Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Knowledge Graph-Based LLM Framework to Identify Potential Contradictions and Regulatory Gap in Environmental Laws of India

Prathamesh Poul
Department of Artificial Intelligence
and Machine Learning
Vishwakarma Institute of Information Technology
Pune, India

Sonal Fatangare
Department of Artificial Intelligence
and Machine Learning
Vishwakarma Institute of Information Technology
Pune, India

***Abstract*—Indian environmental law is dispersed among various acts, rules and bodies. Provisions from various sources may overlap, may duplicate each other in a duty, or may not have a clear owner of the task. These cases can only be found by hand, and knowledge of law. We present a framework with the following components: (1) storing the legal provisions as entities and relationships in a knowledge graph; (2) using graph rules and text similarity to identify related provisions; (3) asking a large language model (LLM) to explain selected candidate pairs.**

The framework was applied to 15 provision pairs of environmental laws in India, giving 105 unique provision pairs for testing purposes. Using preliminary AI annotation, 8 pairs were annotated as positive and 97 as negative. The five approaches we compared were knowledge graph rules, knowledge graph with the LLM's assistance, TF-IDF similarity, keyword overlap, and a hybrid. The highest F1 (0.667) was obtained by TF-IDF, and the hybrid model achieved a recall of 0.750.

These results are preliminary. The data set is small, and there is no legal expert review of the labels. We view the framework as a tool to help human reviewers, and not a guidance system that can replace their review of the law.

***Index Terms*—Environmental legislation, Knowledge graph, Large language models, Legal text mining, Gaps in the law, contradiction detection, Indian environmental law, NLP.**

I. Introduction

Environmental protection in India does not sit in a single law. The main instruments are the Environment (Protection) Act, 1986, the Water (Prevention and Control of Pollution) Act, 1974, the Air (Prevention and Control of Pollution) Act, 1981, the Forest (Conservation) Act, 1980, and the Wild Life (Protection) Act, 1972 [1].

Each statute covers a different area, but their scopes overlap. One industrial activity may involve environmental clearance, pollution control, forest permissions, or wildlife-related restrictions. To see how the relevant provisions fit together, a reader has to compare their scope, the authorities they name, the obligations they set, and how they are enforced.

Doing this by hand is slow. The language is formal, provisions refer to one another, and related ideas are often expressed in different terms, so a keyword search can miss two provisions that cover the same activity. A missing authority or implementation timeline may also turn out to be defined somewhere else, which means checking the surrounding legal context.

Language models have made it more realistic to analyse long legal documents. Transformers capture context within text [2], and LEGAL-BERT shows that adapting a model to legal documents pays off [3]. Knowledge graphs offer a different tool: they record entities and the relationships between them in a structured form [4].

Using an LLM alone carries risk, because it can reach unsupported conclusions or invent legal interpretations. A knowledge graph gives the model structured facts about provisions, authorities, and obligations to work from, and text similarity points to provisions on related subjects. Used together, these make the process easier to inspect.

This paper proposes a knowledge graph-grounded LLM framework for the preliminary detection of potential contradictions, overlapping responsibilities, and structural signs of regulatory gaps in Indian environmental legislation. The framework does not decide whether a provision is invalid, legally contradictory, or unenforceable. It produces candidate relationships that someone with legal training should examine.

A. Research Objectives

Our objectives were to represent selected Indian environmental provisions in a knowledge graph; to find related provisions through graph relationships and textual similarity; to test whether an LLM can examine those candidates in context; to compare graph-based, text-based, and hybrid approaches using preliminary metrics; and to see whether combining structured and unstructured information improves the identification of related provisions.

B. Research Contributions

The paper makes four contributions. First, it gives a structured representation of selected environmental provisions using legal entities and relationships. Second, it presents a multi-stage framework that combines knowledge graph reasoning, textual

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

similarity, and LLM-assisted analysis. Third, it compares five approaches on a small preliminary dataset. Fourth, it discusses the limits of automated legal relationship detection and why expert validation is needed.

II. Related Work

A. Legal NLP

Increasingly NLP applications are found in the area of legal document classification, legal information retrieval and judicial decision analysis. In classical retrieval systems, the terms in the document are given weight according to the TFIDF distribution to identify documents with common vocabulary [5]. These techniques are easy to operate but do not capture the link if two provisions refer to the same concept in different terms.

In the transformer models, contextual representation was improved. Most existing language models are based on the architecture that was introduced by Vaswani et al. [2] and BERT added bidirectional pre-training [6]. The Sentence-BERT generates sentence representations which are used to compare sentences semantically [7].

There is a slightly different vocabulary and structure in legal text, which some models have been designed to take account of. The advantage of domain-specific pre-training is illustrated by LEGAL-BERT, a pre-trained model that improves legal NLP tasks [3] and LexGLUE is a benchmark for multiple legal understanding tasks [8]. Both emphasize the importance of domain specific data and assessment.

In the field of Structured Representation, Knowledge Graphs are used extensively.Knowledge Graphs are widely used in the field of Structured Representation.

A knowledge graph is a graph-based representation of information, in which the nodes are entities and the edges are relationships. It can be used in a legal context to connect provisions, regulatory bodies, activities, obligations, and statutory instruments. Hogan et al. provide an excellent introduction to the concepts and uses of [4].

The methods used to build ontologies are used to identify entities, attributes and the relations between them in the system [9] and both RDF and OWL provide formal means to represent these [10], [15].

Selected provisions are used to represent a knowledge graph and the connections between provisions. This allows provisions to be identified that share an authority, activity, obligation or subject area.

In this section, we primarily examine the application of LLM in legal analysis.In this section, we mainly focus on the application of LLM in legal analysis.

LLMs can summarize, answer questions, and reason over context. Retrieval-augmented generation introduces a retrieval step to better ensure that answers are informed by external knowledge [11].

Benchmarks include IL-TUR [12] and LawBench [16] which evaluate language models on legal knowledge and reasoning. Nevertheless, an explanation of an LLM is not yet a legal conclusion. The model might misinterpret the meaning of the statute, miss one of the exceptions, or make an interpretation that is not supported by anything in the statute.

We use the LLM as an aid, but not as a decision maker, for that reason. Structured and textual methods suggest candidate relationships first and then the LLM considers them within the context.

B. Research Gap

Previous studies have demonstrated that NLP, domain-specific language models and knowledge graphs are helpful for legal information processing. However, determining overlaps and structural gaps in multiple environmental laws requires more than one type of information.

Without knowing who made the provisions or the relationship between them, a text-only approach may be able to match similar wording. A graph only approach may overlook relationships that are expressed in alternative legal terminology. Conclusion without support is possible to produce with an LLM.

Our framework addresses these weaknesses by employing a structured legal representation, graph-based candidate generation, textual similarity between sentences, and LLM-based analysis.

III. Proposed Methodology

A. Framework Overview

The framework includes 5 phases: legal provisions collection and preprocessing, building knowledge graph, building candidate relationship, similarity analysis in text and LLM-assisted analysis, and comparison. It can generate candidate relationships between provisions, highlighting areas where some provisions are similar and some are different in the structure of the provision.

Fig. 1 shows the knowledge graph representation.

B. Legal Provision Collection

The data set consists of 15 provisions being paraphrased from five environmental laws from India, namely, the Environment (Protection) Act, 1986 [1]; the Water (Prevention and Control of Pollution) Act, 1974 [17]; the Air (Prevention and Control of Pollution) Act, 1981 [18]; the Forest (Conservation) Act, 1980 [19]; and the Wild Life (Protection) Act, 1972 [20].

Since then the Forest Act, 1980 has got changed and is now called Van (Sanrakshan Evam Samvardhan) Adhiniyam, 1980. Each user of a specific version or amendment date is required to always check the version and date of the text of law that is to be interpreted under the framework.

Each provision was paraphrased to ensure the major regulatory meaning of the provision. All come with provision identifier, legislation, regulated activity, responsible authority, legal obligation and subject of the regulations.

The experiments reflect a technical evaluation and are not a full set of primary statutory text, but rather paraphrased.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

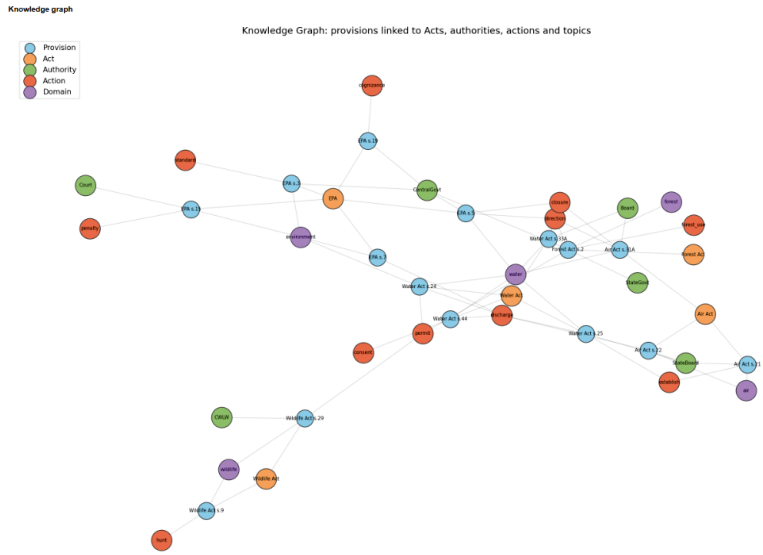


Fig. 1. Knowledge graph representation of the selected environmental legal provisions, illustrating the relationships among environmental laws, regulatory authorities, regulated activities, and associated legal obligations.

C. Knowledge Graph Construction

Selected provisions and their regulatory attributes are symbolized as the knowledge graph. It contains six types of entities: Act (environmental Act), Provision (chosen legal provision), Authority (regulatory or administration body), Activity (regulated environmental activity), Obligation (legal obligation), and Subject (environmental domain).

Labelled edges are the connections between entities: A connection between a provision and its Act takes the form of `BELONGS_TO`; a connection between a provision and an activity takes the form of `REGULATES`; a connection between a provision and an authority takes the form of `ASSIGNED_TO`; a connection between a provision and an obligation takes the form of `REQUIRES`; and a connection between a provision and a regulatory subject takes the form of `RELATED_TO`.

A provision is thus a group of related things and not just a single line of text. This enables to determine provisions that have activities, authorities, or obligations in common.

D. Candidate Relationship Generation

The framework would be utilized for comparison of provisions in order to find possible relationships with candidates. The number of different unordered pairs of n provisions is:

$$N = \frac{n(n-1)}{2} \tag{1}$$

For our dataset:

$$N = \frac{15(15-1)}{2} = 105 \tag{2}$$

This leaves us 105 different provision pairs to look at. During the candidate phase, each pair will explore, common graph relationships, regulatory properties and text.

E. Graph-Based Rule Detection

Graph based rules are rules which compare two provisions with entities that are associated with them. If they are provisions of the same activity then it may be considered whether they might refer to different authorities, for example.

The graph-based candidate score, simplified, would be:

$$S_G(P_i, P_j) = w_a A_{ij} + w_r R_{ij} + w_o O_{ij} \tag{3}$$

$$S_{ij} = w_A A_{ij} + w_R R_{ij} + w_O O_{ij} \tag{4}$$

Here, A_{ij} represents the shared authority information, R_{ij} represents the shared regulatory activity, and O_{ij} represents the shared obligation information. The weights w_A, w_R, and w_O represent their respective contributions to the total score.

Only ”flag” those provisions that may have a relationship. It doesn't prove that they don't meet the law.

In this part, we will be talking about TF-IDF based Textual Similarity.

TF-IDF is a method to measure how similar the text of two provisions is. The weight of a term depends on its distribution in a document and its frequency in the document's corpus [13]. In a document d for a given term t:

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \tag{5}$$

An inverse document frequency is:

$$IDF(t) = \log\left(\frac{N}{DF(t)}\right) \tag{6}$$

In this case N is the total number of documents and $DF(t)$ is the number of documents that contain term t.

The cosine similarity between two provision vectors is calculated:

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

TABLE I
EXPERIMENTAL DATASET SUMMARY

Dataset Attribute	Value
Number of statutes	5
Number of provisions	15
Total unique pairs	105
Positive pairs	8
Negative pairs	97
Positive proportion	7.62%
Negative proportion	92.38%

$$Sim(P_i, P_j) = \frac{V_i \cdot V_j}{\|V_i\|\|V_j\|} \tag{7}$$

where V_i and V_j are the TFIDF vectors of provisions i and j. The higher the score the more the word is similar. There is no contradiction or overlapping in regulation if there is a similarity in wording.

F. LLM-Assisted Contextual Analysis

Selected pairs of candidates are examined by the LLM. It brings paraphrased provisions and the structured information from the knowledge graph. We question it when it regulates an activity that is related, when responsibilities appear to overlap, when other authorities have similar responsibilities, when it appears the text doesn't cover something out, and when it seems it needs further legal investigation.

The output is not meant to be an authoritative interpretation of the Indian law, but rather an explanatory evaluation. It is better to give the model context, according to the motivation in retrieval-grounded generation ref11 which we have done, to make the model more friendly to the context.

A potential regulatory gap identification is conducted.A Potential Regulatory Gap Identification is performed.

The framework also looks for evidence of regulatory gaps in structure – there are two indicators.

The first one is an unassigned authority: a provision is a regulatory responsibility or activity that is mentioned in a provision but not by the provision's extracted representation. The second one is a *missing implementation timeline*: an obligation without explicit provision of a time line.

The two indicators are just hints to look into. In such cases, they do not exhibit a statutory gap, as there might be other information available in the Act, the rules and regulations, the notifications or the instruments.

IV. EXPERIMENTAL SETUP

A. Dataset Description

The data set consists of a set of five environmental laws with 15 paraphrased provisions. Each provision was compared to each of the other provisions, resulting in 105 different pairs. Preliminary annotation – 8 positive, 97 negative.

Labels were assigned through preliminary annotation assisted by artificial Intelligence. They are not legal ground truth, but rather the metrics below are interpreted in that context.

TABLE II
COMPARATIVE PERFORMANCE EVALUATION OF DIFFERENT APPROACHES

Approach	Accuracy	Precision	Recall	F1-Score
KG Rules	0.914	0.400	0.250	0.308
KG + LLM	0.933	0.667	0.250	0.364
TF-IDF	0.952	0.714	0.625	0.667
Keyword Overlap	0.867	0.350	0.875	0.500
Hybrid	0.924	0.500	0.750	0.600

B. Evaluation Approaches

We considered the following five approaches: *KG Rules* identifies candidate pairs by leveraging knowledge graph relationships and predefined rules, *KG + LLM* uses knowledge graph candidates and adds LLM-assisted contextual examination, *TF-IDF* compares the textual vectors by cosine similarity, and *Keyword Overlap* counts common keywords between the two provisions, and *Hybrid* integrates graph-based candidate information with text analysis and contextual examination.

KG + LLM and Hybrid are two different sets up. The first is for candidates generated from a graph using an LLM, and the second is for more than one candidate-identification signal.

C. Evaluation Metrics

Accuracy, precision, recall and F1-score are reported.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

This is because true positive, true negative, false positive, and false negative are represented by the symbols TP, TN, FP, and FN respectively.

The dataset is imbalanced; there are only 8 out of a total of 105 positives, therefore accuracy is not very telling. A method which gives every pair a negative rating will have a correlation of more than 0.92. Precision, recall, and F1 score provide a more comprehensive comparison [14].

V. RESULTS AND DISCUSSION

A. Comparative Performance

Table II gives the performance of the five approaches.

TF-IDF got the highest accuracy and F1 score. The high value of F1 (0.667) is a better balance of precision and recall than any other method in this experiment.

The results showed that keyword overlap recalled best (0.875) with a precision of 0.350. It produced many false positives, and found most of the positive pairs.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

TABLE III
CONFUSION MATRIX COUNTS FOR THE EVALUATED APPROACHES

Approach	TP	FP	FN	TN
KG Rules	2	3	6	94
KG + LLM	2	1	6	96
TF-IDF	5	2	3	95
Keyword Overlap	7	13	1	84
Hybrid	6	6	2	91

The hybrid resulted in a recall of 0.750 and an F1-score of 0.600. This is better than both KG-based configurations for recall, and close to TF-IDF for accuracy and F1, but is still under TF-IDF.

KG Rules scored an F1 of 0.308 and KG + LLM scored 0.364. However, on this dataset, the structured relationships alone did not encompass all the pertinent textual similarities. In the KG setup, adding the LLM did not affect recall, but did improve precision.

B. Confusion Matrix Analysis

The confusion matrix counts for all 5 approaches are given in Table III.

The scores for the positive pairs and the misses are shown, and KG Rules missed six and found two. Two additional ones were found by KG + LLM, with the number of false positives reduced from three to one.

TF-IDF had the highest F1 of the five with five positive pairs and two false positives.

The seven pairs of positive matches were identified but also 13 false positive matches were raised with keyword overlap. Lacking any other support, shared terminology is too shallow of a criteria for identifying legal relationships.

The hybrid approach identified six positive pairs, but failed to identify two, and generated six false positives. It operates at a higher recall when compared to the KG based configurations.

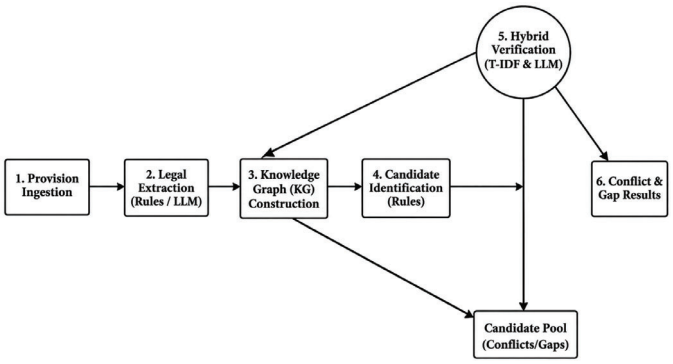
Typical precision versus recall trade-off. If you have a method designed to detect more potential problems, you'll get more pairs that aren't problems, and someone needs to review them.

C. Structural Candidate Analysis

During the structural analysis stage, 13 candidate relationships were identified. Five were linked to potential conflict and eight to potential structural gaps.

17 provision pairs were analysed separately, with 12 provision pairs requiring further consideration in the hybrid analysis. The 13 structural candidates and the 12 pairs of candidates flagged as hybrids are not the same results as the other candidates because they are from different stages and candidate sets.

The structural candidates identify areas where provisions can share activities, responsibilities and/or obligations. Gap candidates are gaps which do not have an explicit authority



Representative results of the knowledge graph and LLM-assisted analysis workflow.

or implementation time. Neither kind establishes a real contradiction or omission in the law. Both must be reviewed in conjunction with the entire legislation and other instruments.

Knowledge Graph and output from LLM

The knowledge graph represents the relationships between provisions, activities, authorities, and obligations. The outputs of the knowledge graph and LLM-assisted analysis are shown in Fig. V-C.

The graph structure can show why two provisions were picked for further analysis. The contextual description of each candidate relationship in the LLM stage allows a reviewer to have a better idea of how two provisions might relate. The descriptions still need to be compared to the original statutes.

D. Pairwise Similarity Analysis

The pairwise similarity analysis provides another perspective of the relationship of provisions in text.

The higher the number, the more text is shared between two provisions under the selected similarity measure; the lower the number, the less the two provisions share. It can be helpful to identify groups of similar provisions in the heatmap. Similarity is not legal equivalence, though. Two provisions may have the same wording but different obligations and two others may have very different wording but the same obligations.

VI. DISCUSSION

Both methods selected covered different aspects of the relationships between provisions.

TF-IDF performed best on this data set, indicating that it was useful for distinguishing the annotated positive pairs, meaning that vocabulary was useful. We would not extrapolate that in general beyond the current case, because the data are limited and are paraphrased provisions.

The highest recall was obtained with keyword overlap. This might be appropriate for an initial screening phase where the expense of missing the relevant provision outweighs the benefit of screening other candidates and the precision is not high enough that it should be used as it is likely to require some filtering.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

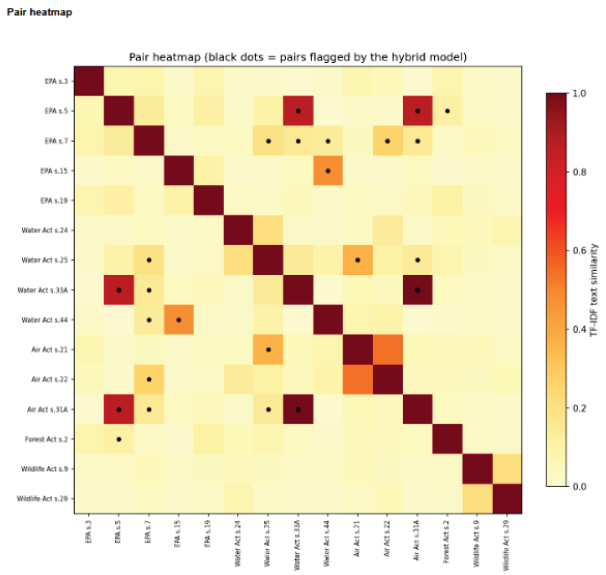


Fig. 2. Pairwise similarity heatmap showing the textual similarity between the selected environmental legal provisions.

The KG-based approaches provide a structured perspective on regulatory relationships and may reveal shared authorities, activities and obligations that may not be apparent in unstructured text. The accuracy and completeness of the graph are crucial to their results.

Precision of KG + LLM was improved while recall remained the same. The LLM might reject some candidates generated by the graphs, and the results are not necessarily improving on all measures.

The hybrid approach achieved the highest recall performance compared to both KG-based approaches, suggesting that the combination of structured and textual signals can identify more of the positive pairs annotated. It was not able to outperform TF-IDF on F1 and accuracy. The method used here was not sufficient, and the methods complement one another.

A. Practical Significance

A system of this nature may contribute to the legal researcher and to the environmental policy analyst to better understand the implications of a large number of provisions and determine whether there may be anything in common. It might indicate provisions that are similar to those in the other act, areas of administrative responsibility that might overlap, provisions that may be missing details about implementation, retrieval across statutes, and structural evidence that can be used as a starting point for human reviewers.

It is used as a preliminary screening and decision support tool. It is not a substitute for legal interpretation and statutory analysis or legal advice.

Limitations and future work are discussed.

B. Dataset Size

There are only 15 paraphrased provisions and 105 pairs in the dataset. This is not enough to demonstrate the effectiveness of the framework in the overall context of Indian environmental laws. The complete statutory provision, rules, amendment and notification is required for a fairer test.

C. Annotation Quality

These 8 positive and 97 negative labels are preliminary labels generated by AI which have not been validated by a legal expert. The reported metrics thus reflect only the level of agreement on those annotations, not the correctness of the legal content. Further experiments should be conducted with labels and criteria that were judged by experts, as well as with clarity of overlap, contradiction, and gaps in regulation.

D. Paraphrased Legal Text

Legal qualifications, exceptions, conditions and cross-references may be omitted when paraphrasing. All future work should include use of the original statutory language, section numbers, amendment history, definitions, and subordinate legislation.

The work of the group revealed contextual and temporal limitations.The group identified contextual and temporal limitations.

Environmental laws are subject to amendment, rules, notifications and interpretation by courts and there are certain provisions that may take various meanings. A temporal model of these changes is yet to be developed within the framework. Later editions should include dates of amendments, commencement dates, repealed provisions, and cross-references between principle Acts and subordinate legislation.

E. Future Research Directions

We then intend to add more of the original statutory provisions and obtain expert reviewed overlap and potential gap labels. Modelling: we would like to experiment with sentence representations for legal domain sentences and language models (LMs) that take the context into account, and to work on

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

completing the graph using automatic entity and relation extraction. We will also incorporate temporal information to capture amendments and changes in the legal responsibility for the content and evaluate using separate training and testing sets; and report on other metrics that are appropriate to imbalanced data. Last, we wish to create an interface for legal profession to accept or reject generated candidates.

These measures would prove if the structure withstands larger and more complex sets of Indian environmental law.

VII. Conclusion

We developed a knowledge graph based LLM system to detect the potential contradiction, overlapping responsibilities and structural indicators of regulatory gaps in the Indian environmental legal landscape. It integrates a knowledge graph, graph-based rules, TF-IDF similarity, keyword overlap, and LLM analysis.

On 15 paraphrased provisions from five statutes (105 pairs), TF-IDF had the best F1-score (0.667) and accuracy (0.952). At precision of 0.350, the highest recall was achieved on Keyword overlap (0.875), and the hybrid had a recall of 0.750. The LLM improved the accuracy of KG Rules from 0.400 to 0.667 without affecting recall (0.250).

In the structured relationships and text analysis it appears as though complementary signals can be provided for the identification of related provisions. However, the conclusions are constrained due to the small data set, the labels generated by the AI and the paraphrased text. The framework should be enriched with larger sets of data, annotated by experts, source documents, and temporal data of amendments before being used for practical legal analysis.

References

[1] Government of India, The Environment (Protection) Act, 1986, Act No. 29 of 1986, India.
[2] A. Vaswani ., “Attention is all you need,” in *Proc. Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.
[3] I. Chalkidis *et al.*, “LEGAL-BERT: The muppets straight out of law school,” in *Proc. Findings of EMNLP*, 2020, pp. 2898–2904.
[4] A. Hogan ., “Knowledge graphs,” *ACM Computing Surveys*, vol. 54, no. 4, pp. 1–37, 2021.
[5] This is a result of the collaboration between G. Salton and C. Buckley, Term-weighting methods for automatic text retrieval,” *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.
[6] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” is the title of the paper. in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
[7] S. I. N. Reimers and I. Gurevych, Sentence-BERT: Sentence embeddings using Siamese BERT-networks,” in *Proc. EMNLP-IJCNLP*, 2019, pp. 3982–3992.
[8] I. Chalkidis *et al.*, “LexGLUE: A benchmark dataset for legal language understanding in English,” is the name of this collection.This collection is called “LexGLUE: A benchmark dataset for legal language understanding in English.”. in *Proc. ACL*, 2022, pp. 4310–4330.
[9] N. F. Noy and D. L. McGuinness, Ontology development 101: A guide to creating your first ontology,” Stanford Knowledge Systems Laboratory, Stanford University, 2001.
[10] (1) R. Cyganiak, D. Wood M. Lanthaler, RDF 1.1 concepts and abstract syntax, W3C Recommendation, Feb. 2014.
[11] P. Lewis *et al.*, “Retrieval-augmented generation for knowledge-intensive NLP tasks,” in *Proc. Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474.
[12] A. Joshi *et al.*, IL-TUR: Benchmark for Indian legal text understanding and reasoning,” is a programmatic framework.IL-TUR: Benchmark for Indian legal text understanding and reasoning,” is a programmatic framework. in *Proc. ACL*, 2024, pp. 11460–11499.
[13] C. D. Manning, P. Raghavan and H. Schutze, Introduction to Information Retrieval. Cambridge, U.K.: Cambridge University Press, 2008.
[14] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
[15] P. Hitzler, M. Krotzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, The OWL: Web Ontology Language primer, W3C Recommendation, Dec. 2012.
[16] Z. Fei *et al.*, “LawBench: Benchmarking legal knowledge of large language models,” in *Proc. EMNLP*, 2024, pp. 7933–7962.
[17] Government of India, The Water (Prevention and Control of Pollution) Act, 1974, and Act No. 6 of 1974, India.
[18] Government of India, The Air (Prevention and Control of Pollution) Act, 1981, and Act No. 14 of 1981, India.
[19] Government of India, *The Forest (Conservation) Act, 1980*, Act No. 69 of 1980, India.
[20] Government of India, The Wild Life (Protection) Act, 1972, Act No. 53 of 1972, India.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)