LexRel: Benchmarking Legal Relation Extraction for Chinese Civil Cases

Anonymous ACL submission

Abstract

Legal relations form the core analytical framework of the civil law system, serving as a crucial foundation for resolving disputes and realizing the values of the rule of law in judicial practice. However, legal relations in Chinese civil cases remain underexplored in the field of legal artificial intelligence (legal AI), largely due to the absence of structured, learnable schemas. In this work, we introduce the first systematic and comprehensive schema for modeling legal relations in Chinese civil cases. Grounded in established Chinese civil law theory, we construct a hierarchical taxonomy of legal relations along with their subjects, objects and content. This taxonomy provides a comprehensive coverage of legal relations in Chinese civil cases, covering 9 broad domains of Chinese civil law and encompassing 265 fine-grained categories of civil legal relations. Based on this taxonomy, we construct **LexRel**, an expert-annotated dataset for legal relation extraction in Chinese civil law, curated from real-world legal cases. We use **LexRel** to evaluate state-of-the-art large language models (LLMs) on the task of extracting legal relations from case facts. Empirical results show that current LLMs exhibit significant limitations in accurately identifying civil legal relations, revealing substantial scope for improvement. Furthermore, we demonstrate that incorporating structured representations of legal relations leads to consistent performance gains on both legal relation extraction and other downstream legal AI tasks.[1]

1 Introduction

Legal relations constitute the foundational analytical unit of the Chinese civil law system. Defined as the relationships between individuals as regulated by legal norms, legal relations serve as the crucial interface between the normative world of law and the empirical world of social facts (Yuan, 2024). In judicial practice, legal professionals frequently rely on legal relations as a key reference point for retrieving relevant legal norms, cases, and doctrinal interpretations (Shen, 2004). Moreover, legal relations constitute essential information for a variety of downstream legal tasks, such as statute prediction and case outcome analysis. In this way, legal relations function not only as the conceptual foundation for modeling legal analysis, but also as a practical interface for organizing and accessing legal knowledge in decision-making contexts.

Despite their central importance to both legal theory and judicial practice, legal relations remain largely underexplored in legal AI research. Ex-

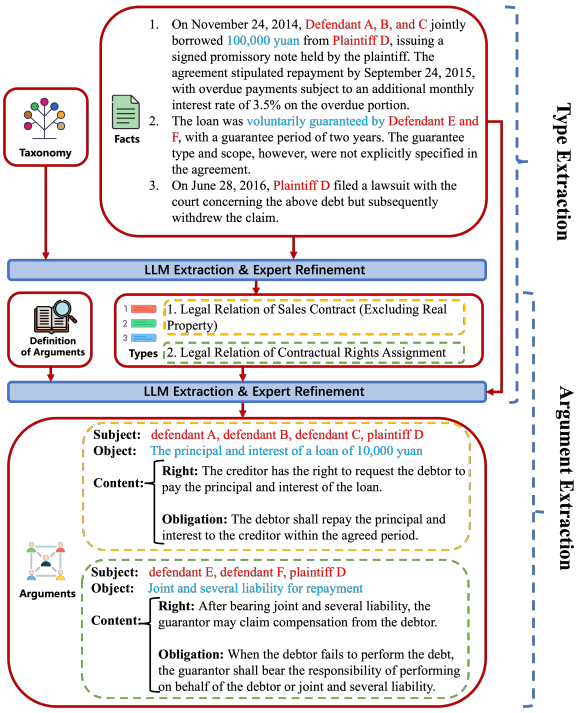

Figure 1: The construction workflow of **LexRel**, consisting of two modules: **type extraction** and **argument extraction**. The English translation of the original Chinese text is presented.

[1] Code, taxonomy, prompts and data will be released upon acceptance.

isting studies on relation modeling in this field have primarily focused on legal information extraction, targeting factual entities (e.g., persons, goods, contracts) or general social relationships (e.g., employment, ownership) (Cui et al., 2024; Huang et al., 2021; Li et al., 2024b, 2021; Xu, 2021). Such approaches overlook that legal relations—a concept grounded in statutory rules and judicial practice—are qualitatively distinct from ordinary semantic associations in natural language. A key reason for this gap is the absence of a well-structured schema of legal relations. In real-world practice, legal analysis and reasoning often proceed through identifying and analyzing the legal relations between parties. However, these relations are rarely made explicit in written judgments or legal documents, where factual analysis and legal reasoning are typically intertwined. As a result, models trained solely on large-scale legal texts struggle to capture the underlying relational structure of legal analysis. Developing a structured schema of legal relations is therefore essential, enabling AI systems to explicitly model and reason over this foundational layer of legal knowledge.

In this paper, we present the first comprehensive schema to model legal relations in Chinese civil cases. Drawing from established legal theory and judicial practice, our schema consists of a hierarchical taxonomy spanning 9 major domains of civil law and 265 relation types, together with precise definitions of their core arguments (subject, object and content).

Building on this schema, we develop **LexRel**, a benchmark of expert-annotated legal relations from real-world cases. To ensure legal accuracy and conceptual rigor, we collaborate with domain experts in Chinese civil law during the development of annotation guidelines, the annotation process itself, and subsequent expert review.

We use **LexRel** to evaluate the capabilities of state-of-the-art large language models (LLMs) on the task of legal relation extraction. In addition, we apply our relation-aware method to other legal AI tasks, where the incorporation of structured legal relation knowledge allows us to assess its contribution to downstream performance.

Our contributions are summarized as follows:

- We propose the first structured schema of legal relations in Chinese civil law, with a taxonomy covering 265 relation types across 9 major domains, and clear definitions of subjects, objects, and content.
- We construct **LexRel**, the first expert-annotated dataset for legal relation extraction in Chinese civil cases.
- We evaluate existing LLMs on **LexRel**, revealing significant model limitations in extracting legal relations from case facts.
- We demonstrate that integrating structured legal relations improves performance on downstream legal AI tasks.

2 Related Work

2.1 Legal Relation

The concept of *legal relation* refers to a social relation regulated by law, defined by the rights and obligations that arise between parties by operation of law (Liang, 2020). As used in this article, the term primarily denotes civil legal relations under the Civil Code of the People's Republic of China (Wan et al., 2022). Within the new Civil Code, legal relations structure the organization of all chapters and form the basis for the logical framework of civil law (Wang, 2018). In judicial practice, legal relations also serve as a core methodological approach: judges situate the parties within a specific relation, analyze the subjects, objects, and their respective rights and obligations, and determine the creation, alteration, or extinction of rights (Wang, 2017). While prior scholarship and doctrine have outlined general definitions of legal relations and their arguments (subject, object and content), a systematic taxonomy and precise operationalization of these elements has not previously been established.

2.2 Legal Knowledge Graph

Legal knowledge graphs (KGs) are widely used in legal AI, supporting applications such as legal reasoning (Luo et al., 2024; Li et al., 2021), question answering (Liang et al., 2024), and decision support (Cui et al., 2024). These graphs typically organize entities (e.g., persons, organizations, statutes) and their relationships into structured formats to enable downstream machine reasoning. However, most existing legal KGs in Chinese civil law focus on general social relations (e.g., employment, kinship, transaction) (Cui et al., 2024; Li et al., 2021) or entity co-occurrence patterns (Li et al., 2024b). That is, while legal KGs often capture who interacts with whom, they rarely characterize how

they are related under law, such as the presence of a creditor-debtor relationship or a contractual obligation. Some works (Xu, 2021) adopt stricter definitions of civil legal relations to construct relation types, but their coverage remains limited and lacks a comprehensive taxonomy.

2.3 Legal Large Language Models

Large language models (LLMs) show remarkable promise in legal applications, including legal consultation, trial assistance, and document analysis (Lai et al., 2023). Several works explore adapting general LLMs to the legal domain through continued pretraining and fine-tuning (Huang et al., 2023; Yue et al., 2023a; Xiao et al., 2021; Yue et al., 2023b). In addition, some works focus on evaluating the capabilities of LLMs in the legal field (Dai et al., 2023; Li et al., 2024a). These advancements demonstrate the growing capability of legal LLMs and highlight their strong potential to support a wide range of legal applications.

3 Schema for Civil Legal Relations

We develop a structured schema for legal relations in Chinese Civil Law, which comprises a hierarchical taxonomy of relation types and precise definitions of their arguments.

3.1 Taxonomy for Civil Legal Relations

We construct the taxonomy of civil legal relations through a combination of draft taxonomy construction and expert-guided refinement:

- **Draft taxonomy construction.** We begin by applying a keyword-matching approach to extract relational expressions from Chinese civil court judgments (China Judgments Online, 2013), producing an initial list of candidate legal relation terms. These candidates are then refined by cross-referencing specialized legal terminology references (Editorial Committee of Zhong hua fa xue da ci dian, 1996), retaining only those whose prefixes correspond to legally recognized concepts. This process yields 123 candidate relation types, which are subsequently clustered and organized into an initial taxonomy grounded in the rights–obligations framework of Chinese civil law (Wan et al., 2022). The resulting draft taxonomy comprises six major domains: *Legal Relations of Personality Rights*, *Legal Relations of Identity Rights*, *Legal Relations of Intellectual Property Rights*, *Legal Relations of Succession Rights*, *Legal Relations of Property Rights*, and *Legal Relations of Obligations*.

- **Expert-guided refinement.** Two senior legal scholars (professors or lecturers in law) then review and refine the draft taxonomy. Drawing on legal textbooks, academic literature, and their professional expertise, they add three new civil domains—*Bill Relations*, *Letter of Credit Relations*, and *Independent Guarantee Relations*—and expand the coverage to 265 fine-grained legal relation types.

These two stages, combining empirical induction from litigated cases with normative grounding in established legal doctrine, ensures that the resulting taxonomy is both data-informed and conceptually rigorous. The finalized taxonomy covers 9 broad domains in Chinese civil law, consisting of 265 fine-grained legal relation types.

Among these domains, *Legal Relations of Obligations* encompasses the largest number of relation types with 200. This predominance reflects the inclusion of two categories that correspond to the most common types of civil disputes in practice—*Contractual Legal Relations* and *Legal Relations of Tort Liability*—which account for the majority of cases in judicial proceedings (Supreme People's Court of the People's Republic of China, 2024).

3.2 Arguments of Civil Legal Relations

In Chinese civil law, legal relation's arguments are defined as a triplet: subject, object, and content (Wang, 2018). However, there are no precise definitions for different arguments in different types of legal relation. With the help of the specific provisions of the Civil Code of the People's Republic of China regarding different categories of rights (Wan et al., 2022) and senior legal scholars, we add precise definitions of the legal relations contained in main domains of our taxonomy.

To ensure semantic clarity and legal fidelity, we standardize 12 distinct types of definitions to subject, object and content, corresponding to 8 out of 9 domains as well as 4 major subtypes under Legal Relations of Obligations: Contractual Legal Relations, Legal Relations of Tort Liability, Legal Relations of Negotiorum Gestio and Legal Relations of Unjust Enrichment. Each relation type is

Definitions of Arguments in Legal Relations of Property Rights	
Subject	The right holder himself/herself, and all subjects related to the rights and obligations relationship are obligors (generally, the right holder and other obligors who bear direct legal obligations in this fact are output).
Object	Corporeal things, including real property, movable property and special items.
Content	In terms of rights, the right-holder legally enjoys ownership, usufructuary right, security interest, etc. The right-holder exercises direct control over the object through powers such as possess, use, seek profits from and dispose of the real property or movable property according to law. Real property rights become effective upon registration, whereas rights in movable property become effective upon delivery. In terms of obligations, it includes completing the registration or delivery procedures in accordance with the law; not infringing upon the legitimate property rights of others; and assuming the protection obligations such as removing obstacles, restoring the original state, and returning the original object when the property rights are unlawfully interfered with.

Table 1: An example of the definitions of the subject, object, and content for *Legal Relations of Property Rights*.

thus linked to a specific combination of different roles and legal meaning.

- Subject refers to the party initiating or holding the right or obligation, typically a natural person, legal entity, or organization.
- Object denotes the counterparty or the legal target of the relation, such as another individual, a piece of property, or an intellectual asset.
- Content captures the substantive nature of the legal relation—e.g., ownership, liability, contractual duty, or tortious responsibility.

Table 1 shows an example of the definition of legal relations. The 12 definitions of subject, object, and content comprehensively cover all legal relation types defined in our taxonomy. Each relation type is associated with a specific configuration of these triplet elements, forming a unified schema that serves as the foundation for dataset construction and model evaluation in this work. All definitions are shown in Appendix C

4 LexRel

As shown in Fig. 1, with the structured legal relation schema in place, we construct the **LexRel** benchmark. Each entry in **LexRel** consists of three components: the factual description of the case, a list of legal relation types, and the corresponding arguments (subject, object, and content) for each legal relation. Based on this structured data, we define the legal relation extraction task as comprising two subtasks: *type extraction*, which aims to identify the relevant legal relation types, and *argument extraction*, which seeks to determine the specific knowledge in each relation. **LexRel** is carefully curated from real-world Chinese civil court decisions through a two-step process: first, use LLMs to annotate candidate relations and their corresponding arguments, and second, legal experts review and refine the annotations to ensure both coverage and correctness.

4.1 Task Definition

We formulate *legal relation extraction* as a two-stage structured prediction task: *type extraction* and *argument extraction*.

4.1.1 Step 1: Type Extraction

Given input case facts x, which consists of the fact statements in a court judgment, the type extraction stage aims to predict the legal relation type $\hat{r}$:

$$\hat{r} = f_{\text{type}}(x), \quad \hat{r} \in \mathcal{R} \tag{1}$$

Here, f_{type} denotes the relation type classifier, and $\mathcal{R}$ is the legal relation set predefined in our taxonomy.

4.1.2 Step 2: Argument Extraction

Conditioned on the predicted relation type $\hat{r}$, the model then extracts the set of associated subjects $\hat{S} = \{\hat{s}_1, \hat{s}_2, \ldots\}$, the set of associated objects $\hat{O} = \{\hat{o}_1, \hat{o}_2, \ldots\}$, and the detailed content (i.e., rights and obligations of the parties) $\hat{c}$. The extractor f_{arg} produces argument sets based on both input facts x and relation type $\hat{r}$:

$$(\hat{S}, \hat{O}, \hat{c}) = f_{\text{arg}}(x, \hat{r}) \tag{2}$$

4.2 Data Construction

In order to ensure the quality of extraction and cost effectiveness, we adopt the method of extracting first with LLM and then correcting by experts.

Type Extraction	Argument Extraction
Please extract the existing legal relations from the following judgment text. Only output the names, no need for explanation or description. **Judgment text: {}** **Legal relation candidate set: {}** Output requirements: 1. Based on the content of the judgment, find the names of the legal relationships existing in the judgment text. 2. Only output the name of the most important legal relationship existing in the judgment text, and the output relationship name should be within 10 characters. 3. Each legal relationship name output must be in the legal relationship set. 4. Output strictly in accordance with the output format, and the output format is: [Legal relation 1, Legal relation 2.....]	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: {}**. **Object definition: {}**. **Content definition: {}**. **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.

Table 2: Prompt templates for *type extraction* and *argument extraction*. English translation for the original Chinese text is presented.

Data preprocessing. We draw on court judgments for real-world adjudicated civil trial cases to construct the **LexRel** dataset. Firstly, We utilize DeepSeek-V3 (DeepSeek-AI, 2025b) to extract legal relation types and their arguments from the full-text judgment as draft annotations to guide human experts during their annotations. The extraction follows the two-step task structure defined above, with prompting instructions provided in Table 2.

And then, we also extract the facts authenticated by the court with DeepSeek-V3 as input for our dataset from each judgment as the final input for our dataset. The LLM is prompted to identify factual paragraphs that reflect the adjudicated evidence, excluding legal analysis and procedural history.

Expert annotation. To ensure the accuracy of the annotations, we engage six legal experts to review and refine them. All annotators hold law degrees and have prior experience with data annotation. The annotation process is overseen by a senior expert in legal AI research with extensive knowledge of legal relations. Annotators are provided with a detailed guideline including our schema and are briefed to ensure consistent understanding. Each annotator is assigned 200 items, each contains its facts and court reasoning extracted from judgments, together with draft relation types and arguments. They are tasked with verifying and refining the input case facts and these labels. After annotation, 60 items without identifiable legal relations are removed. The final **LexRel** benchmark comprises 1,140 annotated samples, each consisting of case facts paired with legal relations and their corresponding arguments, forming the **LexRel** evaluation dataset for assessing models' ability to extract legal relations from case facts.

To validate annotation quality, we randomly sample 100 cases from **LexRel** and ask the senior expert to independently re-annotate them. Comparing the senior expert's annotations with those produced by the annotation team yields high inter-annotator agreement (IAA), with a Cohen's Kappa (Cohen, 1960) of 0.920 for *type extraction* and 0.774 for *argument extraction*. These results indicate strong consistency and confirm the overall reliability of the annotated dataset.

5 Experiment

5.1 Evaluation on LLMs

We firstly evaluate closed-source and open-source state-of-art LLMs on the **LexRel** benchmark under two settings: **zero-shot prompting** and **distillation-based supervised fine-tuning (SFT)**. These settings allow us to assess both the off-the-shelf legal relation extraction capability of recent models and the feasibility of transferring this capability to smaller open-source models.

5.1.1 Settings

Models. We select widely-used state-of-the-art LLMs for evaluation, including closed-source models: GPT-4o (Hurst et al., 2024), o3-mini (OpenAI, 2025) and Claude-Sonnet-4 (Anthropic, 2025), as well as open-source models: DeepSeek-V3 (DeepSeek-AI, 2025b), DeepSeek-R1 (DeepSeek-AI, 2025a), Llama3.1-8B-Instruct, Llama3.1-70B-Instruct (Dubey et al., 2024), InternLM3-8B-Instruct (Cai et al., 2024), MiniCPM4-8B

Models	Methods	Type Extraction				Argument Extraction			
		P	R	micro-F1	macro-F1	P	R	micro-F1	macro-F1
GPT-4o	Zero-Shot	0.582	**0.789**	0.670	0.314	0.188	0.278	0.224	0.068
o3-mini	Zero-Shot	**0.768**	0.756	**0.762**	0.441	**0.413**	0.355	**0.382**	0.129
Claude-Sonnet-4	Zero-Shot	0.619	0.564	0.590	0.330	0.206	0.345	0.258	0.088
DeepSeek-R1	Zero-Shot	0.647	0.746	0.693	0.376	0.299	0.242	0.268	0.065
DeepSeek-V3	Zero-Shot	0.440	0.781	0.563	0.308	0.164	0.430	0.237	0.083
Llama3.1-70B	Zero-Shot	0.192	0.384	0.256	0.074	0.022	0.203	0.039	0.008
Qwen3-32B	Zero-Shot	0.500	0.700	0.583	0.158	0.030	**0.551**	0.057	0.015
Llama3.1-8B-Instruct	Zero-Shot	0.170	0.472	0.250	0.052	0.014	0.337	0.027	0.006
	SFT *w/ GPT-4o*	0.488	0.741	0.588	0.270	0.131	0.287	0.180	0.054
	SFT *w/ DeepSeek-R1*	0.586	0.771	0.666	0.338	0.236	0.339	0.279	0.093
Internlm3-8B-Instruct	Zero-Shot	0.310	0.644	0.418	0.139	0.055	0.042	0.048	0.002
	SFT *w/ GPT-4o*	0.479	0.773	0.592	0.288	0.150	0.349	0.209	0.061
	SFT *w/ DeepSeek-R1*	0.599	**0.817**	0.691	0.324	0.277	0.387	0.323	0.100
MiniCPM4-8B	Zero-Shot	0.353	0.417	0.382	0.218	0.095	0.128	0.109	0.053
	SFT *w/ GPT-4o*	0.519	0.760	0.617	0.314	0.162	0.314	0.213	0.054
	SFT *w/ DeepSeek-R1*	0.599	0.769	0.674	0.343	0.243	0.310	0.272	0.086
Qwen3-8B	Zero-Shot	0.338	0.745	0.464	0.149	0.061	0.275	0.099	0.014
	SFT *w/ GPT-4o*	0.536	0.769	0.632	0.311	0.175	0.346	0.233	0.078
	SFT *w/ DeepSeek-R1*	0.576	0.816	0.675	0.337	0.255	0.375	0.304	0.098
Qwen3-14B	Zero-Shot	0.564	0.603	0.583	0.313	0.072	0.173	0.102	0.035
	SFT *w/ GPT-4o*	0.614	0.749	0.675	0.357	0.212	0.351	0.264	0.096
	SFT *w/ DeepSeek-R1*	**0.719**	0.748	**0.733**	0.430	**0.373**	**0.389**	**0.381**	0.146

Table 3: Evaluation results of state-of-art LLMs on **LexRel**.

(MiniCPM-Team, 2025), Qwen3-8B, Qwen3-14B and Qwen3-32B (Yang et al., 2025).

Prompts. We follow the task formulation introduced above to evaluate on two subtasks: *type extraction* and *argument extraction*. All LLMs use a consistent prompting strategy that incorporates the predefined schema. Detailed prompt templates are provided in Table 2.

Data generation for SFT. We generate pseudo-labels for 5,500 instances collected from the same corpus to construct SFT. We employ our two good-performing zero-shot models, **GPT-4o** and **DeepSeek-R1**, as teacher models. Following the guidance designed for human annotators, we also divide the tasks for LLM annotators into: type annotation and arguments annotation. We first instruct the model to output the types, then have it generate the corresponding arguments based on them. The prompts used are identical to those in Table 2.

Hyper-parameter. In the zero-shot setting, we use a temperature of 0.7 and a maximum generation length of 8192 tokens. In the SFT setting, we firstly generate data using GPT-4o and DeepSeek-R1 for 5,500 items with the full-text judgments, and fine-tune open-source models with 8B and 14B parameters. Fine-tuning is conducted using LoRA (Hu et al., 2022) on 8×A800 GPUs (each with 40GB memory), with a learning rate of 1e-4, a batch size of 4 and train for 3 epochs. We further randomly sample 550 items as the validation set.

Metrics and LLM-as-a-Judge. We use precision, recall, micro-F1 score and macro-F1 score (computing by averaging F1 scores across different relation types) to assess model performance on both *type extraction* and *argument extraction* tasks. For *type extraction*, evaluation is conducted by directly matching each predicted legal relation type against the gold-standard labels. Considering the inherent ambiguity in legal language and the cost of large-scale human evaluation, we adopt an LLM-as-a-Judge strategy for evaluating argument correctness. Specifically, we employ **DeepSeek-V3** to compare each predicted argument against the gold standard.

To validate this choice, we conduct a manual quality check on 100 randomly selected samples. A human expert independently evaluates each case, and we compare the expert's judgments with those produced by DeepSeek-V3. A prediction is considered correct if (1) the subject and object refer to the correct entities or persons, and (2) the content is semantically equivalent. We observe a high degree of agreement, with a Cohen's Kappa (Cohen, 1960) of 0.737, confirming the reliability of the automated LLM-assisted evaluation.

5.1.2 Results and Discussion

Results of our evaluation are shown in Table 3. In zero-shot type extraction task, the best overall performer is o3-mini ($micro{-}F1 = 0.762$), followed by DeepSeek-R1 ($micro - F1 = 0.693$). Among open-source models, the Qwen3 series demonstrates the strongest performance, with Qwen3-8B, Qwen3-14B and Qwen3-32B achieving micro-F1 scores of 0.464 ,0.583 and 0.583 respectively. These results suggest that reasoning LLMs are better suited for zero-shot legal relation type extraction. In contrast, argument extraction proves significantly more challenging. The highest zero-shot micro-F1 is again achieved by o3-mini (0.382), followed by DeepSeek-R1 (0.268) and Claude-Sonnet-4 (0.258). Most open-source models perform poorly ($micro - F1 < 0.2$), indicating the difficulty of identifying multiple argument spans in unstructured legal text under zero-shot conditions.

In the SFT experiments, we distill responses from GPT-4o and DeepSeek-R1 into smaller open-source models, and their performance improves significantly across both tasks [2]. For type extraction, the best distillation result is achieved by Qwen3-14B ($micro{-}F1 = 0.733$), which closely matches the performance of o3-mini. And argument extraction also benefits much from SFT, for example, InternLM3-8B improves from 0.048 to 0.323 in micro-F1. This large gain suggests that SFT can effectively equip smaller models with the ability to learn structured legal relations.

Furthermore, we observe that the macro-F1 scores of all LLMs are lower than their corresponding micro-F1 scores across all settings. This disparity suggests that the LLMs exhibit varying capabilities across different types of legal relations.

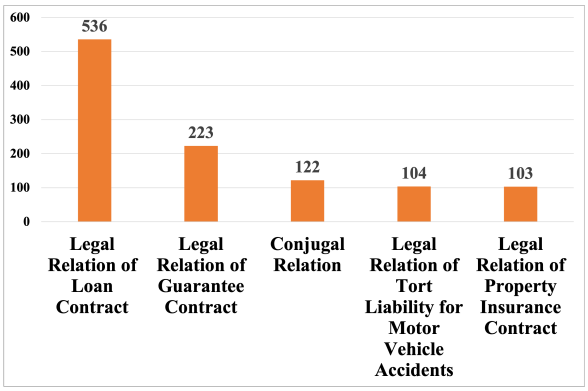

Figure 2: Top 5 legal relation types in **LexRel**.

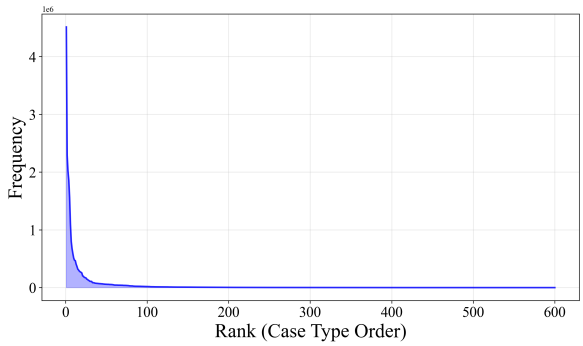

Figure 3: The frequency distribution of causes of action across 26.6 million civil judgments shows a pronounced long-tail pattern.

5.2 Long-Tail Analysis

LexRel contains 1,140 civil cases and 1,863 fully annotated legal relations spanning all major civil law domains. Fig 2 shows the five most frequent legal relation types in **LexRel**, revealing a clear downward trend. We further examine whether this downward pattern faithfully mirrors real-world distributions. Because real-world cases do not explicitly specify legal relations, we use **civil causes of action** as a proxy. We first validate the reliability of this proxy by demonstrating a strong correlation between legal relations and causes of action (χ^2 test, $p < 0.001$; Cramér's $V = 0.528$). We then examine a comprehensive set of 26.6 million real-world civil court judgments from China Judgments Online (China Judgments Online, 2013) and extract the cause of action information for each case. Fig 3 shows that causes of action for real-world civil cases follow a long-tail distribution. Consistently, Pareto analysis of **LexRel** (Fig 4) shows that a small number of frequent relation types account for the majority of instances, while a large number

[2] In this paper, significance is tested through t-test (α=0.05).

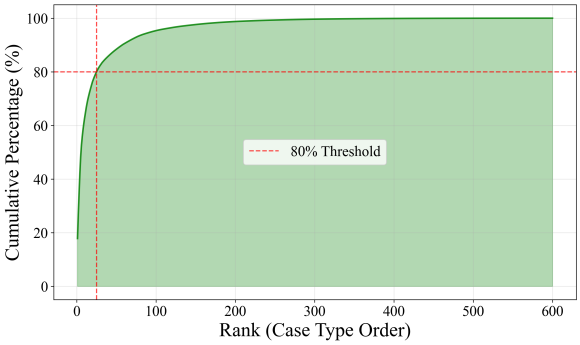

Figure 4: The Pareto distribution of causes of action reveals that the top 25 case types account for 80% of all civil judgments, highlighting a pronounced long-tail pattern.

Models	Document Proofreading	Statute Prediction	Consultation
GPT-4o	22.12	33.16	18.16
w/ LR	**22.62**	32.01	**19.17**
DeepSeek-V3	23.81	34.86	16.39
w/ LR	**25.60**	**35.10**	**16.45**
MiniCPM4-8B-Instruct	5.70	22.10	7.61
w/ LR	**11.78**	**24.36**	**8.41**
Internlm3-8B-Instruct	13.56	9.53	8.97
w/ LR	**16.58**	**17.85**	**11.71**

Table 4: Results on downstream legal AI tasks.

of relations occur sparsely. Taken together, these results indicate that the long-tail distribution of legal relations in **LexRel** echoes the intrinsic pattern observed in real-world litigation data.

5.3 Effects on Downstream Legal AI Tasks

To demonstrate the necessity of incorporating legal relations in legal AI research, we select a set of representative tasks as our evaluation benchmarks. For each task, we extract the legal relation types and arguments mentioned in instructions, and provide them explicitly to the models during inference.

5.3.1 Settings

Tasks. We base our task selection on the categorization provided by LawBench (Fei et al., 2023), which organizes legal AI tasks across various legal domains. Specifically, we focus on tasks situated in the civil law domain that require a high degree of legal knowledge and structured information, including the tasks of Document Proofreading, Statute Prediction and Consultation.

Models. We select four representative models as GPT-4o, DeepSeek-V3, MiniCPM4-8B-Instruct, and Internlm3-8B-Instruct. The model configurations are kept the same to ensure a fair comparison between models.

Hyper-parameter. Considering both efficiency and effectiveness, we use legal relation distilled from DeepSeek-V3. Specifically, we extract relation types and arguments of the input and insert them directly into the task instructions during inference. All other evaluation parameters remain consistent with the previous section.

Metrics. We follow the official evaluation guidelines of LawBench, where each task is scored on a scale from 0 to 100 based on correctness and completeness of the model's output. We compare results under two settings: (1) vanilla zero-shot, where the model receives only the task instruction and input; and (2) relation-aware zero-shot, where the input is enhanced with distilled legal relation types and arguments.

5.3.2 Results and Analysis

As shown in Table 4, incorporating legal relations ("w/ LR") consistently enhances performance across nearly all tasks and models. In the Document Proofreading task, contextualized legal relations enable models to more effectively detect redundant or inaccurate statements by anchoring the analysis in the underlying legal structure. For Statute Prediction and Consultation, these relations provide essential cues for identifying and reasoning over relevant statutory knowledge. Together, these findings indicate that explicitly modeling legal relations strengthens the model's ability to perform knowledge-intensive legal tasks with greater accuracy and interpretability.

6 Conclusion

In this work, we explicitly model legal relations in the context of Chinese civil law. By constructing a structured schema of civil legal relations and curating an evaluation dataset for the legal relation extraction task, we evaluate large language models on their ability to effectively extract legal relations and associated arguments, and further demonstrate the benefits of incorporating such structured knowledge across multiple tasks.

7 Limitations

Our schema and dataset focus specifically on civil legal relations under Chinese law, which may limit direct transferability to other jurisdictions. Although the conceptual structure of legal relations is broadly consistent across civil law systems, each jurisdiction embodies distinct doctrines and terminology, requiring localized adaptation.

8 Ethical Considerations

All data used in this study are sourced from publicly available Chinese civil court decisions released through official judicial channels (China Judgments Online, 2013). We strictly adhere to data privacy and research ethics standards, ensuring that no sensitive or confidential information is disclosed. All human annotators participated voluntarily and were compensated fairly according to institutional guidelines.

References

Anthropic. 2025. Introducing Claude 4. https://www.anthropic.com/news/claude-4.

Zheng Cai, Maosong Cao, Haojiong Chen, Kai Chen, Keyu Chen, Xin Chen, Xun Chen, Zehui Chen, Zhi Chen, Pei Chu, and 1 others. 2024. InternLM2 technical report. *arXiv preprint arXiv:2403.17297*.

China Judgments Online. 2013. China Judgments Online. https://wenshu.court.gov.cn.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

Jiaxi Cui, Munan Ning, Zongjian Li, Bohua Chen, Yang Yan, Hao Li, Bin Ling, Yonghong Tian, and Li Yuan. 2024. ChatLaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model. *Preprint*, arXiv:2306.16092.

Yongfu Dai, Duanyu Feng, Jimin Huang, Haochen Jia, Qianqian Xie, Yifang Zhang, Weiguang Han, Wei Tian, and Hao Wang. 2023. LAiW: A Chinese legal large language models benchmark. *arXiv preprint arXiv:2310.05620*.

DeepSeek-AI. 2025a. DeepSeek-R1: Incentivizing reasoning capability in llms via reinforcement learning. *Preprint*, arXiv:2501.12948.

DeepSeek-AI. 2025b. DeepSeek-V3 technical report. *Preprint*, arXiv:2412.19437.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, and 1 others. 2024. The Llama 3 herd of models. *arXiv e-prints*, pages arXiv–2407.

Editorial Committee of Zhong hua fa xue da ci dian. 1996. *Zhong hua fa xue da ci dian*. Zhong guo jian cha chu ban she, Beijing.

Zhiwei Fei, Xiaoyu Shen, Dawei Zhu, Fengzhe Zhou, Zhuo Han, Songyang Zhang, Kai Chen, Zongwen Shen, and Jidong Ge. 2023. LawBench: Benchmarking legal knowledge of large language models. *arXiv preprint arXiv:2309.16289*.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, and 1 others. 2022. LoRA: Low-rank adaptation of large language models. *The International Conference on Learning Representations*, 1(2):3.

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

Zhigang Huang, Xinqiang Xie, Dong Ge, Chenqiu Cai, Lili Dou, and Tianyi Wang. 2021. Case recommendation based on knowledge graph of judicial cases. *Journal of Nanjing University Natural Science*, 57(6):1053–1063.

Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, and 1 others. 2024. GPT-4o system card. *arXiv preprint arXiv:2410.21276*.

Jinqi Lai, Wensheng Gan, Jiayang Wu, Zhenlian Qi, and Philip S. Yu. 2023. Large language models in law: A survey. *Preprint*, arXiv:2312.03718.

Haitao Li, You Chen, Qingyao Ai, Yueyue Wu, Ruizhe Zhang, and Yiqun Liu. 2024a. LexEval: A comprehensive chinese legal benchmark for evaluating large language models. *Advances in Neural Information Processing Systems*, 37:25061–25094.

Jun Li, Lu Qian, Peifeng Liu, and Taoxiong Liu. 2024b. Construction of legal knowledge graph based on knowledge-enhanced large language models. *Information*.

Luoqiu Li, Zhen Bi, Hongbin Ye, Shumin Deng, Hui Chen, and Huaixiao Tou. 2021. Text-guided legal knowledge graph reasoning. *Preprint*, arXiv:2104.02284.

Huixing Liang. 2020. *General Principles of Civil Law*. Law Press China.

Lei Liang, Mengshu Sun, Zhengke Gui, Zhongshu Zhu, Zhouyu Jiang, Ling Zhong, Yuan Qu, Peilong Zhao, Zhongpu Bo, Jin Yang, Huaidong Xiong, Lin Yuan, Jun Xu, Zaoyang Wang, Zhiqiang Zhang, Wen Zhang, Huajun Chen, Wenguang Chen, and

Jun Zhou. 2024. Kag: Boosting llms in professional domains via knowledge augmented generation. *Preprint*, arXiv:2409.13731.

Linhao Luo, Yuan-Fang Li, Gholamreza Haffari, and Shirui Pan. 2024. Reasoning on graphs: Faithful and interpretable large language model reasoning. *Preprint*, arXiv:2310.01061.

MiniCPM-Team. 2025. MiniCPM4: Ultra-efficient LLMs on end devices. *Preprint*, arXiv:2506.07900.

OpenAI. 2025. OpenAI o3-mini. https://openai.com/index/openai-o3-mini/.

Weixing Shen. 2004. Reconsideration on the constitution of civil legal relationship. *Journal of Comparative Law*, (1):42–54.

Supreme People's Court of the People's Republic of China. 2024. Gazette of the supreme people's court (2024). http://gongbao.court.gov.cn/Details/3ece7439305fa8bf1a7aae143d1598.html. Accessed: 2025-10-06.

Meng Wan, Feng Zhu, Benedict Amour, Hailong Tang, and 1 others. 2022. *The Civil Code of the People's Republic of China*. Springer.

Liming Wang. 2017. *General Principles of Civil Law*. Renmin University Press.

Liming Wang, editor. 2018. *Civil Law*. Renmin University Press.

Chaojun Xiao, Xueyu Hu, Zhiyuan Liu, Cunchao Tu, and Maosong Sun. 2021. Lawformer: A pre-trained language model for Chinese legal long documents. *AI Open*, 2:79–84.

Jiayu Xu. 2021. Legal text knowledge mining for civil law. Master's thesis, Harbin Institute of Technology.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, and 1 others. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Zhonghua Yuan. 2024. Reflections on the paradigm of legal relations. *Chinese Journal of Law*, 46(1):129–147.

Shengbin Yue, Wei Chen, Siyuan Wang, Bingxuan Li, Chenchen Shen, Shujun Liu, Yuxuan Zhou, Yao Xiao, Song Yun, Xuanjing Huang, and Zhongyu Wei. 2023a. DISC-LawLLM: Fine-tuning large language models for intelligent legal services. *Preprint*, arXiv:2309.11325.

Shengbin Yue, Wei Chen, Siyuan Wang, Bingxuan Li, Chenchen Shen, Shujun Liu, Yuxuan Zhou, Yao Xiao, Song Yun, Xuanjing Huang, and 1 others. 2023b. Disc-LawLLM: Fine-tuning large language models for intelligent legal services. *arXiv preprint arXiv:2309.11325*.

Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, Zheyan Luo, Zhangchi Feng, and Yongqiang Ma. 2024. LlamaFactory: Unified efficient fine-tuning of 100+ language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*.

A Taxonomy

The complete content of the Taxonomy can be found in the **Taxonomy** folder in attachment, which includes 265 different relationship types across 9 domains

B LexRel and SFT data

The LexRel we constructed and SFT data generated by GPT-4o and DeepSeek-R1 are included in the **LexRel** folder in attachment.

C Prompts

The complete definition of arguments and the prompts used in *arguments extraction* are as follows in Tables 5-8. The original Chinese version is in the **Prompts** folder of the file.

D Code

The source code we use is located in the **Code** folder. It includes code for generating via API and vllm to perform type and argument extraction, as well as code for conducting two types of evaluations. Additionally, it contains code for evaluating other downstream tasks. The training is implemented using the Llama-Factory (Zheng et al., 2024) framework, and the SFT configuration files are also included.

Types	Prompt
Legal Relations of Personality Rights	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The right holder himself/herself and all other subjects are the obligors (generally only the right holder is output).** **Object definition: The corresponding personal interests (output xx rights interests only).** **Content definition: In terms of rights, the subject enjoys the right to have their life, body, health, name, portrait, reputation, privacy and other personal interests not infringed upon by law. In terms of obligations, other civil subjects bear negative obligations such as not infringing upon, not insulting or defaming, and not disclosing relevant information. When necessary, they also have the obligation to actively protect and provide relief.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relations of Status Rights	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The right holder themselves and all other subjects are the obligors (generally, multiple right holders and all obligors related to the facts should be output simultaneously)** **Object Definition: Corresponding identity interests (output xx relationship interests only)** **Content definition: In terms of rights, the subject enjoys the rights of spousal, parental, guardianship, support and inheritance, etc. in accordance with the law due to the identity relationship. In terms of obligations, the corresponding subjects undertake identity obligations such as loyalty, assistance, support, education, and guardianship.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Tortious Legal Relations	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The infringer is the subject that implements the infringement act, and the infringed party is the subject that suffers damage due to the infringement act (generally, both should be output simultaneously).** **Object definition: The legitimate rights and interests of the infringed party (just output the name of the infringed right).** **Content definition: The rights enjoyed by the creditor: The right to request the debtor to perform or refrain from acting in accordance with the agreement or law. When a debtor fails to exercise his rights or infringes upon his creditor's rights, the right of subrogation or revocation may be exercised in accordance with the law. The obligations borne by the debtor: They should fulfill the obligation of payment as agreed, including delivering money, property or providing services, and may also bear the obligation of inaction.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.

Table 5: Definition and prompts for Arguments Extraction.

Types	Prompt
Legal Relations of Intellectual Property	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The right holder himself/herself and all other subjects are the obligors (generally only the right holder is output).** **Object definition: Intellectual achievements, including works, inventions and creations, trademarks, trade secrets, etc. (Output the name of the intellectual achievement and the corresponding rights to this achievement).** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relations of Real Rights	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The right holder himself/herself and all subjects related to the rights and obligations relationship are the obligors (generally the output right holder and other obligors who bear direct legal obligations in this fact).** **Object definition: Tangible objects, including immovable property, movable property and special objects. (Name of the output).** **Content definition: In terms of rights, the right holder enjoys ownership, usufructuary rights, security interests, etc. in accordance with the law, exercises direct control over the object through rights such as possession, use, income, and disposal, and can legally exclude interference from others. Immovable property rights come into effect through registration, while movable property rights take delivery as a prerequisite for their effectiveness. In terms of obligations, it includes completing the registration or delivery procedures in accordance with the law; It is prohibited to infringe upon the legitimate property rights of others. When property rights are illegally interfered with, it undertakes the obligation of protection such as removing obstacles, restoring the original state and returning the original property.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relations of Success	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: Heir and decedent. (Generally, both need to be output simultaneously). Object definition: Specific heritage content.** **Content definition: In terms of rights, heirs enjoy rights such as inheriting the estate, accepting bequests, and subrogation inheritance in accordance with the law or will. In terms of obligations, the heirs should divide the property in the spirit of friendly consultation, fulfill the obligations stipulated in the conditional will, notify the relevant parties and properly keep the estate, and also settle debts and taxes in accordance with the law.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.

Table 6: Definition and prompts for Arguments Extraction.

Types	Prompt
Contractual Legal Relations	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The parties to the contract (generally all parties need to be output simultaneously).** **Object Definition: The subject matter pointed to by the contractual obligations, including goods, acts, rights, etc. (output the performance of the contract). Content Definition: The rights enjoyed by the creditor: The right to request the debtor to perform or refrain from acting in accordance with the agreement or law. When a debtor fails to exercise his rights or infringes upon his creditor's rights, the right of subrogation or revocation may be exercised in accordance with the law. The obligations borne by the debtor: They should fulfill the obligation of payment as agreed, including delivering money, property or providing services, and may also bear the obligation of inaction.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relation of Unjust Enrichment	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: A beneficiary is the subject that gains benefits without legal basis, while a victim is the subject that suffers losses due to the beneficiary's acquisition of benefits (generally, both should be output simultaneously).** **Object definition: Interest, which refers to the property or benefit obtained by the beneficiary without legal basis. (Content of Output Benefits)** **Content definition: The rights enjoyed by the creditor: The right to request the debtor to perform or refrain from acting in accordance with the agreement or law; When a debtor fails to exercise his rights or infringes upon his creditor's rights, the right of subrogation or revocation may be exercised in accordance with the law. The obligations borne by the debtor: They should fulfill the obligation of payment as agreed, including delivering money, property or providing services, and may also bear the obligation of inaction.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relation of Negotiorum Gestio	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: A manager (creditor) refers to a person who, without being entrusted or bound by legal obligations, takes the initiative to manage the affairs of others to prevent their interests from being damaged. The beneficiary (debtor) refers to the party whose interests are attributed to the management of the affairs (usually both of them should be output simultaneously).** **Object definition: The actions of the manager (such as repairing houses on behalf of others, providing assistance to the sick and injured, etc.) and the necessary expenses or compensation for losses arising therefrom. (Content of the Output act)** **Content definition: The rights enjoyed by the creditor: The right to request the debtor to perform the act or refrain from acting in accordance with the agreement or law; When a debtor fails to exercise his rights or infringes upon his creditor's rights, the right of subrogation or revocation may be exercised in accordance with the law. The obligations borne by the debtor: They should fulfill the obligation of payment as agreed, including delivering money, property or providing services, and may also bear the obligation of inaction.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.

Table 7: Definition and prompts for Arguments Extraction.

Types	Prompt
Legal Relation of Letters of Credit	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: The primary parties are the Applicant, the Issuing Bank, and the Beneficiary, but there are eight parties in total, including: the Advising Bank, the Confirming Bank, the Negotiating Bank, the Paying Bank, and the Reimbursing Bank.** **Object definition: The delivery of documents and the promise of payment.** **Content Definition: 1. The contractual sales relationship between the Applicant and the Beneficiary. 2. The principal-agent contractual relationship between the Issuing Bank and the Applicant. 3. The relationship between the Issuing Bank and the Beneficiary (a subject of theoretical and practical debate). Generally, when a revocable letter of credit is issued, there is no contractual relationship between the Issuing Bank and the Beneficiary. However, when an irrevocable letter of credit is issued and received by the Beneficiary, a binding independent contractual relationship is formed between them. 4. The principal-agent relationship between the Advising Bank and the Issuing Bank. 5. The contractual relationship between the Issuing Bank and the Paying/Negotiating Bank arises when the Issuing Bank designates or authorizes another bank to pay or negotiate, and that bank accepts the role. Otherwise, no legal relationship exists. 6. Typically, there is no legal relationship between the Beneficiary and the Paying/Negotiating Bank, unless a bank (other than the Issuing Bank) pays or negotiates the Beneficiary's draft under the Issuing Bank's authorization. In such cases, a creditor-debtor legal relationship is established under negotiable instruments law.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relation of Independent Guarantees	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: Issuer (bank/financial institution), beneficiary, applicant (can be an instructor).** **Object definition: The "document" on which the claim request is based.** **Content definition: It refers to the bill rights that the subject enjoys and the bill obligations it should undertake in accordance with the law. The right to a bill is the right enjoyed by the right holder to request the obligor to pay the amount of the bill, including the right to request payment and the right of recourse. Bill obligations refer to the responsibilities that the obligor must fulfill to satisfy the rights claims enjoyed by the right holder based on the bill, such as payment obligations, acceptance obligations, and guarantee payment obligations, etc.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.
Legal Relation of Bills	Please extract the subject, object, and content of **{relation type}** from the given judgment documents without explanation or description. **Subject definition: A subject is specific and mainly includes: the drawer, the payee, the payer, the holder of the bill, the acceptor, the endorser, the guarantor, the participant, etc.** **Object definition: A certain amount of currency.** **Content definition: The "underlying transaction contract" relationship between the applicant and the beneficiary; The "independent guarantee contract" relationship between the issuer and the beneficiary; The "application/recovery contract" relationship between the issuer and the applicant. considerations correspond to: the issuer's unconditional payment obligation, the principle of document review and "apparent conformity", the right of recovery after payment, the right of defense against fraud exceptions, and the conditions for termination or invalidation of the guarantee.** **Judgment text: {}**. Output requirements: 1. Each set of results is presented as a standard JSON object. 2. Required fields: {'subject': '', 'object': '', 'content': ''}. 3. Each set is in a separate paragraph without numbering or sorting. 4. Only one set needs to be output for the same subject and object.

Table 8: Definition and prompts for Arguments Extraction.