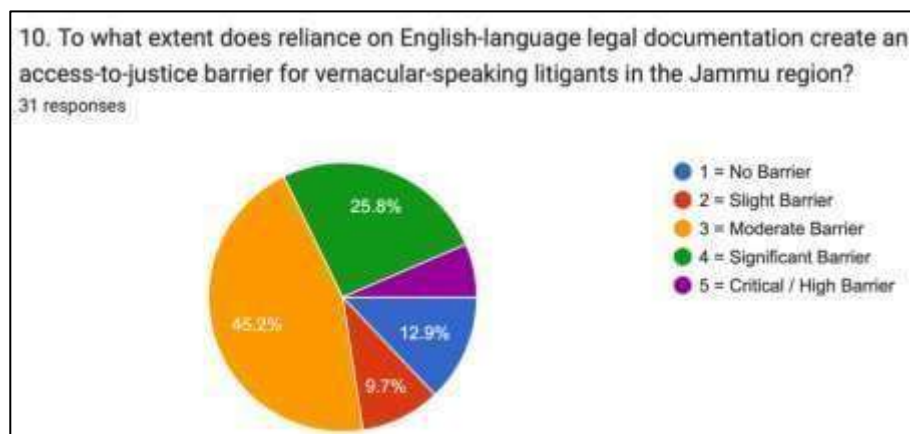
The research demonstrates that accessibility is closely tied to reliability. While 58.1% of the participants found the accuracy of existing AI-based translation systems to be moderate, and 32.3% of respondents found it to be good.



Likewise, 41.9% of the respondents said that they had often faced problems relating to AI-based legal summarization, such as omission of precedents, errors in statutory citations, and incorrect factual summarization. In addition, 29% of the respondents mentioned that they had faced such issues occasionally, whereas 25.8% had faced them rarely, and 3.2% very often.

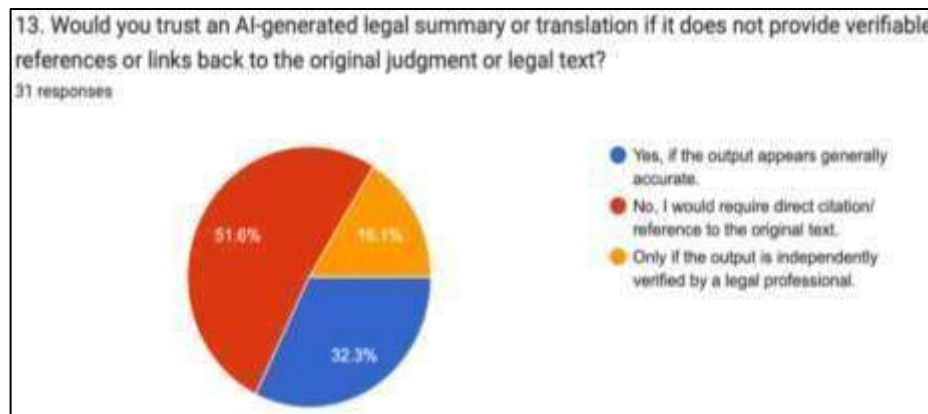


From the answers it is apparent that respondents believe it is vital for the AI summary software tool to be able to present the various aspects of the case which have legal significance separately and not just provide a summary. 51.6% of the respondents felt that this was very important, while 16.1% felt that it was absolutely necessary or mandatory for trust. Another 12.9% were neutral, while 16.1% believed it to be slightly important. Only 3.2% stated that almost always. This means that a good AI summary of a case would involve reducing its size as well as highlighting all of its elements.

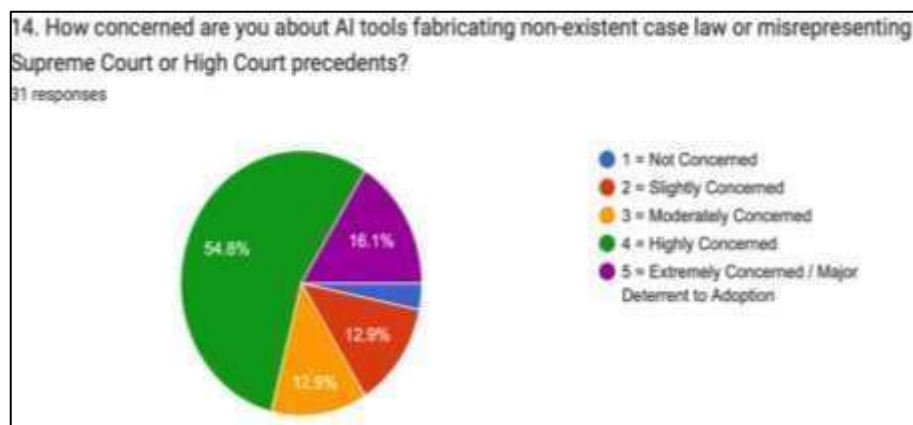


51.6% of the respondents said that they would need a citation to the original judgment or legal text before believing an AI-generated summary or translation of a legal document. 16.1% said that they would believe the AI-generated content only after it had been independently verified by a lawyer. The remaining 32.3 percent would have trusted the output if it seemed generally accurate.





Similar worries are echoed by the 54.8% of participants who had high levels of concern about the fabrication of case law or misrepresentation of judicial precedents using AI technology. Over 16.1% were extremely concerned to the point that it became a deterrent to adoption, 12.9% were slightly concerned and another 12.9% moderately concerned. Only 3.2% were not concerned. This shows that the use of AI to improve access to justice should be regarded as the tool that makes legal information accessible, especially between languages, but the results of which should be transparent and accurate.

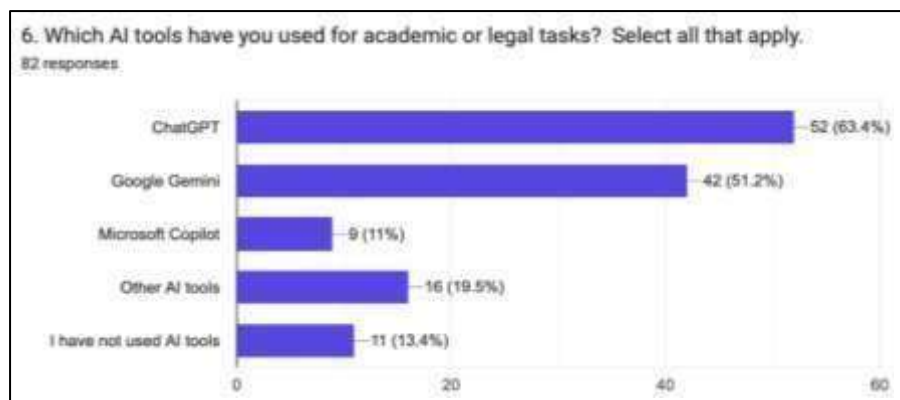


B. STUDENTS

The second step of the empirical component of this paper involved administering the Google Form to students. The responses provide a useful second dimension to the study as they show how AI is being experienced by the actual learners who are expected to use these tools to gather legal information and understand judicial decisions.

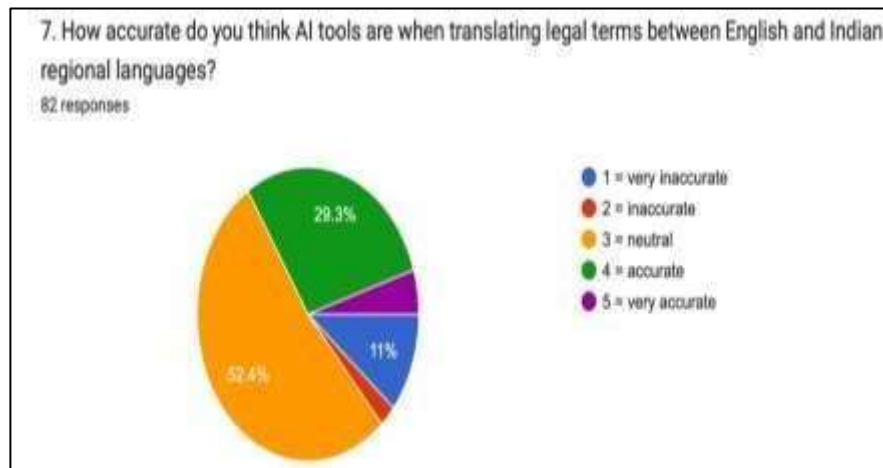
1. Frequency and Kinds of Tools Used

The findings show that AI has already become an important part of the academic practices of law students. 45.1% of the respondents reported using AI tools daily and another 30.5% reported using them weekly, 14.6% stated monthly/rarely while only 9.8% stated that they had never used such tools. ChatGPT was the most commonly used tool (63.4%), followed by Google Gemini (51.2%), Other AI tools (19.5%), Microsoft Copilot (11%) Interestingly, over 13.4% students stated that they have not used any AI tools.



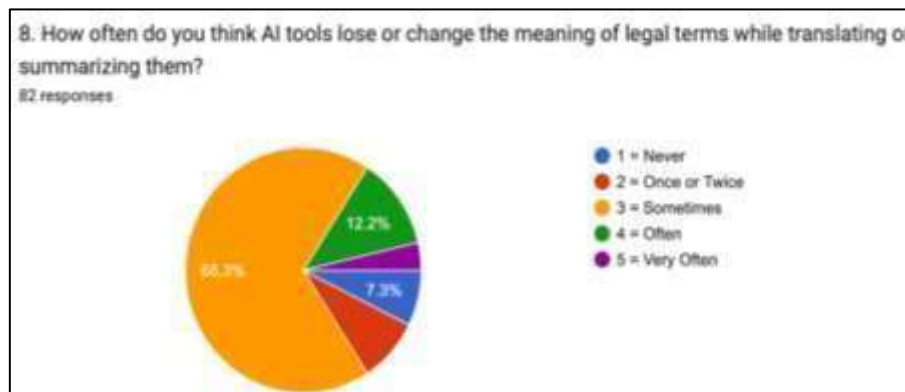
2. Identification of Language Barriers

The language barrier also emerged as an important concern, with 50% of students stating that language barriers sometimes make it difficult for people in Jammu and Kashmir to understand or access legal information. Further, 22% stated that this happens often and 8.5% stated that it happens mostly, while only 17.1% stated that language barriers do not create such difficulties. These concerns become particularly salient when viewed against the linguistic profile of Jammu and Kashmir, where speakers of Gojri and Pahari constitute a significant demographic. For such speakers, the barrier is not merely between English and an official language such as Urdu or Dogri, but also between English and their own mother tongue, for which AI-based translation and summarisation resources remain virtually non-existent.

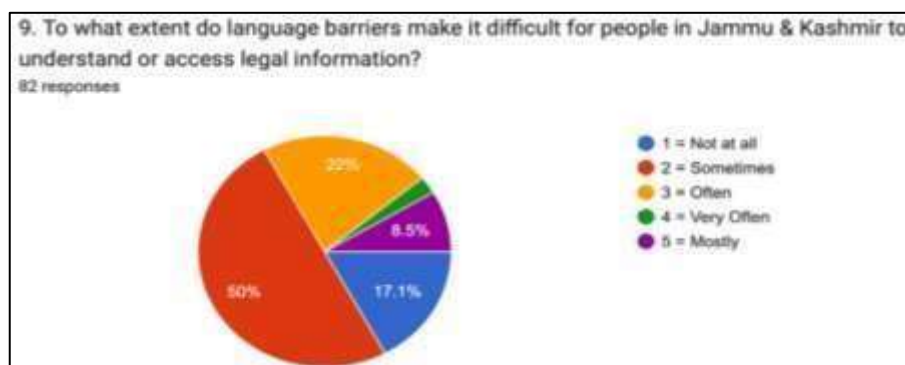


3. Reliability of AI tools

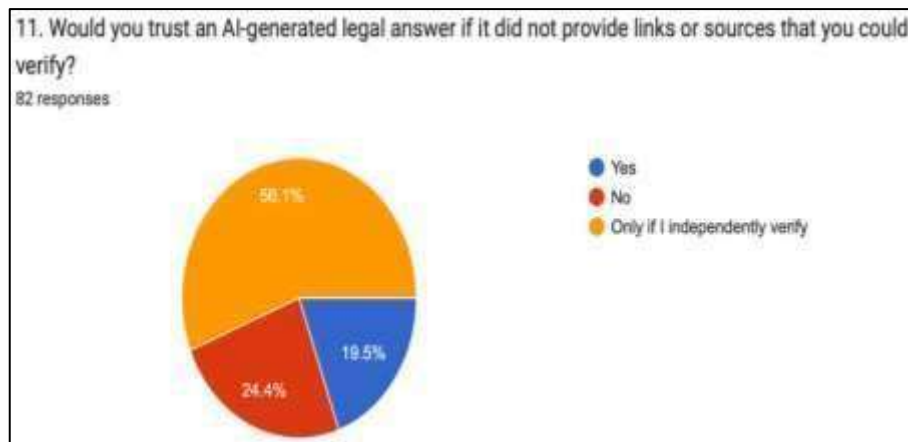
At the same time, the responses indicate that students remain cautious about the reliability of AI for legal work. The accuracy of AI tools when translating legal terms between English and Indian regional languages was then assessed. While 29.3% considered AI translation of legal terms accurate, a majority of 52.4% remained neutral about its accuracy. 4.9% stated that AI is very accurate while, 2.4% stated that it is inaccurate.



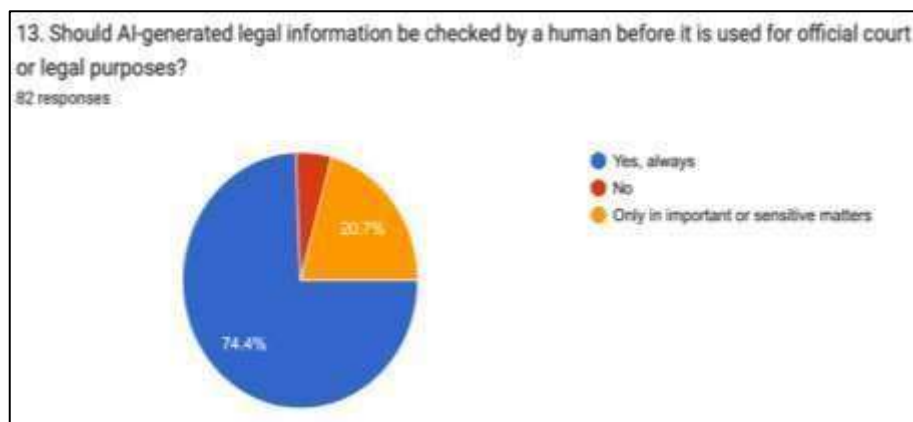
When asked whether AI tools lose or change the meaning of legal terms while translating or summarising them over 68.3% stated that it happens sometimes while 12.2% stated that it happens often and 7.3% stated that it never happens.



The responses further show that students do not view AI-generated legal information as something that can simply be accepted without verification. 56.1% would trust an AI-generated legal answer only if they independently verified it, while 24.4% would not trust it without sources or links.

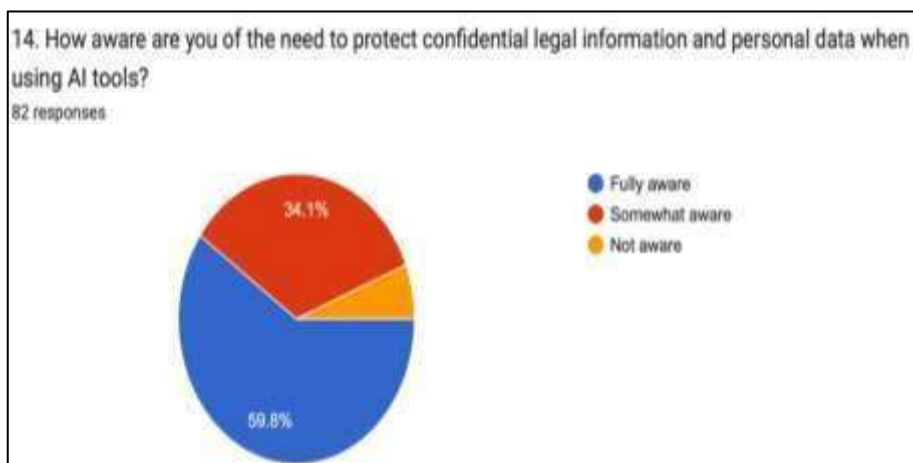


Further, 74.4% of the Respondents believed that AI-generated legal information should always be checked by a human before being used for official court or legal purposes. Further, 20.7% stated that in important and sensitive matters the information should be checked.

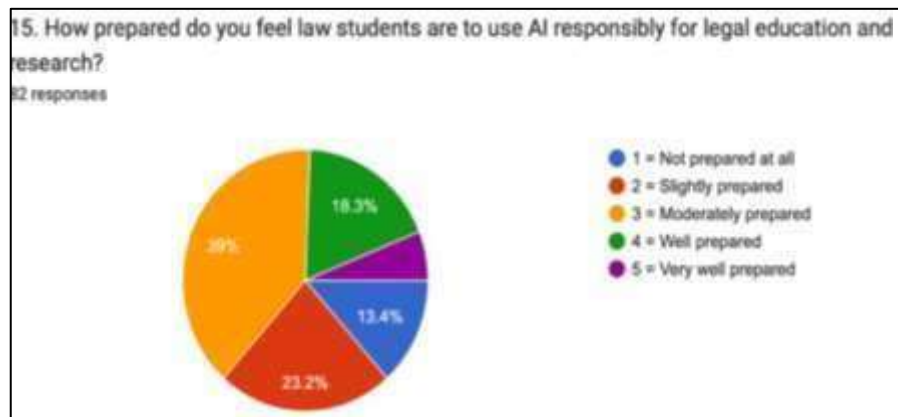


The responses also reveal a need for greater preparedness among students: only 6.1% considered themselves well prepared, 23.2% considered themselves slightly prepared and 13.4% not prepared at all to use AI responsibly, while 39% considered themselves only moderately prepared.

At the same time, 59.8% were fully aware of the need to protect confidential legal information and personal data when using AI tools. 34.1% were somewhat aware and only 6.1% were not aware.



These results indicate that students acknowledge the value of AI in legal studies, but they also understand that using AI requires verification, law interpretation, and increased understanding of responsible AI usage.



The survey result that 51.6% of teachers found it very important and another 16.1% of them thought that the separate identification of facts, key arguments, issues, ratio decidendi and final order was essential or even mandatory for trust is therefore quite meaningful. It shows that the worries of the respondents correspond with one of the problems identified in the Legal NLP research⁹. The legal summary should not treat every sentence of a judicial decision equally. Instead, it should recognize its role in the judgment and retain the parts of the decision which are significant legally. It becomes especially important when the user relies on the summary generated by the AI in order to understand the judgment and does not need to read the whole decision.

These empirical findings prove that the implementation of multilingual LLMs for access to justice should be viewed through the prism of the assistive model rather than the substitutional one. In other words, the AI could help the individual to understand the decision, identify the issues in it, translate the legal terminology or find the relevant part of the decision. Yet, the original decision should stay the official version and the AI-generated output should still be verifiable.

6. REGULATORY FRAMEWORK

The premise of this paper is that meaningful access to justice should not only be understood in terms of availability of legal remedies and off Court but also the ability of an individual to understand the law and judicial decisions which determine his or her rights and obligations. In this way, it flows right through the Constitutional mandate. Article 14 of the Constitution guarantees equality before the law,¹⁰ while Article 21 protects life and personal liberty.¹¹ Article 39A further provides that “The State shall secure that the operation of the legal system promotes justice, on a basis of equal opportunity, and shall, in particular, provide free legal aid, by suitable legislation or schemes or in any other way, to ensure that opportunities for securing justice are not denied to any citizen by reason of economic or other disabilities.”¹²

The problem of language becomes all the more pertinent in the light of Jammu & Kashmir. Jammu and Kashmir Official Languages Act, 2020 acknowledges Kashmiri, Dogri, Urdu, Hindi, and English as the official languages of the Union Territory.¹³ On the other hand, English continues to maintain its relevance in the higher judiciary and legal documents. There emerges a linguistic gap between the language in which the legal data is made available and the language in which the people feel comfortable comprehending such data. Translation through AI can become one way of bridging this linguistic gap. However, the translation should retain the meaning of the original legal data and cannot be considered as an independent source of law.

⁹ Shubham Kumar Nigam, Tanmay Dubey, Govind Sharma, Noel Shallum, Kripabandhu Ghosh & Arnab Bhattacharya, “LegalSeg: Unlocking the Structure of Indian Legal Judgments Through Rhetorical Role Classification”, Findings of the Association for Computational Linguistics: NAACL 2025 1129–1144 (2025), DOI: 10.18653/v1/2025.findings-naacl.63, available at: <https://aclanthology.org/2025.findings-naacl.63/>.

¹⁰ The Constitution of India, 1950, art. 14.

¹¹ *Id.*, art. 21.

¹² *Id.*, art. 39A.

¹³ The Jammu and Kashmir Official Languages Act, 2020 (Act No. 23 of 2020), s. 3.

The language barrier in the higher judiciary is also reflected in the constitutional position under Article 348,¹⁴ under which English continues to be the language of proceedings before the Supreme Court and, subject to the constitutional framework, the High Courts.¹⁵ It is important to note that the 2018-19 Annual Report of the Supreme Court recognized the importance of translations into vernacular languages in certain categories of cases where the litigants might be less familiar with the English language¹⁶. This includes labour, consumer protection and eviction matters.

The Indian Judiciary has already started using AI for the same. The Supreme Court has come out with the Supreme Court Vidhik Anuvaad Software (SUVAS) for translating judicial documents, and the Supreme Court has provided this software to High Courts to translate their judgments.¹⁷ In addition, the Supreme Court has also recognised that human intervention would be required to validate the output of the AI in terms of translation. In addition, it has been reported by the Government that the application of AI in translation has also been used to translate Supreme Court judgments into Indian languages.¹⁸

Furthermore, the Supreme Court has begun to produce Landmark Judgment Summaries to facilitate the comprehension of critical judgments.¹⁹ Notably, the Supreme Court itself makes clear that the summary is not part of the judgment and the reason of the Court and is not to be used for legal purposes. This point becomes highly significant for the current research. It shows that accessibility and legal force are two different things. A summary can help one to comprehend the meaning of the judgment, but it does not affect its legal force in any way.

The problem of accuracy acquires special significance in light of the Supreme Court's judgment in *Pooja Ramesh Singh v. Jammu & Kashmir Bank Ltd.*²⁰ This judgment related to reliance on fabricated AI-produced case law by NCLT and NCLAT. The Supreme Court took a zero-tolerance attitude to the use of AI-created or hallucinated materials as authority without any verification and stressed the necessity of controlling the use of AI in the process of adjudication. At the same time, the Supreme Court made it clear that its judgment did not restrict the proper use of AI; it only addressed the problems of providing false or hallucinated materials as genuine authorities. It should be noted that this judgment is directly related to the research under consideration since it proves that the worries about fabricated authorities voiced by teachers and students are not mere theoretical concerns.

The use of AI technologies is also a cause for concern in relation to the use of personal and confidential data. Digital Personal Data Protection Act, 2023²¹ sets out the processing of digital personal data and recognizes the right of individuals to the protection of personal data. It becomes important in cases where legal documents with personal data are uploaded to external AI software. What becomes problematic here is not the fact that using commercial AI tools is forbidden per se, but the consideration whether the uploaded data may be processed by a specific system in accordance with the law.

Accordingly, any form of regulation of AI in the legal domain needs to take into account the distinct functions of AI. While AI may prove useful in the context of translation, summarization, retrieving and explaining legal information and thus making the law more understandable, AI-generated material, unverified as being the law, entails higher risks. The recent approach of the Supreme Court, too, acknowledge this distinction. Through the Draft Regulations for Use of Artificial Intelligence (AI) in Courts,

¹⁴ The Constitution of India, 1950, art. 348.

¹⁵ Supreme Court Observer, "Language Barriers", 24 June 2020, *available at*: <https://www.scobserver.in/journal/language-barriers/>.

¹⁶ Supreme Court of India, Annual Report 2018–19, at 88, *available at*: <https://cdnbbsr.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/uploads/2024/07/2018-19.pdf>.

¹⁷ In Re: Policy Strategy for Grant of Bail, Suo Motu Writ Petition (Criminal) No. 4 of 2021, Order dated 8 May 2025, Supreme Court of India, *available at*: https://api.sci.gov.in/supremecourt/2021/26546/26546_2021_4_17_61634_Order_08-May-2025.pdf.

¹⁸ Department of Justice, Ministry of Law and Justice, Government of India, "Measures to Translate and Publish Proceedings and Judgments of Supreme Court and High Courts", Rajya Sabha Unstarred Question No. 435, answered 28 November 2024, *available at*:

<https://www.doj.gov.in/static/uploads/2025/09/04d247ddf5a1a6df49e2e61b32b58f15.pdf>

¹⁹ Supreme Court of India, "Landmark Judgment Summaries", *available at*: <https://www.sci.gov.in/landmark-judgment-summaries/>.

²⁰ *Pooja Ramesh Singh v. Jammu & Kashmir Bank Ltd.*, Civil Appeal No. 11950 of 2025, 2026 INSC 668 (Supreme Court of India, decided July 2, 2026).

²¹ The Digital Personal Data Protection Act, 2023 (Act No. 22 of 2023).

2026,²² the Supreme Court has proposed a framework which recognises the need for human supervision, transparency, accountability, and protection of data when AI is used in the courts. Therefore, the Draft Regulations consider AI as a tool assisting in legal and judicial proceedings rather than a substitute for human judge.

7. Limitations Of The Study

There are a few limitations to this study. Empirically, the study was conducted in the Union Territory of Jammu and Kashmir with 31 teachers and law educators and 82 students participating in the survey. The results of the study can only be said to indicate the perceptions of the participants and cannot be generalised to the whole Indian legal education and judicial system. Second, this study is purely descriptive and looks into the use and perception of AI and does not test independently the technical accuracy of any one AI model. Third, this study does not conduct any technical comparison between various LLMs and translation systems.

8. SUGGESTIONS

Based on the findings of the study, the following suggestions may be considered.

1. First, AI-generated legal summaries and translations should provide a direct reference to the original judgment or legal provision. This would allow the user to verify the information and would prevent the AI-generated output from being treated as an independent source of law.
2. Second, AI summarisation tools used for legal judgments should separately identify the facts, issues, arguments, applicable law, reasoning, ratio decidendi and final order. This would make the summary more useful for legal understanding and would reduce the possibility of mixing the legally binding reasoning of the Court with other observations contained in the judgment.
3. Third, there should be a human-in-the-loop system for AI-generated legal translations and summaries which are intended for official or professional use. The role of the human reviewer should include verification of case citations, statutory provisions, legal terminology and the meaning conveyed through translation.
4. Fourth, AI literacy should be incorporated into legal education. The responses show that students are already using AI regularly, but their preparedness to use it responsibly remains comparatively limited. Legal education should therefore include training on verification of AI-generated authorities, identification of hallucinations, protection of confidential information and the limitations of AI-generated legal information.
5. Fifth, greater attention should be given to the development of multilingual legal tools for Indian languages. In the context of Jammu and Kashmir, the development of tools capable of dealing with Dogri, Kashmiri and Urdu, along with Hindi and English, can contribute towards making legal information more accessible. Beyond the officially recognised languages, tools should also be developed for Gojri and Pahari, which are spoken by significant communities in the Union Territory but presently lack adequate digital and legal-linguistic resources. Such tools should, however, be evaluated not only for linguistic accuracy but also for their ability to preserve the legal meaning of the original text.
6. Sixth, institutions should formulate clear guidelines regarding the use of external AI systems for legal documents containing personal or confidential information. Students, teachers and legal professionals should be informed about the risks involved in uploading such material to third-party systems and should be encouraged to use institutionally approved and secure systems wherever available.
7. Finally, the performance of multilingual legal AI should be evaluated through legal experts as well as technical metrics. The literature already demonstrates that automated measures alone may not adequately capture the legal quality of a summary. Future systems should therefore be assessed on accuracy, preservation of legal meaning, completeness, correct identification of legal roles and reliability of citations.

9. Conclusion

This research was conducted with the purpose of examining how AI and multilingual Large Language Models can help in bridging the vernacular divide in access to legal information in Jammu and Kashmir. It turns out that currently AI is actively utilized by legal educators as well as by students. Nevertheless, the respondents acknowledge that there are still difficulties associated with understanding legal information because of language differences and that multilingual AI can be a possible solution to this problem. Nevertheless, it turns out that accessibility cannot be considered as the only goal in itself. The point is that legal information needs to become more accessible without any alteration of its original meaning. As it can

²² Supreme Court of India, Draft Regulations for Use of Artificial Intelligence (AI) in Courts, 2026, Notice dated 3 June 2026 inviting comments and suggestions, *available at*: <https://www.sci.gov.in/notices-and-circulars/>.

be seen from the concerns of the respondents, related to inaccurate translations, missing precedents, wrong citations of statutes and fake cases, the use of AI in the legal domain requires additional care.

Thus, the study lends support to a model of an assisting bridge that is played by AI in connecting the individual to the law rather than AI being the substitute for the law and human judgment. AI can interpret and translate the legal information but the original judgment should always be the authority and the outputs of AI should also be verifiable. The current application of AI-assisted translation by the Indian judiciary and the decision of the Supreme Court of India towards AI generated hallucinations further reinforce the necessity of such a model. After all, the aim is not to increase the availability of legal information using the AI but to make the law accessible and reliable. Thereby, facilitating meaningful access to justice.

The advent of such multilingual AI systems in Jammu & Kashmir that are developed through human verification and legal interpretation while keeping confidentiality of personal information, could be a valuable step towards reducing the vernacular gap in Jammu and Kashmir.