Supplementary Material for JurisGraph Insight Engine 1.0v

Haiguang Zhang MAXWELLZHG@GMAIL.COM
The Chinese University of Hong Kong, Shenzhen, Guangdong, China

1. Criminal Knowledge Graph Details

1.1. Knowledge Graphs

Through careful categorization and linkage, the knowledge graph provides a comprehensive overview of the complex relationships and dependencies within the domains of criminal behavior. The drug-related segment delineates various aspects, including types of illicit substances, distribution networks, and affiliated criminal organizations. It captures detailed information on drug trafficking routes, smuggling techniques, and money laundering strategies, offering valuable insights into the intricate ecosystem of the illicit drug trade.

Similarly, the theft-related portion of the graph encompasses diverse elements such as target demographics, theft modus operandi, and channels for disposing of stolen goods. By visualizing connections among theft hotspots, fencing operations, and underground markets, the graph reveals the underlying mechanisms driving theft-related activities. Together, this knowledge graph serves as an essential tool for understanding the multifaceted nature of drug and theft crimes, supporting informed decision-making and enabling targeted interventions. Furthermore, we compare the entity and relationship data between the drug and theft knowledge graphs, with detailed statistics presented in Tables 1, 2, 3, and 4.

Table 1: Node and Attribute Counts of the Criminal Drug Knowledge Graph (CDKG). * denotes attributes. Time: Case occurrence date/time. Location: Case location.

Name	Court	Time*	Location*	Suspect	Drug	Weight*	Total
Count	2665	2665	2154	4293	2751	1951	16479

Table 2: Edge Types and Counts in the Criminal Drug Knowledge Graph (CDKG).

Edge	Main	Sell Drug To	Traffic In	Provide Shelter For	Possess	Total
Count	2665	1231	1130	1043	589	6658

1.2. Failure Case Analysis

To complement the quantitative results, we further examined several failure cases. One representative example is shown in Table 6, where the system failed to correctly answer a

© 2025 H. Zhang.

Table 3: Node and Attribute Counts of the Criminal Theft Knowledge Graph (CTKG). * denotes attributes. Sus: Suspect; Item: Stolen item; Time: Case occurrence date/time; Loc: Case location; Value: Value of stolen items; Amo: Amount of money.

Name	Court	Sus	Victim	Item (other)	Time	Loc	Value*	Amo	Item (cash)	Sus (other)	Total
Count	6449	6449	3114	5780	2756	3500	2086	481	913	31	31559

Table 4: Edge and Attribute Counts of the Criminal Theft Knowledge Graph (CTKG). * denotes attributes.

Name	Theft	Tool*	Possess	Traffic	Accomplice	Total
Count	3763	732	2704	94	922	8215

query involving multi-jurisdictional drug trafficking. Although the correct answer required synthesizing both the Narcotics Control Law and relevant sections of the Criminal Law, the model generated a response grounded solely in the latter, omitting key clauses on cross-border intent and international cooperation.

This failure highlights two limitations: (1) insufficient cross graph reasoning when multiple legal knowledge subgraphs (CLKG and CDKG) must be traversed; and (2) ambiguity in user intent, where the input did not clearly specify whether domestic or international law should dominate. Such errors motivate future improvements in graph level fusion and intent clarification mechanisms.

Table 7 summarizes the key differences between JurisGraph and existing systems such as ChatLaw and DISC-LawLLM. Our system uniquely integrates several practical features, including interactive subgraph visualization, entity-ID based retrieval, sensitivity filtering, and expert curated legal knowledge graphs. These components collectively support real world deployment scenarios requiring legal compliance and user trust, which sets JurisGraph apart from ChatLaw and DISC-LawLLM, both primarily focused on open domain QA tasks. Specifically, ChatLaw offers partial support for entity ID retrieval through limited modules, whereas DISC-LawLLM lacks this functionality. Additionally, while ChatLaw and DISC-LawLLM build knowledge graphs predominantly via automated extraction methods, JurisGraph relies on expert curated ontologies aligned with criminal law, ensuring higher accuracy and domain relevance. Our system also supports private deployment and multitask capabilities beyond QA, including case search and entity queries. We will include this detailed comparison in the final version to clearly demonstrate our system's unique strengths and its orientation toward practical, trustworthy legal applications.

Table 5: CDKG

Name	Court	Time*	Location*	Suspect	Drug	Weight*	Total
Count	2665	2665	2154	4293	2751	1951	16479

Table 6: Example failure case in legal QA.

User Query: A suspect was caught transporting methamphetamines from Vietnam into China. What is the likely sentence under current law?
Ground Truth: Based on both the Criminal Law and the Narcotics Control Law, cross-border trafficking of methamphetamines can result in a sentence of 15 years to life, depending on quantity and cooperation.
Model Prediction: According to the Criminal Law, drug trafficking is punishable by 5 to 15 years of imprisonment or more, depending on the circumstances.
Error Analysis: The model failed to reference the Narcotics Control Law and to consider the international dimension of the case.

Table 7: System level comparison of recent KG-LLM legal QA models

Feature	**ChatLaw**	**DISC-LawLLM**	**JurisGraph (Ours)**
Core Knowledge Graph and Retrieval Features			
Legal Subgraph Integration	✗	✗	✓
Entity ID Based Retrieval	✗	✗	✓
Hybrid Retrieval + QA	✓	✓	✓
Deployment and Safety			
Local Deployment Option	✗	✗	✓
Sensitive Question Rejection	✗	✗	✓
Explainability and User Support			
Graph-Based Output Explanation	✗	✗	✓
Legal Expert Involvement	✗	✗	✓
Multi-task Support	QA Focused	QA Focused	✓
Private Deployment Support	Limited	Limited	✓

1.3. Training and Evaluation of Large Model

The software and hardware environment used in this system are as follows: the graphics card is NVIDIA RTX3090 with a memory capacity of 24GB, operating system is Linux 5.15.0-91-generic, Python version is 3.8.0, and Torch version is 1.13.1. The large model used in this paper is based on Alibaba Cloud's Qwen-7B. Due to the high time and computational cost of training large language models, the total training epochs are set to 1, with a time consumption of approximately 25 hours. The specific training parameters are as follows: fine-tuning method is QLoRA, learning rate is 0.0002, batch size is 4, maximum sentence length is 1024, LoRA rank is 16, LoRA alpha is 16, LoRA dropout is 0.05, and Warm steps is 700. To demonstrate the effectiveness of our proposed model, we compare it with several strong baseline models for RTE:

- ChatGLM (Du et al., 2021) is a powerful natural language generation model. It is based on the Transformer architecture and learns the syntax, semantics, and contextual information of language through training on large-scale corpora.
- Baichuan-Chat (Baichuan, 2023) is trained on 26 trillion tokens of high-quality corpora, achieving the best performance for its size on authoritative Chinese and English benchmarks.
- Chinese-Alpaca-2 (Cui et al., 2023b) utilizes large-scale Chinese data for incremental pre-training, further enhancing its understanding of Chinese basic semantics and instructions, leading to significant performance improvements compared to similar models.
- GPT-3.5-turbo (OpenAI, 2022) is a member of OpenAI's natural language processing model series, representing an improved version of GPT-3, likely featuring performance optimizations and accuracy enhancements.
- LexiLaw (University, 2023b) is a fine-tuned large-scale Chinese legal model based on the ChatGLM-6B architecture, which has been fine-tuned on legal datasets to enhance its performance and expertise in providing legal consultation and support.
- LawGPT (University, 2023a) boosts its understanding of legal semantics by expanding general Chinese base models like Chinese-LLaMA and ChatGLM, and improves its ability to comprehend and execute legal content through fine-tuning on specific legal datasets.
- Lawyer LLaMa (Huang et al., 2023) is continuously pretrained on a large legal corpus, allowing it to systematically grasp China's legal knowledge system. We then fine-tuned the model using ChatGPT, analyzing responses from China's National Unified Legal Professional Qualification Examination and legal consultations, enabling it to apply legal knowledge effectively in various scenarios.
- ChatLaw (Cui et al., 2023a) employs fine-tuning based on Ziya-LLaMA and lora techniques, achieving notable success in legal inquiries.
- DISC-LawLLM (Yue et al., 2023) introduces a large language model-driven intelligent legal system in Chinese, offering diverse legal services to various user groups.
- InternLM-Law(Fei et al., 2025) is a Chinese legal large language model designed to excel across diverse legal NLP tasks, offering state-of-the-art performance and open-source resources to advance research in this domain.
- Deepseek(DeepSeek-AI et al., 2025) is a high performance open-source LLM excelling in reasoning, knowledge-based QA.

References

Baichuan. Baichuan 2: Open large-scale language models. *arXiv preprint arXiv:2309.10305*, 2023. URL https://arxiv.org/abs/2309.10305.

Jiaxi Cui, Zongjian Li, Yang Yan, Bohua Chen, and Li Yuan. Chatlaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*, 2023a.

Yiming Cui, Ziqing Yang, and Xin Yao. Efficient and effective text encoding for chinese llama and alpaca. *arXiv preprint arXiv:2304.08177*, 2023b.

DeepSeek-AI, Daya Guo, Dejian Yang, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning, 2025.

Zhengxiao Du, Yujie Qian, Xiao Liu, Ming Ding, Jiezhong Qiu, Zhilin Yang, and Jie Tang. Glm: General language model pretraining with autoregressive blank infilling. *arXiv preprint arXiv:2103.10360*, 2021.

Zhiwei Fei, Songyang Zhang, Xiaoyu Shen, et al. InternLM-law: An open-sourced Chinese legal large language model. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 9376–9392, January 2025.

Quzhe Huang, Mingxu Tao, Zhenwei An, Chen Zhang, Cong Jiang, Zhibin Chen, Zirui Wu, and Yansong Feng. Lawyer llama technical report. *arXiv preprint arXiv:2305.15062*, 2023.

OpenAI. Openai: Introducing chatgpt. 2022. URL https://openai.com/blog/chatgpt.

Nanjing University. Lawgpt. 2023a. URL https://github.com/pengxiao-song/LaWGPT.

Tsinghua University. Lexilaw. 2023b. URL https://github.com/CSHaitao/LexiLaw.

Shengbin Yue, Wei Chen, Siyuan Wang, Bingxuan Li, Chenchen Shen, Shujun Liu, Yuxuan Zhou, Yao Xiao, Song Yun, Wei Lin, et al. Disc-lawllm: Fine-tuning large language models for intelligent legal services. *arXiv preprint arXiv:2309.11325*, 2023.