

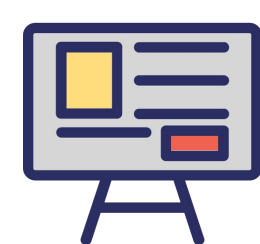
# Exploring Large Language Models for Analyzing Open Source License Conflicts: How Far Are We?

## 探索大语言模型在开源许可证冲突分析中的应用：我们还有多远？

崔星 吴敬征 凌祥 罗天悦 杨牧天 欧文祥

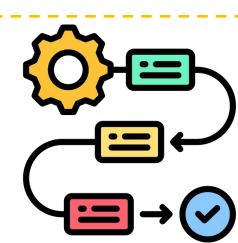
2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE-Companion 2025)

联系人：崔星，13051316652，cuixing@iscas.ac.cn



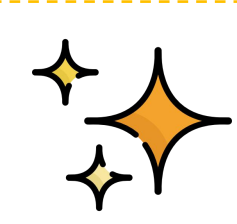
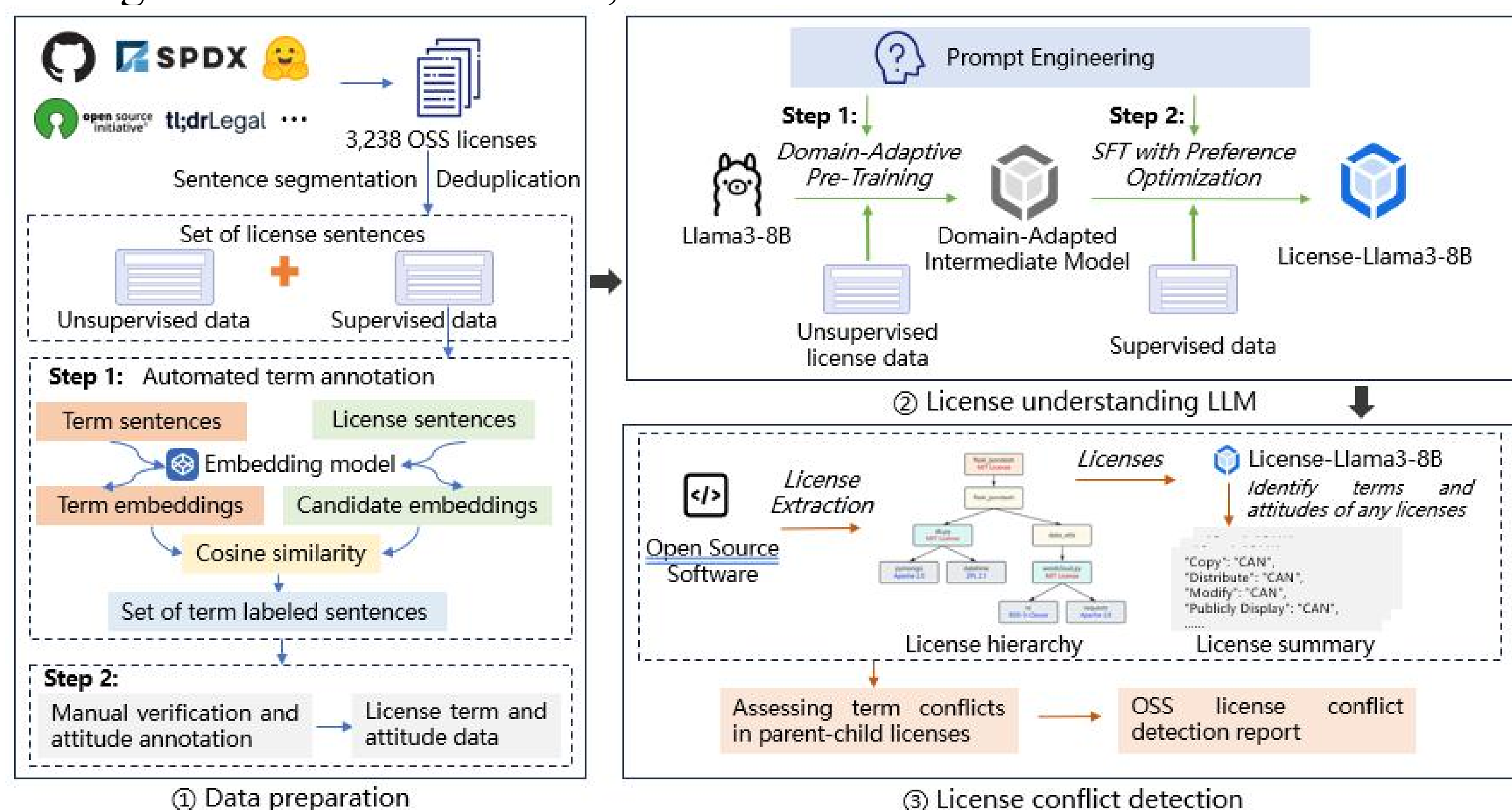
### Introduction

Open source software (OSS) has become the dominant model in modern development, but over 53% of projects face license conflicts, posing risks to sustainability and legal compliance. Existing detection methods often lack coverage and accuracy. We present **L<sup>3</sup>icNexus**, an LLM-based system for automated license conflict detection. It builds a dataset of 3,238 OSS licenses using a joint labeling approach and trains License-Llama3-8B through domain-adaptive pretraining and supervised fine-tuning. The model enables end-to-end license understanding and conflict detection, achieving an F1-score of 85.58 percent, outperforming existing approaches by 20.69 percent. An evaluation on 500 GitHub projects confirms its low false positive and false negative rates, demonstrating the effectiveness of LLMs in OSS license analysis.



### Methodology

This work presents **L<sup>3</sup>icNexus**, an end-to-end LLM-based framework for OSS license analysis, integrating license data preparation, domain-adaptive training of License-Llama3-8B, and hierarchical conflict detection.



### Contributions

- We propose **L<sup>3</sup>icNexus**, an LLM-based system that extracts license terms and attitudes and detects conflicts using license hierarchies.
- Experiments show **L<sup>3</sup>icNexus** achieves 88.04% F1-score in term extraction, 85.58% in license understanding, and low false positive (5.88%) and false negative (2.47%) rates, outperforming existing methods.
- We release the License-Llama3-8B model and its corresponding OSS license dataset on Hugging Face to support community research and development.



### Evaluation and Analysis

Method	P (%)	R (%)	F1 (%)
Text Similarity	18.79	48.99	27.16
FOSS-LTE	17.48	67.45	27.76
DIKE	84.98	60.74	70.84
LiDetector	86.47	66.85	75.40
LiResolver	76.37	72.13	74.19
GPT-4o	48.64	78.19	59.97
GPT-3.5-Turbo	23.23	77.18	35.71
Llama3-70B-Instruct	35.08	66.04	45.82
<b>L<sup>3</sup>icNexus</b>	<b>92.63</b>	<b>83.89</b>	<b>88.04</b>

Table 1: Comparison of term identification

Method	P (%)	R (%)	F1 (%)
Text Similarity	15.19	39.60	21.95
FOSS-LTE	15.83	61.07	25.14
DIKE	72.30	51.68	60.27
LiDetector	73.15	58.31	64.89
LiResolver	63.74	64.35	64.04
GPT-4o	45.49	71.81	55.70
GPT-3.5-Turbo	21.92	72.82	33.70
Llama3-70B-Instruct	32.14	61.34	42.18
<b>L<sup>3</sup>icNexus</b>	<b>90.04</b>	<b>81.55</b>	<b>85.58</b>

Table 2: Comparison of license understanding

Method	Term Identification			License Understanding		
	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)
Llama3-8B	7.99	68.79	14.31	7.44	64.09	13.33
L <sup>3</sup> icNexus(w/o SFT)	11.52	68.12	19.71	10.95	62.08	18.62
L <sup>3</sup> icNexus(w/o DAPT)	86.17	76.82	81.23	84.27	75.16	79.45
L <sup>3</sup> icNexus(w/o ORPO)	92.03	80.91	86.11	87.39	76.84	81.78
<b>L<sup>3</sup>icNexus</b>	<b>92.63</b>	<b>83.89</b>	<b>88.04</b>	<b>90.04</b>	<b>81.55</b>	<b>85.58</b>

Table 3: Performance comparison of removing different modules

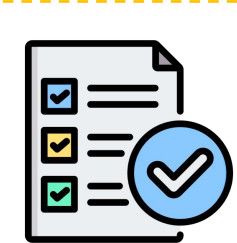
Method	Conflict Issues			Conflict Projects		
	#Report	FPR(%)	FNR(%)	#Report	FPR(%)	FNR(%)
SPDX-VT	13,467	17.58	86.43	96	12.50	57.69
Librariesio	15,703	20.48	63.72	172	15.44	42.31
DIKE	21,992	33.35	47.39	267	25.74	27.20
LiDetector	24,699	26.48	29.56	354	10.29	10.71
LiResolver	29,406	23.73	17.32	361	8.82	5.49
<b>L<sup>3</sup>icNexus</b>	<b>38,847</b>	<b>10.08</b>	<b>8.04</b>	<b>383</b>	<b>5.88</b>	<b>2.47</b>

Table 4: Comparison on license conflict issues and projects



### Findings

- L<sup>3</sup>icNexus** achieves 90.44% accuracy and 85.58% F1-score in extracting license terms and attitudes, and enhances conflict detection through structured summaries of rights and obligations.
- In **L<sup>3</sup>icNexus**, SFT improves understanding, ORPO reduces hallucinations, and DAPT enhances adaptation to OSS license tasks.
- L<sup>3</sup>icNexus** achieves 5.88% false positives and 2.47% false negatives, outperforming five baselines on 500 OSS projects.



### Conclusion

This paper explores the use of LLMs for OSS license conflict detection and presents **L<sup>3</sup>icNexus**, the first LLM-based tool for this task. **L<sup>3</sup>icNexus** builds a dataset of 3,238 licenses and introduces AdaFine to train License-Llama3-8B for accurate term extraction and attitude inference. By constructing a knowledge base and analyzing license hierarchies, it detects conflicts effectively. Evaluated on 500 OSS projects, **L<sup>3</sup>icNexus** outperforms existing methods. Future work will focus on license conflict resolution, recommendation, and customization.



This paper is supported by

Open Source Map