

Towards Understanding Causality in Judicial Narratives: A Simple Approach using Event-Centric Knowledge Graphs

Ngoc-Duy Mai^{1,2,3,4,*}, Ha-Thanh Nguyen⁵, Ken Satoh³ and Hideaki Takeda^{1,2}

¹The Graduate University for Advanced Studies, SOKENDAI, Kanagawa, Japan

²National Institute of Informatics, Tokyo, Japan

³Center for Juris-Informatics, ROIS-DS, Tokyo, Japan

⁴Research and Development Center for Large Language Models, NII, Tokyo, Japan

⁵VinUniversity, Hanoi, Vietnam

Abstract

Judicial decisions are typically presented as complex narratives that describe sequences of events, the actors involved, and the reasons underlying legal outcomes. Transforming such narratives into structured representations remains challenging, especially when causal and justificatory relations are implicit in text. We propose a simple event-centric causality knowledge graph framework for judicial narratives. The framework integrates named entity recognition, coreference resolution, event extraction, temporal ordering, and causality detection to construct a graph connecting entities, events, time expressions, and explicit causal explanations. We evaluate rule-enhanced and LLM-based extraction methods on annotated cases from the ECtHR-A dataset and analyze their strengths and limitations. Results show that rule-enhanced methods provide stable structural extraction, while LLMs better capture some semantic attributes but remain inconsistent for coreference and cross-event causality. Competency question analysis further suggests that the resulting graph can support core factual information needs in legal case analysis, providing an interpretable foundation for causality-aware legal knowledge representation.

Keywords

Knowledge Graph, Causality, Event-Centric, Legal, Natural Language Processing

1. Introduction

The legal domain faces growing challenges due to the rapid expansion of legal information, increasing regulatory complexity, and rising demands for timely and well-founded decisions. Legal professionals must analyze large volumes of materials, including statutes, judicial decisions, and evidence. However, much of this information is expressed in unstructured natural language, where facts, events, and legal relations are embedded in complex narratives, making large-scale computational analysis difficult. Legal Knowledge Graphs (LKGs) have been proposed to transform such texts into structured, machine-readable representations for information access, analysis, and explainable decision-making. However, constructing effective LKGs remains challenging due to complex legal language, implicit relations, and heterogeneous narrative styles. These difficulties are especially important for causality, which plays a central role in legal analysis [1]. Existing work [2, 3] has largely focused on entities or isolated relations, while explicit causal structure integrated with event representations remains limited.

In this paper, we propose an event-centric causality knowledge graph framework for judicial narratives. The framework jointly models events, temporal relations, and causal explanations by integrating information extraction, event modeling, and causality detection. We focus on the factual and event-centric layer of judicial narratives, treating causal relations as text-grounded narrative explanations rather than formal legal reasoning. This work serves as a first step toward connecting real-world actions to legal actions through an event-centric KG, supporting interpretable querying over case facts and causal relationships.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the proposed framework. Section 4 reports experimental results. Section 5 analyzes competency questions for the constructed KG. Finally, Section 6 concludes the paper and discusses future directions.

Posters, Demos, Blue Sky, and Tutorials at SEMANTiCS 2026, Sep 2026, Ghent, Belgium

*Corresponding author.

✉ duymn@nii.ac.jp (N. Mai)

ORCID 0009-0008-8525-6238 (N. Mai); 0000-0003-2794-7010 (H. Nguyen); 0000-0002-9309-4602 (K. Satoh); 0000-0002-2909-7163 (H. Takeda)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. Related Work

Recent work on LKGs has moved from manually curated, ontology-driven resources toward more automated construction pipelines that extract entities, relations, and structures from case law and statutory text using neural models and LLM-assisted extraction [4, 5]. LKGs are also used as grounding structures for legal NLP and LLM-based systems, where statutory, citation, and graph-based information can improve retrieval and coverage [6].

Beyond entity-centric KGs, Event Knowledge Graphs (EKGs) model events as fundamental units of representation, together with temporal links and event–entity participation. Representative resources such as EventKG [7] and ASER [8] show how event-centric structures can support semantic search, question answering, recommendation, and text generation [9]. In the legal domain, judicial case knowledge graphs centered on event extraction have been proposed to structure case facts and support legal knowledge discovery [10]. However, these approaches mainly model event structure, such as temporal order and participation, while leaving causal and justificatory relations largely implicit.

From a knowledge graph viewpoint, Causal Knowledge Graphs (CKGs) explicitly represent cause–effect relations as graph primitives, enabling systems to identify, traverse, and reason over causal chains rather than relying on isolated textual cues [11]. This idea has been explored in several domains, such as finance, where FinCaKG [12] and FinCaKG-Onto [13] construct causal graphs from financial texts, and medicine, where causal relations are modeled to support domain-specific reasoning [14]. Within legal AI, causality has also been incorporated into task-oriented graph constructions, including explainable charge prediction, robust legal judgment prediction, and legal case retrieval [2, 15, 16]. However, these approaches typically use causal graphs for specific downstream tasks or domain settings. Up until now, existing work has not yet integrated event representations, temporal structure, and causal explanations into a unified knowledge graph for judicial narratives. The construction of such event-centric causality KGs for lengthy legal cases, involving multiple actors, events, and implicit or defeasible causal relations, therefore remains underexplored.

3. Method

Figure 1 illustrates the pipeline for constructing an event-centric causality knowledge graph from judicial narratives. After basic preprocessing, we apply named entity recognition, coreference resolution, event extraction, and causality detection. The extracted entities, events, temporal information, and causal relations are then integrated into a structured graph.

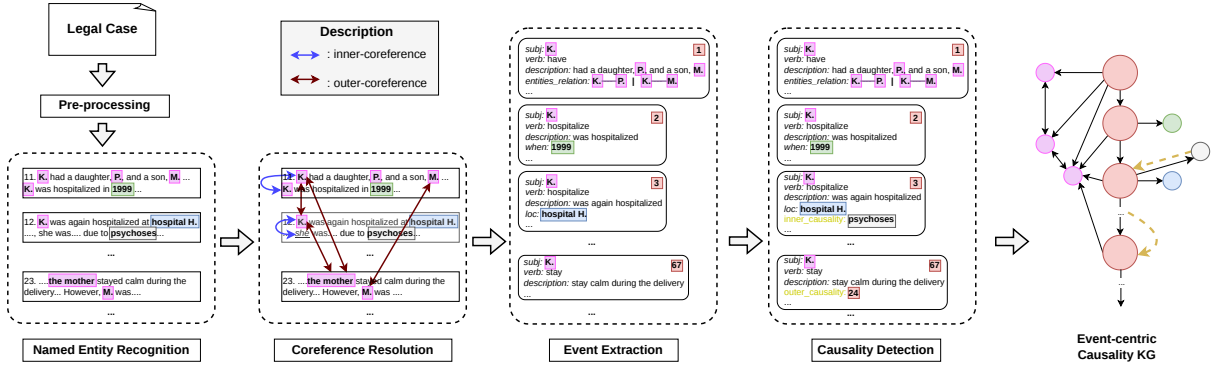


Figure 1: Pipeline of constructing the Event-centric Causality Knowledge Graph.

3.1. Named Entity Recognition

NER is used to identify entities and contextual information required for event and causal modeling. The label set includes PER, ORG, LOC, TIME, and OTHERS. The OTHERS category covers legally or causally relevant concepts that do not fit standard NER labels, such as laws, legal provisions, legal measures, illnesses, diagnoses, conditions, and circumstances.

We compare two approaches. The rule-enhanced approach augments spaCy [17]’s `en_core_web_sm` model with domain-specific rules for abbreviated names, grouped mentions, and legal roles such as

“the mother” or “the authorities”. The LLM-based approach uses zero-shot prompting to extract entities in a structured format¹.

3.2. Coreference Resolution

Coreference resolution links different mentions of the same entity, such as abbreviated names, pronouns, and descriptive expressions. We again compare rule-enhanced and LLM-based approaches. The rule-enhanced method uses Stanza [18] with preprocessing for anonymized legal references and post-processing to clean and align mention clusters. The LLM-based method uses zero-shot prompting to cluster coreferent mentions within factual paragraphs¹. We first resolve *inner-coreference* within each factual paragraph, then apply *outer-coreference* to align entities across paragraphs in the same case. For ambiguous generic references such as “the mother”, we select the most recent compatible referent, assuming that nearby mentions are more likely to refer to the same entity.

3.3. Event Extraction

Event extraction captures legal procedures and circumstances described in the narrative. Each event is represented by core attributes, including subj, verb, and description, together with participating entities, temporal attributes, and event-level metadata. Participating entities are derived from NER and coreference outputs, while temporal attributes include when_date, when_period, and when_frequency. Events with explicit dates or periods are marked as *main_event*; otherwise, an *expected_when* value is assigned from the most recent compatible time expression.

We also record higher-level attributes, including *event_type*, *decision_event*, and *level*. Following prior work [19], *event_type* distinguishes *circumstance* and *procedure*. Procedural events containing decision-related expressions are marked as *decision_event*. The *level* attribute distinguishes factual events from reporting or mental-state events. Event extraction is implemented using both dependency-based rule-enhanced extraction and zero-shot LLM extraction¹.

3.4. Causality Detection

We distinguish two types of causality. *Inner-causality* occurs within a single event and is usually signaled by expressions such as *because*, *since*, or *on account of*. *Outer-causality* captures causal or justificatory relations between distinct events, where one event explains, supports, or leads to another, especially in decision-related contexts. To reduce the number of candidate event pairs, heuristic rules first select potential causal spans and event pairs. LLMs are then prompted to perform causal span extraction for inner-causality and binary cause–effect classification for outer-causality¹.

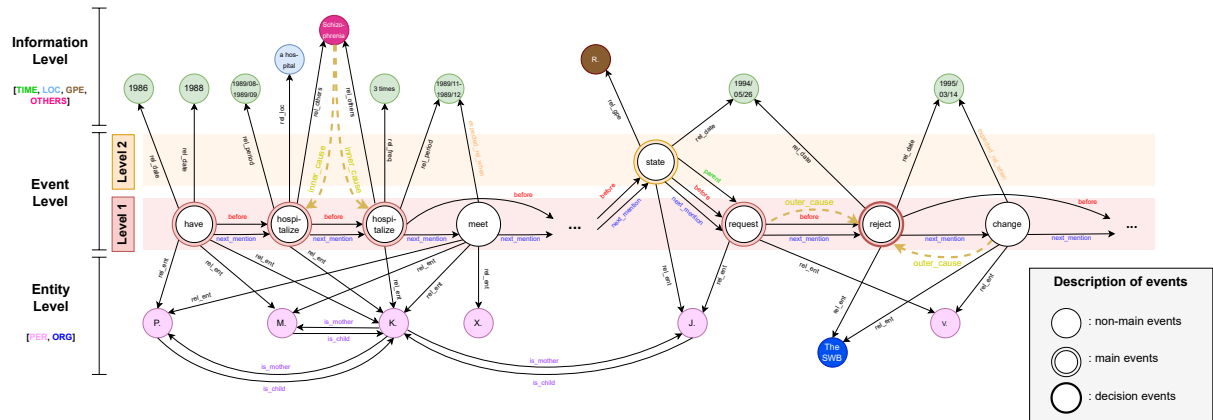


Figure 2: Example of the constructed Event-centric Causality Knowledge Graph from a case.

Figure 2 shows the constructed graph structure. The **information level** contains TIME, LOC, GPE, and OTHERS nodes. The **event level** represents the case timeline using *before* and *next_mention* relations, links level 2 events to level 1 events through *parent* relations, and encodes causal relations between events or information nodes. The **entity level** contains PER and ORG entities and their explicit relationships when available.

¹<https://github.com/AESir2412/legal-causality-kg-prompt/blob/main/Prompt%20LCKG.pdf>

4. Results

We evaluate the framework on the ECtHR-A dataset [20], which contains approximately 11,000 European Court of Human Rights cases. Due to the cost of manual annotation, we randomly select five training cases and annotate entities, coreference clusters, events, temporal information, and causal relations using Label Studio. For LLM-based extraction, cases are processed in batches of five factual paragraphs due to input length constraints.

Table 1

Statistics of selected cases and constructed KGs.

Case	Case statistics				Entities		Information				Causality		Temporal edges		KG size	
	Facts	Words / fact	Events	Events / fact	PER	ORG	TIME	LOC	GPE	OTHERS	Inner	Outer	next_mention	before	Nodes	Edges
0	83	82.9	394	4.75	66	23	113	24	4	9	14	26	393	176	337	1726
451	16	72.2	71	4.44	17	6	25	1	0	3	3	17	70	34	70	328
713	45	107.8	250	5.56	52	11	61	17	3	7	8	33	249	77	192	1185
752	14	42.4	43	3.07	7	8	39	0	2	0	3	0	42	38	66	191
1036	14	127.6	66	4.71	17	15	22	1	4	2	1	6	65	31	77	249

Dataset statistics. Table 1 shows that the selected cases vary substantially in length and graph size. Each fact contains multiple events on average, confirming the dense event structure of judicial narratives. Outer-causality is generally more frequent than inner-causality, suggesting that many causal links connect distinct factual or procedural events.

Table 2

NER performance.

Metric	Rule	GPT	Gemini	Qwen
Micro-P	0.825	0.857	0.816	0.785
Micro-R	0.610	0.533	0.723	0.602
Micro-F ₁	0.702	0.657	0.767	0.681
Macro-P	0.472	0.619	0.427	0.671
Macro-R	0.473	0.408	0.432	0.554
Macro-F ₁	0.443	0.477	0.422	0.595

Table 3

Coreference resolution performance.

Metric	Rule	GPT	Gemini	Qwen
MUC F ₁	0.253	0.124	0.378	0.108
B ³ F ₁	0.782	0.598	0.622	0.576
CEAF _e F ₁	0.273	0.240	0.382	0.256
CoNLL F ₁	0.436	0.321	0.461	0.313

Named entity recognition. Table 2 shows that Gemini achieves the best micro-F₁, indicating strong overall performance, while Qwen obtains the best macro-F₁, suggesting more balanced extraction across entity categories.

Coreference resolution. As shown in Table 3, Gemini achieves the best CoNLL F₁, while the rule-enhanced method remains competitive. However, all systems remain below 0.5 CoNLL F₁, showing that abbreviated names, indirect references, and long-distance mentions remain difficult.

Table 4

F₁ scores of event detection and event attribute extraction.

Core and entity attributes					Temporal and higher-level attributes				
Attr.	Rule	GPT	Gemini	Qwen	Attr.	Rule	GPT	Gemini	Qwen
Event Detection	0.790	0.765	0.479	0.534	when_date	0.652	0.561	0.220	0.397
subj	0.694	0.703	0.427	0.457	when_period	0.083	0.057	0.014	0.026
verb	0.787	0.678	0.425	0.475	when_frequency	0.039	0.041	0.000	0.027
description	0.687	0.582	0.225	0.363	event_type	0.352	0.580	0.359	0.180
subj_contain	0.611	0.680	0.450	0.468	main_event	0.765	0.713	0.358	0.423
description_contain	0.327	0.365	0.285	0.277	decision_event	0.733	0.700	0.441	0.468
other_entities_contain	0.000	0.018	0.010	0.000	expected_when	0.422	0.264	0.027	0.029
entities_relation	0.000	0.005	0.003	0.005	level	0.769	0.699	0.451	0.414
					voice	0.783	0.741	0.461	0.504
					polarity	0.782	0.677	0.422	0.323

Event extraction. Table 4 shows that the rule-enhanced method performs best on event detection and most structural attributes, such as verb, description, main_event, and decision_event. GPT performs better on semantically richer attributes such as subj and event_type, while Gemini and Qwen generally struggle to produce complete event structures.

Table 5

Performance of LLMs on causality detection.

Causality Type	GPT	Gemini	Qwen
Inner-causality	0.8000	0.6774	0.7935
Outer-causality	0.4331	0.3268	0.4929

Causality detection. Table 4 shows that LLMs perform well on inner-causality, where explicit causal markers are often available. In contrast, outer-causality remains difficult for all models, suggesting that cross-event causal reasoning requires deeper contextual understanding.

5. Competency Questions Analysis

To assess the practical relevance of the proposed design of the KG, we formulate competency questions based on discussions with legal experts. These questions reflect common information needs in legal case analysis and indicate whether the current graph structure can support them.

Table 6

Competency questions and support by the proposed knowledge graph.

CQ	Question	Support	Reason	CQ	Question	Support	Reason
CQ1	What actions occurred, and when?	Supported	Events and time attributes.	CQ6	What legal consequences follow?	Supported	Event causal links.
CQ2	Who performed a specific action?	Supported	Entity–event links.	CQ7	Relationships between legal actions?	Not supported	Requires legal reasoning.
CQ3	What events are causally related?	Supported	Causal relations.	CQ8	Is there sufficient proof?	Partially	No proof standard.
CQ4	What evidence or justification is provided?	Supported	Causality and OTHERS.	CQ9	Why was an ECHR right violated?	Not supported	Requires norms and provisions.
CQ5	Is an action legally relevant?	Not supported	Requires legal norms.	CQ10	Who are the main actors?	Partially	Approximated by connectivity.

Most supported questions concern factual event analysis: identifying actions, actors, temporal order, causal links, and justifications. Partially supported questions can be approximated using graph structure, such as entity connectivity or causal links, but are not fully resolved. Unsupported questions require legal norms, statutory provisions, doctrinal reasoning, or standards of proof, which are outside the current factual and text-grounded scope of the graph. Nevertheless, the supported questions suggest that the KG could serve as an exploratory tool for inspecting case timelines, actors, and explicit causal explanations before deeper legal analysis.

6. Conclusion

This paper presents a simple event-centric causality knowledge graph framework for transforming judicial narratives into structured representations of events, temporal relations, and explicit causal explanations. By integrating entity recognition, coreference resolution, event extraction, and causality detection, the framework enables interpretable querying over the factual progression and causal structure of legal cases. Experimental results show that rule-enhanced methods provide stable extraction of structural information, while LLM-based methods better capture some semantic attributes but remain inconsistent, especially for coreference resolution and cross-event causality. Competency question analysis further shows that the graph supports core factual information needs in legal case analysis. However, the current evaluation is based on a small number of annotated cases, and the construction methods remain relatively simple at this stage. The current framework is also limited to the factual layer and does not yet incorporate legal norms, statutory provisions, or doctrinal reasoning required to model legal conclusions. Moreover, legal causality involves deeper notions such as proximate cause and foreseeability, which often require commonsense knowledge and professional judgment [1]. Future work will focus on scaling the evaluation, integrating structured extraction with richer legal reasoning, and testing this with other legal domains.

Acknowledgments

This work was supported by the “R&D Hub Aimed at Ensuring Transparency and Reliability of Generative AI Models” project of the MEXT, by JSPS KAKENHI Grant Numbers, 25H00522 and 25H01112, and by JST as part of Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE), Grant Number JPMJAP25B2.

Declaration on Generative AI

For the preparation, we used ChatGPT to assist with paraphrasing and refining parts of the manuscript. All content was then reviewed and edited, and we take full responsibility for the final version.

References

- [1] J. Hellner, Causality and causation in law, *Scandinavian studies in law* 40 (2000) 111–134.
- [2] X. Liu, D. Yin, Y. Feng, Y. Wu, D. Zhao, Everything has a cause: Leveraging causal inference in legal text analysis, *arXiv preprint arXiv:2104.09420* (2021).
- [3] P. Prince Tritto, H. Ponce, Causal artificial intelligence in legal language processing: A systematic review, *Entropy* 27 (2025) 351.
- [4] J. Zhou, X. Chen, H. Zhang, Z. Li, Automatic knowledge graph construction for judicial cases, *arXiv preprint arXiv:2404.09416* (2024).
- [5] J. Li, L. Qian, P. Liu, T. Liu, Construction of legal knowledge graph based on knowledge-enhanced large language models, *Information* 15 (2024) 666.
- [6] S. Shukla, T. Tyagi, A. S. Bisht, A. Sharma, B. Agarwal, Nyaygraph: A knowledge graph enhanced approach for legal statute identification in indian law using large language models, in: *Proceedings of the Natural Legal Language Processing Workshop 2025*, 2025, pp. 147–156. URL: <https://aclanthology.org/2025.nllp-1.11/>.
- [7] S. Gottschalk, E. Demidova, Eventkg: A multilingual event-centric temporal knowledge graph, in: *European semantic web conference*, Springer, 2018, pp. 272–287.
- [8] H. Zhang, X. Liu, H. Pan, Y. Song, C. W.-K. Leung, Aser: A large-scale eventuality knowledge graph, in: *Proceedings of the web conference 2020*, 2020, pp. 201–211.
- [9] S. Guan, X. Cheng, L. Bai, F. Zhang, Z. Li, Y. Zeng, X. Jin, J. Guo, What is event knowledge graph: A survey, *IEEE Transactions on Knowledge and Data Engineering* 35 (2022) 7569–7589.
- [10] B. Zhao, Y. Zhao, Y. Mao, A method for judicial case knowledge graph construction based on event extraction, in: *Proceedings of the 2024 9th International Conference on Intelligent Information Technology*, 2024, pp. 62–69.
- [11] U. Jaimini, A. Sheth, Causalkg: Causal knowledge graph explainability using interventional and counterfactual reasoning, *IEEE Internet Computing* 26 (2022) 43–50.
- [12] Z. Xu, H. Takamura, R. Ichise, Fincakg: A framework to construct financial causality knowledge graph from text, *International Journal of Semantic Computing* (2025) 1–20.
- [13] Z. Xu, R. Ichise, Fincakg-onto: the financial expertise depiction via causality knowledge graph and domain ontology, *Applied Intelligence* 55 (2025) 1–17.
- [14] S. Gopalakrishnan, L. Garbayo, W. Zadrozny, Causality extraction from medical text using large language models (llms), *Information* 16 (2024) 13.
- [15] H. Chen, L. Zhang, Y. Liu, F. Chen, Y. Yu, Knowledge is power: understanding causality makes legal judgment prediction models more generalizable and robust, *arXiv preprint arXiv:2211.03046* (2022).
- [16] K. Zhang, C. Chen, Y. Wang, Q. Tian, L. Bai, Cfgl-lcr: A counterfactual graph learning framework for legal case retrieval, in: *Proceedings of the 29th ACM SIGKDD Conference on knowledge discovery and data mining*, 2023, pp. 3332–3341.
- [17] M. Honnibal, I. Montani, S. Van Landeghem, A. Boyd, et al., *spacy: Industrial-strength natural language processing in python* (2020).
- [18] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, C. D. Manning, Stanza: A python natural language processing toolkit for many human languages, *arXiv preprint arXiv:2003.07082* (2020).
- [19] E. Filtz, M. Navas-Loro, C. Santos, A. Polleres, S. Kirrane, Events matter: Extraction of events from court decisions, in: *Legal Knowledge and Information Systems*, IOS Press, 2020, pp. 33–42.
- [20] I. Chalkidis, I. Androutsopoulos, N. Aletras, Neural legal judgment prediction in english, *arXiv preprint arXiv:1906.02059* (2019).