

# SHIELD: LLM-Driven Schema Induction for Predictive Analytics in EV Battery Supply Chain Disruptions

Zhi-Qi Cheng<sup>1</sup> Yifei Dong<sup>2</sup> Aike Shi<sup>3</sup> Wei Liu<sup>4</sup> Yuzhi Hu<sup>5</sup>  
 Jason O'Connor<sup>1</sup> Alexander Hauptmann<sup>1</sup> Kate Whitefoot<sup>1</sup>

<sup>1</sup>Carnegie Mellon University <sup>2</sup>Columbia University <sup>3</sup>Georgia Institute of Technology  
<sup>4</sup>University of Michigan, Ann Arbor <sup>5</sup>Boston University

Project Page: <https://f1y1113.github.io/MFI/>

## Abstract

The electric vehicle (EV) battery supply chain's vulnerability to disruptions necessitates advanced predictive analytics. We present SHIELD (Schema-based Hierarchical Induction for EV supply chain Disruption), a system integrating Large Language Models (LLMs) with domain expertise for EV battery supply chain risk assessment. SHIELD combines: (1) LLM-driven schema learning to construct a comprehensive knowledge library, (2) a disruption analysis system utilizing fine-tuned language models for event extraction, multi-dimensional similarity matching for schema matching, and Graph Convolutional Networks (GCNs) with logical constraints for prediction, and (3) an interactive interface for visualizing results and incorporating expert feedback to enhance decision-making. Evaluated on 12,070 paragraphs from 365 sources (2022–2023), SHIELD outperforms baseline GCNs and LLM+prompt methods (e.g. GPT-4o) in disruption prediction. These results demonstrate SHIELD’s effectiveness in combining LLM capabilities with domain expertise for enhanced supply chain risk assessment.

## 1 Introduction

The expected widespread adoption of electric vehicles (EVs) is threatened by risks associated with the geographic and economic concentration of critical battery minerals, such as lithium, cobalt, and nickel. To enhance the resilience of the EV battery supply chain, manufacturers must anticipate disruptions caused by natural disasters and geopolitical tensions. Proactive strategies and supply diversification are essential to mitigate these risks<sup>1</sup>.

Completed by Y. Dong and Y. Hu during their remote visit, and by A. Shi and W. Liu during their internships at CMU. Z. Cheng, Y. Dong, A. Shi, W. Liu, and Y. Hu contributed equally. J. O'Connor, A. Hauptmann, and K. Whitefoot provided guidance. Detailed contributions are in Appx. L.

<sup>1</sup>[https://nncta.org/\\_files/documents/chapter4-energy-critical-materials.pdf](https://nncta.org/_files/documents/chapter4-energy-critical-materials.pdf)

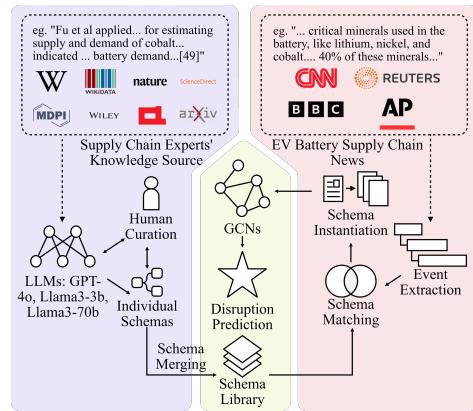


Figure 1: SHIELD’s process for EV battery supply chain disruption prediction. The framework integrates LLM-driven schema learning with expert curation, enabling robust event extraction and prediction from diverse news sources. This approach uniquely combines LLM capabilities with domain expertise, enhancing both predictive accuracy and interpretability for proactive supply chain risk management.

Traditional supply chain risk management approaches, which rely on rule-based reasoning and agent-based simulations (Gallab et al., 2019; Pino et al., 2010; Giannakis and Louis, 2011, 2016; Blos et al., 2015), often fall short in predictive accuracy and adaptability to dynamic market conditions. While machine learning (ML) and deep learning (DL) techniques have enhanced predictive performance (Hegde and Rokseth, 2020; Ruz et al., 2020; Aljohani, 2023; Silva et al., 2017; Garvey et al., 2015; Carbonneau et al., 2008), they frequently sacrifice interpretability, limiting their practical application. Recent studies employing large language models (LLMs) in supply chain management (Ray, 2023; Wang et al., 2022a; Du et al., 2022; Shi et al., 2024; Dror et al., 2022; Li et al., 2023) have focused on improving predictions but struggle to fully grasp complex domain-specific supply chain knowledge. This limitation often leads to hallucinations and inaccuracies which, coupled with limited interpretability, hinder the generation of actionable insights crucial for effective risk management.

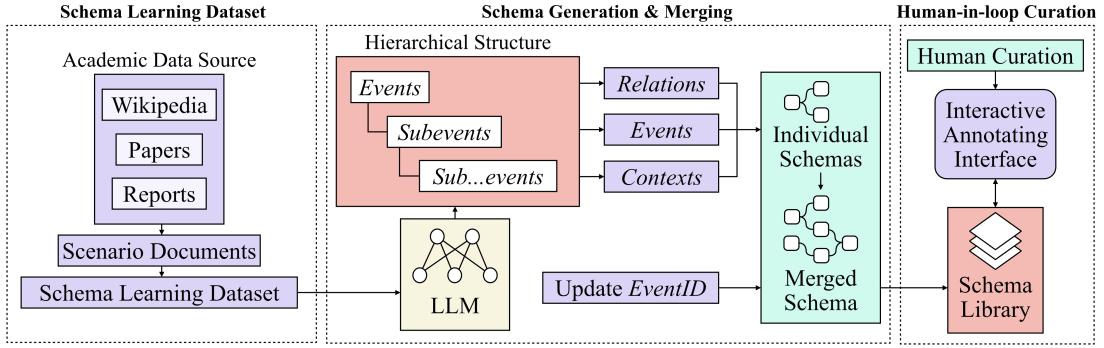


Figure 2: Overview of the supply chain schema construction process, illustrating the collection of diverse sources, schema extraction using large language models, and the integration into a unified schema library.

To address these challenges, we introduce SHIELD (Schema-based Hierarchical Induction for EV supply chain Disruption), a two-stage framework that integrates LLMs and domain expertise for predictive analytics in EV battery supply chains (Fig. 1):

1. *Schema Learning* (Sec. 2): We leverage LLMs (GPT-4o, Llama3-3b, Llama3-70b) to construct a comprehensive schema library—a structured representation of supply chain components and their relationships—from diverse sources. An interactive system integrates expert knowledge, distilling supply chain expertise from specialized documents. This approach ensures analyses align with domain knowledge, capturing EV battery supply chain complexities for accurate, interpretable predictions that adapt to industry dynamics through continuous refinement.
2. *Disruption Analysis* (Sec. 3): Building on our schema learning, we develop a comprehensive disruption prediction system. This system integrates fine-tuned RoBERTa (Liu et al., 2019) for event detection, multi-dimensional similarity for matching, and Graph Convolutional Networks (GCNs) with logical constraints for impact analysis. The resulting end-to-end system enables precise event extraction and reliable predictions in complex supply chains, mitigating LLM hallucination risks while maintaining efficiency. This approach offers a scalable solution for real-time supply chain risk assessment and mitigation.

Evaluated on 12,070 paragraphs from 365 sources (2022-2023) (Sec. 4), SHIELD outperforms baseline GCNs and LLM+prompt methods (e.g., GPT-4o) in disruption prediction. By integrating LLM capabilities with domain expertise, this framework enhances supply chain risk assessment. Key contributions include: (1) an LLM-expert integration methodology for accurate, interpretable predictions; (2) a schema learning and news evaluation dataset spanning the EV battery lifecycle; (3) an interactive schema curation system; and (4) advanced analytical techniques for supply chain analysis. SHIELD offers a promising approach in supply chain risk management, addressing evolving challenges across the EV industry and beyond.

butions include: (1) an LLM-expert integration methodology for accurate, interpretable predictions; (2) a schema learning and news evaluation dataset spanning the EV battery lifecycle; (3) an interactive schema curation system; and (4) advanced analytical techniques for supply chain analysis. SHIELD offers a promising approach in supply chain risk management, addressing evolving challenges across the EV industry and beyond.

## 2 Schema Learning for EV Supply Chains

**Schema Learning Dataset.** Our dataset comprises 239 diverse sources: 200 academic papers, 22 industry reports, and 17 Wikipedia entries (Fig. 5 and Fig. 6). This collection provides an up-to-date view of the EV battery supply chain, covering advanced battery technologies (e.g. LFP, NiMH), production processes, and six key raw materials. We categorized events into 8 categories, three with long-term impacts, subdivided into 18 subcategories. Our analysis includes five-year price trends for all materials, correlated with 39 significant supply chain events. Industry expert feedback refined our categorization into 11 main categories with 27 subcategories, each illustrated with 1-2 real-world events (Tab. 6). The academic dataset was distilled from 239 sources to 125 highly relevant entries. This dataset of over 1,000 events spans the EV battery lifecycle, enabling our methods to acquire expert knowledge for accurate, real-world predictions. More details are in Appx. B.1.

**Schema Generation & Merging.** Building upon our collected dataset, our Schema Learning System facilitates the extraction, visualization, and management of schemas from the 125 diverse textual sources (Fig. 2). The process begins with data cleaning using regular expressions and a locally deployed Llama3-8b model. Subsequently, we em-

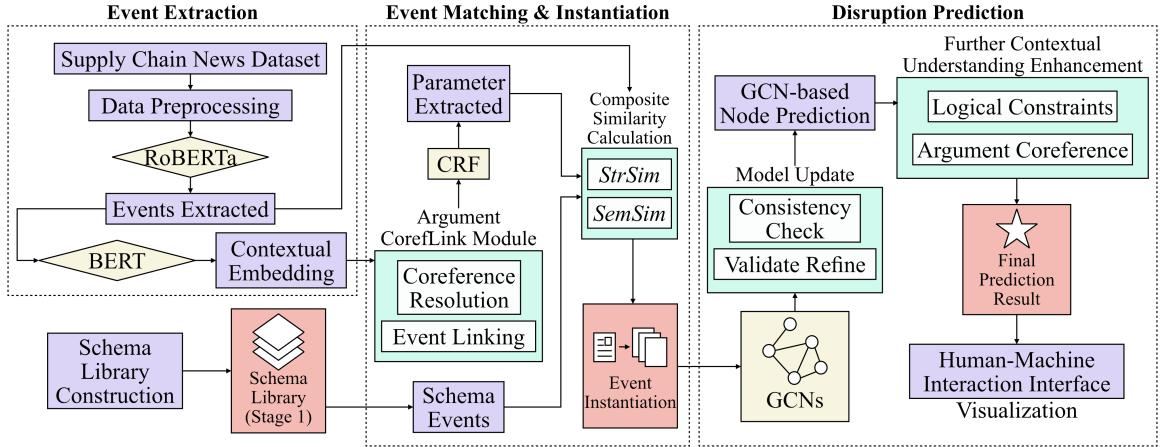


Figure 3: Overview of the supply chain disruption prediction pipeline, illustrating the integration of GCN-based predictions, constrained prediction refinement, and argument coreference resolution.

ploy GPT-4o, Llama3-3b, and Llama3-70b with specific prompts to extract hierarchical structures (**H**) capturing main events (**E**) and sub-events ( $E_{sub}$ ). More details are in Appx. C.

The extracted structures are then converted into individual schemas ( $S_i$ ) and visualized as graphs, demonstrating the hierarchical nature of the schemas and the relationships between main events and sub-events. These schemas are then integrated into a single library ( $S_{final}$ ), aggregating contexts ( $C_{final} = \bigcup_{i=1}^n C_i$ ), merging events ( $E_{final} = \bigcup_{i=1}^n E_i$ ), and updating event IDs for relations ( $R_{final} = \bigcup_{i=1}^n R_i$ ). The detailed schema generation and merging algorithm is provided in Appx. D.

To ensure efficient retrieval and updates, a dedicated Database & Storage module manages schema storage, while the Schema Management System incorporates a Schema Viewer, Editor, collaboration tools, and AI-driven suggestions are built to manage and annotate schemas (Appx. E). This human-in-the-loop curated framework streamlines schema extraction and management, enabling interactive knowledge extraction from structured documents, leveraging supply chain experts’ insights.

### 3 Dynamic Disruption Analysis

**Supply Chain News Dataset.** We developed an EV Supply Chain News Dataset (January 2022 - December 2023) to evaluate our system’s real-world performance (Appx. B.2). The dataset comprises 247 articles from major news outlets and 118 enterprise reports from EV battery-related companies (Fig. 7 and Fig. 8). After preprocessing—including text extraction, language standardization, and noise reduction—we obtained Meta data with 3,022 para-

graphs. We then fused international news with contemporaneous corporate stories in the meta data, creating 354 diverse documents comprising 12,070 paragraphs. The final dataset contains approximately 660K words (Table 9), providing a robust foundation for evaluating supply chain disruption detection and analysis. Comprehensive replication details, including the full dataset and preprocessing pipeline, are provided to ensure reproducibility.

**Event Extraction.** Our pipeline extracts multi-faceted events from textual data, focusing on their impact on the EV battery supply chain. We begin with custom-trained SpaCy models<sup>2</sup> for tokenization, sentence segmentation, named entity recognition, and dependency parsing (Appx. F).

Building on this, we deploy a fine-tuned RoBERTa model for cross-sentence event detection:

$$\text{EventDetect}_{\text{multi-sentence}}(\mathbf{T}) \rightarrow \mathbf{E}_C \quad (1)$$

where  $\mathbf{T}$  represents the input text and  $\mathbf{E}_C$  the detected events. These events are then enriched with contextual information using BERT:

$$\text{BERT}_{\text{context}}(\mathbf{E}_C) \rightarrow \mathbf{C}_E \quad (2)$$

generating contextual embeddings  $\mathbf{C}_E$ . To enhance analytical coherence, we implement coreference resolution and event linking:

$$\text{CorefLink}(\mathbf{E}_C) \rightarrow \mathbf{E}_L \quad (3)$$

This critical step, yielding linked events  $\mathbf{E}_L$ , maintains contextual continuity across documents. Subsequently, Conditional Random Fields (CRFs) extract event parameters  $\mathbf{P}_C$ :

$$\text{CRF}(\mathbf{E}_L) \rightarrow \mathbf{P}_C \quad (4)$$

<sup>2</sup><https://spacy.io/models>

Leveraging Graph Convolutional Networks (GCNs), we model complex event relationships and score each event's impact as:

$$\text{ImpactScore}(e_i) = \text{Centrality}(e_i) + \text{Magnitude}(e_i) \quad (5)$$

This scoring mechanism balances two crucial factors.  $\text{Centrality}(e_i)$  represents the event's importance within the network, reflecting its centrality or influence in the supply chain context. Meanwhile,  $\text{Magnitude}(e_i)$  quantifies the event's impact intensity, indicating its severity or significance.

Finally, we apply logical constraints and argument coreference to ensure robustness:

$$\text{LogicCoref}(\mathbf{P}_C) \rightarrow \mathbf{P}_F \quad (6)$$

producing a refined, logically consistent set of event parameters  $\mathbf{P}_F$ . More implementation details are in Appx. F.

**Event Matching & Instantiation.** We link extracted events with schema library to detect supply chain disruption patterns using a multi-dimensional approach of semantic and structural similarities. We align each extracted event  $E_{\text{ext}} \in \mathbf{E}_{\text{ext}}$  (extracted events) with each schema event  $E_{\text{schema}} \in \mathbf{E}_{\text{schema}}$  (schema events) using a composite similarity:

$$\begin{aligned} \text{Sim}(E_{\text{ext}}, E_{\text{schema}}) = & \alpha \cdot \text{SemSim}(E_{\text{ext}}, E_{\text{schema}}) \\ & + \beta \cdot \text{StrSim}(E_{\text{ext}}, E_{\text{schema}}) \end{aligned} \quad (7)$$

where  $\text{SemSim}$  captures contextual meaning using BERT embeddings, and  $\text{StrSim}$  assesses structural similarity. Specifically, semantic similarity measures contextual alignment using cosine similarity between BERT embeddings:

$$\text{SemSim}(E_{\text{ext}}, E_{\text{schema}}) = \frac{\mathbf{v}_{\text{ext}} \cdot \mathbf{v}_{\text{schema}}}{\|\mathbf{v}_{\text{ext}}\| \|\mathbf{v}_{\text{schema}}\|} \quad (8)$$

where  $\mathbf{v}_{\text{ext}}$  and  $\mathbf{v}_{\text{schema}}$  are BERT embeddings of extracted and schema events. Similarly, structural similarity evaluates parameter overlap using Jaccard similarity:

$$\text{StrSim}(E_{\text{ext}}, E_{\text{schema}}) = \frac{|\mathbf{P}_{\text{ext}} \cap \mathbf{P}_{\text{schema}}|}{|\mathbf{P}_{\text{ext}} \cup \mathbf{P}_{\text{schema}}|} \quad (9)$$

where  $\mathbf{P}_{\text{ext}}$  and  $\mathbf{P}_{\text{schema}}$  are the parameter sets for the extracted and schema events.

Following the calculation of semantic and structural similarities, we refine matching using heuristic rules from annotated datasets. Successful

matches lead to event instantiation, enriching the event representation with schema attributes:

$$\text{Instantiate}(E_{\text{matched}}, \mathbf{S}_{\text{schema}}) \rightarrow \mathbf{E}_{\text{inst}} \quad (10)$$

where  $E_{\text{matched}}$  represents the matched event,  $\mathbf{S}_{\text{schema}}$  yields the schema library, and  $\mathbf{E}_{\text{inst}}$  refers to the instantiated event with enriched attributes.

To ensure logical adherence to schema constraints, we perform consistency checks. These checks validate the instantiated events against the schema library, ensuring they conform to predefined logical and structural constraints:

$$\text{ConsistencyCheck}(\mathbf{E}_{\text{inst}}, \mathbf{S}_{\text{schema}}) \quad (11)$$

This step is crucial for maintaining the integrity of the schema and the reliability of the predictions.

Finally, we incorporate a continuous improvement process through manual review and feedback. Feedback from domain experts is used to update and refine the models, ensuring they adapt to new patterns and maintain high performance. The complete process is summarized in Algorithm 3. More implementation details are in Appx. G.

**Disruption Prediction.** Building on the extracted and matched events, we employ Graph Convolutional Networks (GCNs), logical constraints, and argument coreference resolution to predict supply chain disruptions. Note that the events are represented as nodes and interactions as edges using GCNs with the propagation rule:

$$\mathbf{H}^{(l+1)} = \sigma(\mathbf{A}\mathbf{H}^{(l)}\mathbf{W}^{(l)}) \quad (12)$$

where  $\mathbf{H}^{(l)}$  is the hidden state at layer  $l$ ,  $\mathbf{A}$  is the adjacency matrix,  $\mathbf{W}^{(l)}$  is the weight matrix, and  $\sigma$  is a non-linear activation function. We optimize using mean squared error loss with L2 regularization:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \|\mathbf{W}\|^2 \quad (13)$$

where  $y_i$  and  $\hat{y}_i$  are actual and predicted disruption scores, and  $\lambda$  is a regularization parameter. This approach balances prediction accuracy and model complexity, preventing overfitting.

To ensure consistency with domain knowledge, we apply logical constraints, refining initial predictions  $(\hat{y})$  to produce final predictions  $(\hat{y}')$  that adhere to known rules:

$$\begin{aligned} \hat{y}' &= \arg \min_{\hat{y}' \in \mathcal{Y}} \text{Constrain}(\hat{y}) \\ \text{subject to } \mathcal{C}(\hat{y}') &= \text{true} \end{aligned} \quad (14)$$

Table 1: Performance comparison of different LLMs on schema learning in stage 1.

	ChatGPT4o		Llama3-3b		Llama3-70b	
	Individual Schemas	Integrated Library	Individual Schemas	Integrated Library	Individual Schemas	Integrated Library
Precision	<b>0.637</b>	<b>0.184</b>	0.198	0.018	0.353	0.019
Recall	<b>0.695</b>	<b>0.336</b>	0.047	0.014	0.133	0.022
F-score	<b>0.652</b>	<b>0.238</b>	0.068	0.016	0.175	0.020

Table 2: Subjective evaluation by domain experts.

Model	Consistency	Accuracy	Completeness
GPT-4o	<b>4.5</b>	<b>4.3</b>	<b>4.6</b>
Llama3-3b	1.8	1.5	1.9
Llama3-70b	3.0	2.7	3.1

where  $\mathcal{C}$  represents constraint sets. For example, a constraint might ensure that a major supplier’s disruption increases risk for dependent manufacturers.

To further enhance the model’s contextual understanding, we incorporate argument coreference:

$$R_{ij} = \arg, \max_{E_i, E_j \in \mathcal{E}}; \text{Coref}(E_i, E_j) \quad (15)$$

subject to  $\text{Coref}(E_i, E_j) = \text{true}$

where  $(E_i, E_j)$  denotes each event pair and  $R_{ij}$  represents their relation. This AllenNLP-based model links entities across event mentions, recognizing when different descriptions refer to the same incident, thereby improving prediction accuracy and context comprehension. Algorithm 1 outlines our approach, combining GCN-based predictions, logical constraints, and argument coreference resolution. Detailed examples and implementation guidelines are provided in Appx. H.

#### Algorithm 1 Supply Chain Disruption Prediction

- 1: **Input:** Historical supply chain events  $\mathbf{E}$ , adjacency matrix  $\mathbf{A}$ , initial predictions  $\hat{y}$
- 2: **Output:** Refined predictions  $\hat{y}'$
- 3: **GCN-based Prediction**  $\triangleright$  Initial prediction using GCN
- 4: **for**  $l = 1$  to  $L$  **do**
- 5:      $\mathbf{H}^{(l+1)} = \sigma(\mathbf{AH}^{(l)}\mathbf{W}^{(l)})$   $\triangleright$  Refer to Eq. 12
- 6: **end for**
- 7:  $\hat{y} \leftarrow \mathbf{H}^{(L)}$
- 8: **Constrained Prediction**  $\triangleright$  Apply logical constraints
- 9: **for** each prediction  $\hat{y}_i$  **do**
- 10:     $\hat{y}'_i \leftarrow \text{Constrain}(\hat{y}_i)$
- 11:    such that  $\mathcal{C}(\hat{y}'_i) = \text{true}$   $\triangleright$  Refer to Eq. 14
- 12: **end for**
- 13: **Coreference Resolution**  $\triangleright$  Link related events
- 14: **for** each pair of events  $(E_i, E_j)$  **do**
- 15:     $R_{ij} \leftarrow \text{Coref}(E_i, E_j)$   $\triangleright$  Refer to Eq. 15
- 16:    **if**  $R_{ij}$  is coreferential **then**
- 17:       Link  $E_i$  and  $E_j$
- 18:    **end if**
- 19: **end for**
- 20: **Return:** Refined predictions  $\hat{y}$

## 4 Experiments

Our evaluation comprises two parts: (1) *Schema Learning Assessment* and (2) *Supply Chain Disruption Prediction*. We assess learned schemas against expert knowledge and evaluate our schema induction process’s effectiveness in predicting supply chain events. Detailed experimental setup and evaluation metrics are in Appx. I.

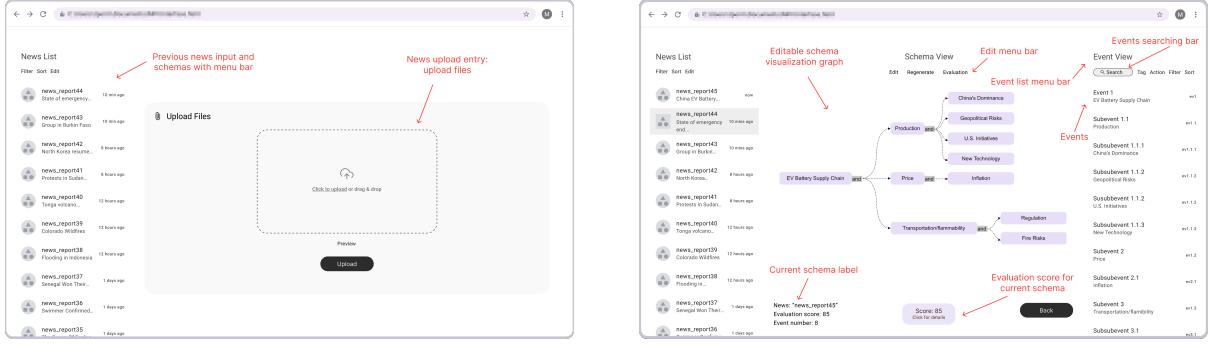
### 4.1 Schema Learning Performance

We evaluate GPT-4o, Llama3-3b, and Llama3-70b for schema learning, comparing individual schema extraction and integrated library generation. Tables 1 and 2 present quantitative metrics and subjective evaluations by domain experts. GPT-4o outperforms Llama models, achieving F-scores of 0.652 and 0.238 for individual schemas and integrated library generation, respectively. All models perform better in individual schema extraction than integrated library generation, indicating challenges in schema integration. Subjective assessments align with quantitative metrics, with GPT-4o scoring highest across all criteria (consistency: 4.5, accuracy: 4.3, completeness: 4.6). Individual schemas show strong consistency and completeness but slightly lower accuracy, suggesting a trade-off between comprehensive coverage and precise detail representation.

### 4.2 Disruption Detection Performance

**Event Extraction & Matching.** Table 3 presents quarterly results for 2022 and 2023 on event extraction and matching using a supply chain news dataset. Our system maintains consistent performance across quarters, with F-scores ranging from 0.671 to 0.700. This stability suggests robust generalization across different time periods and varying event types. The slight improvement in 2023 (average F-score 0.687 vs 0.683 in 2022) indicates potential refinement in our model’s ability to adapt to evolving supply chain dynamics.

**Disruption Detection.** Our advanced GCNs model, augmented with logical constraints and coreference resolution, was rigorously evaluated against



(a) User interface for inputting news reports.

(b) Visualization and editing of final prediction results.

Figure 4: User interface for online disruption analysis in stage 2, showing the process from news report input to the visualization and editing of prediction results. More examples are in Appx. K.

Table 3: Event extraction and matching in supply chain news dataset.

Year	2022				2023			
	Quarter	Q1 (Jan-Mar)	Q2 (Apr-Jun)	Q3 (Jul-Sep)	Q4 (Oct-Dec)	Q1 (Jan-Mar)	Q2 (Apr-Jun)	Q3 (Jul-Sep)
Precision	0.714	0.692	0.705	0.689	0.712	0.698	0.703	0.690
Recall	0.675	0.662	0.678	0.655	0.688	0.670	0.681	0.657
F-score	0.694	0.677	0.691	0.671	0.700	0.684	0.692	0.673

Table 4: Performance comparison of different models on disruption prediction.

Model	Precision	Recall	F-score
Our System (GCNs only)	0.701	0.670	0.685
Our System (GCNs + Logical Constraints)	0.724	0.691	0.707
Our System (GCNs + Logical Constraints + Coreference)	0.754	0.712	0.732

Table 5: Performance comparison of direct human interaction with LLMs on disruption prediction.

Model	Precision	Recall	F-score
GPT-4o	0.641	0.608	0.624
Llama3-3b	0.522	0.489	0.505
Llama3-70b	0.557	0.523	0.540
Our Method	<b>0.754</b>	<b>0.712</b>	<b>0.732</b>

ablation versions and LLM+prompt methods. Tables 4 and 5 present the comparative performance metrics. The full system achieved the highest F-score (0.732), significantly outperforming both ablation versions (GCNs+Logical Constraints: 0.707, GCNs only: 0.685) and LLM+prompt methods (GPT-4o: 0.624). However, the incremental improvement from the GCNs-only model to our full system (0.685 to 0.732) suggests that while the additional components significantly enhance performance, there remains substantial potential for further optimization in the future.

### 4.3 Qualitative Analysis & User Interface

Our qualitative analysis of SHIELD’s disruption predictions, focusing on real-world case studies (detailed in Appx. J), complements the quantitative findings and further illuminates the system’s practical utility. A particularly salient example emerged

in SHIELD’s accurate prediction of a semiconductor shortage resulting from geopolitical tensions, made three weeks prior to widespread reporting. This early insight enabled proactive adjustments to procurement strategies, thereby demonstrating the system’s considerable potential in mitigating complex supply chain risks. We have developed an interactive user interface (Fig. 4) for online disruption analysis. This interface allows users to upload news report texts (Fig. 4a), evaluate prediction scores, and edit visualization results for the final disruption analysis (Fig. 4b). More details can be found in Appx. K.

## 5 Conclusion

We present SHIELD, a two-stage framework that integrates Large Language Models (LLMs) with domain expertise, yielding promising results in EV battery supply chain analytics and risk assessment. While demonstrating particular strength in early disruption detection and event prediction for critical battery materials, *significant challenges remain in schema integration, real-time adaptability, and error reduction*. Future research will systematically address these limitations, enhance the system’s robustness, and explore broader applications across diverse industries and supply chain ecosystems.

## References

- David Ahn. 2006. The stages of event extraction. In *Proceedings of the Workshop on Annotating and Reasoning about Time and Events*, pages 1–8.
- Abeer Aljohani. 2023. Predictive analytics and machine learning for real-time supply chain risk mitigation and agility. *Sustainability*, 15(20):15088.
- John R Anderson, Paul J Kline, and Charles M Beasley Jr. 1979. A general learning theory and its application to schema abstraction1. In *Psychology of learning and motivation*, volume 13, pages 277–318. Elsevier.
- Anjali Awasthi, Kannan Govindan, and Stefan Gold. 2018. Multi-tier sustainable global supplier selection using a fuzzy ahp-vikor based approach. *International Journal of Production Economics*, 195:106–117.
- Hülya Behret, Başar Öztaş, and Cengiz Kahraman. 2012. A fuzzy inference system for supply chain risk management. In *Practical Applications of Intelligent Systems: Proceedings of the Sixth International Conference on Intelligent Systems and Knowledge Engineering, Shanghai, China, Dec 2011 (ISKE2011)*, pages 429–438. Springer.
- Maurício F Blos, Robson M Da Silva, and Paulo E Miyagi. 2015. Application of an agent-based supply chain to mitigate supply chain disruptions. *IFAC-PapersOnLine*, 48(3):640–645.
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.
- Shu Cai and Kevin Knight. 2013. Smatch: an evaluation metric for semantic feature structures. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 748–752.
- Alvaro Camarillo, José Ríos, and Klaus-Dieter Althoff. 2018. Knowledge-based multi-agent system for manufacturing problem solving process in production plants. *Journal of manufacturing systems*, 47:115–127.
- Real Carbonneau, Kevin Laframboise, and Rustam Vahidov. 2008. Application of machine learning techniques for supply chain demand forecasting. *European journal of operational research*, 184(3):1140–1154.
- Angel X Chang and Christopher D Manning. 2012. Sutime: A library for recognizing and normalizing time expressions. In *Lrec*, volume 12, pages 3735–3740.
- Yubo Chen, Liheng Xu, Kang Liu, Daojian Zeng, and Jun Zhao. 2015. Event extraction via dynamic multi-pooling convolutional neural networks. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 167–176.
- Rotem Dror, Haoyu Wang, and Dan Roth. 2022. Zero-shot on-the-fly event schema induction. *arXiv preprint arXiv:2210.06254*.
- Xinya Du, Zixuan Zhang, Sha Li, Pengfei Yu, Hongwei Wang, Tuan Lai, Xudong Lin, Ziqi Wang, Iris Liu, Ben Zhou, et al. 2022. Resin-11: Schema-guided event prediction for 11 newsworthy scenarios. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations*, pages 54–63.
- Selby H Evans. 1967. A brief statement of schema theory. *Psychonomic Science*, 8(2):87–88.
- Maryam Gallab, Hafida Bouloiz, Youssef Lamrani Alaoui, and Mohamed Tkouat. 2019. Risk assessment of maintenance activities using fuzzy logic. *Procedia computer science*, 148:226–235.
- A Deiva Ganesh and P Kalpana. 2022. Future of artificial intelligence and its influence on supply chain risk management—a systematic review. *Computers & Industrial Engineering*, 169:108206.
- Myles D Garvey, Steven Carnovale, and Sengun Yeniyurt. 2015. An analytical framework for supply network risk propagation: A bayesian network approach. *European Journal of Operational Research*, 243(2):618–627.
- Mihalis Giannakis and Michalis Louis. 2011. A multi-agent based framework for supply chain risk management. *Journal of Purchasing and Supply Management*, 17(1):23–31.
- Mihalis Giannakis and Michalis Louis. 2016. A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, 29(5):706–727.
- Jeevith Hegde and Børge Rokseth. 2020. Applications of machine learning methods for engineering risk assessment—a review. *Safety science*, 122:104492.
- Christian Hendriksen. 2023. Artificial intelligence for supply chain management: Disruptive innovation or innovative disruption? *Journal of Supply Chain Management*, 59(3):65–76.
- Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, and Clare R Voss. 2018. Zero-shot transfer learning for event extraction. In *56th Annual Meeting of the Association for Computational Linguistics, ACL 2018*, pages 2160–2170. Association for Computational Linguistics (ACL).
- Andrew K Lampinen, Ishita Dasgupta, Stephanie CY Chan, Kory Matthewson, Michael Henry Tessler, Antonia Creswell, James L McClelland, Jane X Wang, and Felix Hill. 2022. Can language models learn from explanations in context? *arXiv preprint arXiv:2204.02329*.
- Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han, and Clare Voss. 2021. The future is not one-dimensional: Complex

- event schema induction by graph modeling for event prediction. *arXiv preprint arXiv:2104.06344*.
- Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers, and Clare Voss. 2020. Connecting the dots: Event graph schema induction with path language modeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 684–695.
- Sha Li, Ruining Zhao, Manling Li, Heng Ji, Chris Callison-Burch, and Jiawei Han. 2023. Open-domain hierarchical event schema induction by incremental prompting and verification. *arXiv preprint arXiv:2307.01972*.
- Jian Liu, Yubo Chen, Kang Liu, Wei Bi, and Xiaojiang Liu. 2020. Event extraction as machine reading comprehension. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pages 1641–1651.
- Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Georgios Makridis, Philip Mavrepis, and Dimosthenis Kyriazis. 2023. A deep learning approach using natural language processing and time-series forecasting towards enhanced food safety. *Machine Learning*, 112(4):1287–1313.
- Ishani Mondal, Michelle Yuan, Aparna Garimella, Francis Ferraro, Andrew Blair-Stanek, Benjamin Van Durme, Jordan Boyd-Graber, et al. 2023. Interactiveie: Towards assessing the strength of human-ai collaboration in improving the performance of information extraction. *arXiv preprint arXiv:2305.14659*.
- Thien Huu Nguyen, Kyunghyun Cho, and Ralph Grishman. 2016. Joint event extraction via recurrent neural networks. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, pages 300–309.
- Sanjoy Kumar Paul. 2015. Supplier selection for managing supply risks in supply chain: a fuzzy approach. *The International Journal of Advanced Manufacturing Technology*, 79:657–664.
- Sanjoy Kumar Paul, Ruhul Sarker, and Daryl Essam. 2017. A quantitative model for disruption mitigation in a supply chain. *European Journal of Operational Research*, 257(3):881–895.
- Raúl Pino, Isabel Fernández, David de la Fuente, José Parreño, and Paolo Priore. 2010. Supply chain modelling using a multi-agent system. *Journal of Advances in Management Research*, 7(2):149–162.
- Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Hwee Tou Ng, Anders Björkelund, Olga Uryupina, Yuchen Zhang, and Zhi Zhong. 2013. Towards robust linguistic analysis using ontonotes. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pages 143–152.
- Jack W Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, et al. 2021. Scaling language models: Methods, analysis & insights from training gopher. *arXiv preprint arXiv:2112.11446*.
- Partha Pratim Ray. 2023. Leveraging deep learning and language models in revolutionizing water resource management, research, and policy making: A case for chatgpt. *ACS ES&T Water*, 3(8):1984–1986.
- Gonzalo A Ruz, Pablo A Henríquez, and Aldo Mascaró. 2020. Sentiment analysis of twitter data during critical events through bayesian networks classifiers. *Future Generation Computer Systems*, 106:92–104.
- Alexander Schrijver. 1998. *Theory of linear and integer programming*. John Wiley & Sons.
- Lei Sha, Feng Qian, Baobao Chang, and Zhifang Sui. 2018. Jointly extracting event triggers and arguments by dependency-bridge rnn and tensor-based argument interaction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.
- Xiaoming Shi, Siqiao Xue, Kangrui Wang, Fan Zhou, James Zhang, Jun Zhou, Chenhao Tan, and Hongyuan Mei. 2024. Language models can improve event prediction by few-shot abductive reasoning. *Advances in Neural Information Processing Systems*, 36.
- Nathalie Silva, Luís Miguel DF Ferreira, Cristóvão Silva, Vanessa Magalhães, and Pedro Neto. 2017. Improving supply chain visibility with artificial neural networks. *Procedia Manufacturing*, 11:2083–2090.
- Jannik Strötgen and Michael Gertz. 2013. Multilingual and cross-domain temporal tagging. *Language Resources and Evaluation*, 47:269–298.
- Naushad UzZaman, Hector Llorens, James Allen, Leon Derczynski, Marc Verhagen, and James Pustejovsky. 2012. Tempeval-3: Evaluating events, time expressions, and temporal relations. *arXiv preprint arXiv:1206.5333*.
- Hongwei Wang, Zixuan Zhang, Sha Li, Jiawei Han, Yizhou Sun, Hanghang Tong, Joseph P Olive, and Heng Ji. 2022a. Schema-guided event graph completion. *arXiv preprint arXiv:2206.02921*.
- Sijia Wang, Mo Yu, Shiyu Chang, Lichao Sun, and Lifu Huang. 2021. Query and extract: Refining event extraction as type-oriented binary decoding. *arXiv preprint arXiv:2110.07476*.
- Xiaozhi Wang, Ziqi Wang, Xu Han, Zhiyuan Liu, Juanzi Li, Peng Li, Maosong Sun, Jie Zhou, and Xiang Ren. 2019. Hmeae: Hierarchical modular event argument extraction. In *Proceedings of the 2019 Conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pages 5777–5783.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, and Denny Zhou. 2022b. Rationale-augmented ensembles in language models. *arXiv preprint arXiv:2207.00747*.

Jason Wei, Maarten Bosma, Vincent Y Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M Dai, and Quoc V Le. 2021. Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.

Diyi Yang, Sherry Tongshuang Wu, and Marti A Hearst. 2024. Human-ai interaction in the age of llms. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 5: Tutorial Abstracts)*, pages 34–38.

Tianyi Zhang, Isaac Tham, Zhaoyi Hou, Jiaxuan Ren, Liyang Zhou, Hainiu Xu, Li Zhang, Lara J Martin, Rotem Dror, Sha Li, et al. 2023. Human-in-the-loop schema induction. *arXiv preprint arXiv:2302.13048*.

Zixuan Zhang and Heng Ji. 2021. Abstract meaning representation guided graph encoding and decoding for joint information extraction. In *Proc. The 2021 Conference of the North American Chapter of the Association for Computational Linguistics-Human Language Technologies (NAACL-HLT2021)*.

## A Related Work

**Supply Chain Risk Management.** AI techniques have been increasingly applied to predict and mitigate supply chain risks (Ganesh and Kalpana, 2022). While agent-based approaches (Pino et al., 2010; Giannakis and Louis, 2011, 2016; Blos et al., 2015) enable inter-agent communication for forecasting, they often lack robust predictive capabilities and have limited parameter sets. Rule-based reasoning methods (Gallab et al., 2019; Behret et al., 2012; Paul, 2015; Paul et al., 2017; Awasthi et al., 2018; Camarillo et al., 2018) offer decision-making frameworks but provide minimal quantitative insights. To address these limitations, Machine Learning (DL) and Deep Learning (DL) techniques have been employed (Silva et al., 2017; Hegde and Rokseth, 2020; Garvey et al., 2015; Ruz et al., 2020; Aljohani, 2023; Carboneau et al., 2008), enhancing demand forecasting and disruption prediction (Hendriksen, 2023; Makridis et al., 2023). Recent studies have begun exploring the potential of large language models (LLMs) in supply chain management (Ray, 2023). However, most current works prioritize predictive performance over interpretability, hindering practitioners' ability to make informed decisions. Our approach addresses this gap by integrating LLMs for schema induction, extracting hierarchical knowledge-graph structures from academic resources to predict supply chain disruptions, thereby enhancing both predictive performance and interpretability.

**Schema Induction & Learning.** Building on foundational works (Anderson et al., 1979; Evans, 1967), recent advancements in language modeling have revolutionized schema induction. Large-scale language models (LLMs) (Brown et al., 2020; Rae et al., 2021) have demonstrated remarkable capabilities in learning and generating schemas with minimal supervision. Researchers have explored various strategies to enhance these models, including contextual explanations (Wei et al., 2021; Lampinen et al., 2022), rationale-augmented ensembles (Wang et al., 2022b), and incremental prompting (Li et al., 2023). Transformer-based approaches (Li et al., 2020, 2021) have proven particularly effective in managing schema generation for complex scenarios, representing schemas as graphs. Integrating human feedback (Mondal et al., 2023; Yang et al., 2024; Zhang et al., 2023) has been crucial in refining schema induction processes, addressing the limitations of automated methods. Our

approach leverages these advancements by employing an LLM-driven framework that integrates human feedback and expert knowledge into a human-in-loop system, thereby enhancing the practical accuracy and relevance of induced schemas.

**Event Extraction & Analysis.** Event extraction has evolved from manually crafted features (Ahn, 2006) to neural models, including recurrent networks (Nguyen et al., 2016; Sha et al., 2018), convolutional networks (Chen et al., 2015), graph networks (Zhang and Ji, 2021), and transformers (Liu et al., 2020). Recent research has focused on event argument extraction (Wang et al., 2019) and explored zero-shot learning (Huang et al., 2018) and weak supervision (Chen et al., 2015) to enhance performance. Our approach incorporates various event extraction techniques, utilizing fine-tuned RoBERTa models and graph convolutional networks (GCNs) to capture and analyze complex event relationships and their cascading impacts. This approach enables a deeper understanding of supply chain disruptions, distinguishing our system from traditional extraction techniques.

## B Dataset

### B.1 Schema Learning Dataset

Our research began by examining the current state of EV batteries, focusing on the predominant types in use, such as lithium iron phosphate and nickel lithium batteries. We analyzed the battery production process and identified key raw materials, including lithium, cobalt, nickel, and graphite. Subsequently, we investigated the primary sources and production volumes of these materials. Through an extensive review of literature and statistical data, we categorized significant supply chain events into eight groups, three of which have long-term impacts. Each category was further divided into sub-categories, and real-world events were identified to illustrate their impact on raw material supplies.

We also analyzed price trends for key raw materials over the past five years, using data from the London Metal Exchange (LME)<sup>3</sup>, to assess how news events influenced these prices. This research produced an initial scenario document listing the primary raw materials for EV batteries, their price trends, and an analysis of events causing supply chain issues and price fluctuations. Each category

<sup>3</sup><https://www.lme.com/en/>

included at least one real-world example to demonstrate its impact.

The initial document was then submitted for review by a supply chain expert. Based on the expert’s feedback, we refined the events affecting the EV battery supply chain into 11 main categories, three with long-term impacts, and subdivided them into 27 subcategories. Each subcategory was illustrated with 1-2 real-world events, and raw materials were further subdivided, such as different grades of nickel and types of lithium. Categories with minimal impact were removed, resulting in a comprehensive and refined scenario document.

Based on the scenario document, we identified the raw materials and events related to the EV battery supply chain and began collecting an academic document dataset. Our data sources included Wikipedia entries, supply chain-related papers, and industrial reports. After obtaining the raw data, we manually removed redundant information and noise, retaining only the paragraphs most relevant to the EV battery supply chain. Through meticulous organization, we compiled an academic document dataset consisting of 125 entries, distilled from 239 diverse sources: 200 academic papers, 22 industry reports, and 17 Wikipedia entries. This curated dataset provides a focused knowledge base essential for analyzing and understanding the complexities of the EV battery supply chain.

The resulting dataset encompasses a wealth of knowledge related to the EV battery supply chain, covering aspects such as raw material procurement, manufacturing processes, supply chain logistics, and market dynamics. Table 6 presents the event categories and example events. Events marked with \* indicate potential long-term impacts, highlighting the various types of disruptions and their implications for the supply chain. Fig. 5 and 6 illustrate the sources of the academic papers, Wikipedia entries, and industry reports used in compiling the dataset, demonstrating the breadth and diversity of our data sources. By synthesizing this information, we aim to provide a robust foundation for understanding the complexities and challenges associated with the EV battery supply chain.

## B.2 Supply Chain News Dataset

To comprehensively test our system, we constructed an EV Supply Chain News Dataset covering the period from January 2022 to December

2023. We initially developed a Python crawler using the requests and BeautifulSoup packages to scrape news titles and summaries from multiple websites, such as Google News<sup>4</sup> and Infoplease<sup>5</sup>. This resulted in a collection of 643 records. To filter out news unrelated to the supply chain, we designed a prompt leveraging GPT-4o’s language capabilities. Using the summaries from the list, GPT-4o helped categorize events into various types, such as natural disasters, wars, trade policy, and political issues, tagging the relevant countries and regions.

Subsequently, we employed large language models (LLMs) to evaluate the relevance of each news event to the EV battery supply chain based on the following criteria, each worth 25 points:

1. Whether natural disasters or humanitarian crises occurred in raw material production areas, such as China, Australia, Indonesia, Congo, Chile, Canada, or in EV production countries, such as China, Japan, South Korea, and the United States.
2. Whether the event could affect trade relations in the aforementioned countries, including trade issues, sanctions, or wars.
3. Whether the event could potentially disrupt international shipping routes due to conflicts or natural disasters near these routes.
4. Whether the event is directly related to international trade.

Events scoring below 25 points were initially eliminated, followed by a manual review of the remaining events, resulting in a refined list of 247 supply chain-related news events. The text data was sourced from reputable media outlets, including Reuters<sup>6</sup>, BBC<sup>7</sup>, and CNN<sup>8</sup>. Additionally, to gather contemporaneous supply chain status information, we scraped company news and analysis reports from EV battery-related companies like Ford, Volkswagen, and CATL, as well as supply chain-related websites, totaling 118 reports. The raw data, including titles, publication dates, and content, was organized chronologically.

The raw data contained invalid information and

<sup>4</sup><https://news.google.com/>

<sup>5</sup><https://www.infoplease.com/>

<sup>6</sup><https://www.reuters.com/>

<sup>7</sup><https://www.bbc.com/>

<sup>8</sup><https://edition.cnn.com/>

Table 6: Event Categories and Example Events. Events marked with \* indicate potential long-term impacts.

Event Category	Subcategory	Example
Acquisition and Investment*	Investment from U.S. or Ally Investment from Other Country	U.S. invests in EV battery industry in Canada China invests in cobalt mines in DRC
Changes in Supply and Demand	Demand Change Supply Change	Demand for ore from the Philippines increases Tight supplies of nickel ore in Indonesia
Enterprise Issue	Production Halt or Reduction	Katanga halts cobalt mining
	Enterprise Crisis	Katanga faces an equity crisis
	Production Plan Adjustment	Kellyton Graphite increases production by 15%
Economic Environment	Macroeconomy Competition and Market Structure	The U.S. and EU face continued inflation Competition from China's low-priced EVs
EV Battery Technology Progress*	Product Upgrading Production Technology Progress	CATL releases Kirin battery Development of graphene batteries
Humanitarian and Ethical Crisis	Forced Labor	Forced labor in production
	Use of Child Labor	Child labor in cobalt mining in DRC
	Human Rights Issue	Large numbers of refugees enter Europe
Natural Disaster	Production Affected by Disaster Transportation Affected by Disaster	Australia floods affect lithium mining Tsunami destroys ports, disrupts shipping
Political Issue*	Regional Tension	Tensions between North and South Korea
	Changes in International Relations	China's relations with the West deteriorate
	Industry Nationalization	Nationalization of the lithium industry in Chile
	Government Intervention	Europe promotes EVs for environmental reasons
Sign a Supply Agreement	Sign a Supply Agreement	PE signs EV battery supply agreement with Tesla
Trade Policy	Export Controls	China restricts graphite exports
	Tax and Duties	China's tax rebates to EV companies
	Trade Barriers	US tariffs on Chinese EV batteries
War and Conflict	Internal Disorder or Rebellion	Civil unrest in Yemen
	War Between Nations	Russo-Ukrainian War
	Geopolitical Crisis	Houthi rebels attack merchant ship

advertisements, which were cleaned using regular expressions to remove most invalid information. We deployed Llama3-8b to filter out embedded advertisements, ensuring the dataset's purity and accuracy. After cleaning, irrelevant content was reduced by 15%, and all data was systematically stored in a database, resulting in a refined meta dataset of 365 news documents. The metadata contains approximately 152,000 words and 3,000 paragraphs. To validate our system's ability to detect connections between events, we randomly merged international news with contemporaneous corporate stories from the same quarter, creating 354 fused documents for a more diverse and challenging dataset. The final fusion dataset contains approximately 660,000 words and 12,000 paragraphs.

Upon completing dataset collection, we conducted preliminary statistics and analysis on the news dataset. Table 8 shows the number of event types included in each quarter in the news dataset, providing a comprehensive overview of the various events tracked over time. Table 9 details the number of words and paragraphs in the dataset, highlighting

the extensive scope of the collected data. Fig. 7 presents the categories and examples of news articles in the dataset, while Fig. 8 shows the distribution of sources in the news dataset, emphasizing the dataset's diversity and comprehensiveness. The dataset covers global events that could impact the supply chain, such as natural disasters, trade issues, wars, enterprise issues, etc.

## C Hierarchical Structure Extraction

We utilize large language models (LLMs) to extract hierarchical structures ( $\mathbf{H}$ ) that capture main events ( $\mathbf{E}$ ) and sub-events ( $\mathbf{E}_{\text{sub}}$ ) based on our prompt, as illustrated in Fig. 9.

In a hierarchical structure ( $\mathbf{H}$ ):

- An *event* ( $\mathbf{E}$ ) refers to anything that happens related to the EV battery supply chain. There can be multiple events  $\langle \mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n \rangle$  in one hierarchical structure  $\mathbf{H}$ .
- An *event\_id* is a unique identifier code assigned for each specific event.

• Academic Paper

- Electric vehicle battery supply chain and critical materials: a brief survey of state of the art
- Challenges and recent developments in supply and value chains of electric vehicle batteries: A sustainability perspective
- Cost-effective supply chain for electric vehicle battery remanufacturing
- The supply chain for electric vehicle batteries
- Critical issues in the supply chain of lithium for electric vehicle batteries
- Building a Robust and Resilient U.S. Lithium Battery Supply Chain
- At the mining or extraction stage, major risks include the location of the deposit, cost, geopolitical environment, and mining regulations
- A SWOT Analysis of the UK EV Battery Supply Chain
- Lithium-ion battery supply chain: enabling national electric vehicle and renewables targets
- Sustainable electric vehicle batteries for a sustainable world: perspectives on battery cathodes, environment, supply chain, manufacturing, life cycle, and policy
- Developing pricing strategy to optimise total profits in an electric vehicle battery closed loop supply chain
- Optimising quantity of manufacturing and remanufacturing in an electric vehicle battery closed-loop supply chain
- Life-cycle implications and supply chain logistics of electric vehicle battery recycling in California
- Mining challenges for sustainable supply chain of electric vehicle batteries using a hybrid approach of Delphi and Best-Worst Method
- Determining requirements and challenges for a sustainable and circular electric vehicle battery supply chain: A mixed-methods approach
- Graphic resources, and their potential to support battery supply chains, in Africa
- Lithium-ion battery supply chain considerations: analysis of potential bottlenecks in critical metals
- The case for recycling: Overview and challenges in the material supply chain for automotive li-ion batteries
- Global Value Chains: Graphite in Lithium-Ion Batteries for Electric Vehicles
- Traceability methods for cobalt, lithium, and graphite production in battery supply chains
- The global cycle of Graphite - A dynamic Material Flow Analysis (2020-2050) of the natural, synthetic and recycled graphite value chains to understand the supply of LIB anodes
- Natural graphite demand and supply—Implications for electric vehicle battery requirements
- Material System Analysis of five battery-related raw materials: Cobalt, Lithium, Manganese, Natural Graphite, Nickel
- Life cycle assessment of natural graphite production for lithium-ion battery anodes based on industrial primary data
- Sustainability challenges throughout the electric vehicle battery value chain
- Electric vehicle battery chemistry affects supply chain vulnerability
- Global competition in the lithium-ion battery supply chain: a novel perspective for criticality analysis
- Supply risks of lithium-ion battery materials: An entire supply chain estimation
- Vulnerable links in the lithium-ion battery supply chain
- The global battery arms race: lithium-ion battery gigafactories and their supply chain
- Towards the lithium-ion battery production network: Thinking beyond mineral supply chains
- Challenges and opportunities in lithium-ion battery supply
- Identifying supply risks by mapping the cobalt supply chain
- The Cobalt Supply Chain Environmental Life Cycle Impacts of Lithium-Ion Battery Energy Storage Systems
- Environmental sustainability and supply resilience of cobalt
- Perspectives on cobalt supply through 2030 in the face of changing demand
- An integrated supply chain analysis for cobalt and rare earth elements under global electrification and constrained resources
- Sources of uncertainty in the closed-loop supply chain of lithium-ion batteries for electric vehicles
- Battery technology and recycling alone will not save the electric mobility transition from future cobalt shortages
- Global value chains: cobalt in lithium-ion batteries for electric vehicles
- Global electrification of vehicles and intertwined material supply chains of cobalt, copper and nickel
- The cobalt supply chain and life cycle assessment
- Towards the lithium-ion battery production network: Thinking beyond mineral supply chains
- Vertically Integrated Supply Chain of Batteries, Electric Vehicles, and Charging Infrastructure: A Review of Three Milestone Projects From Theory of Constraints Perspective
- The behavioural evolution of the smart electric vehicle battery reverse supply chain under government supervision
- Managing resource dependencies in electric vehicle supply chains: a multi-tier case study
- Application of sustainable supply chain finance in end-of-life electric vehicle battery management: a literature review
- Structural characteristics and disruption ripple effect in a meso-level electric vehicle Lithium-ion battery supply chain network
- Battery global value chain and its technological challenges for electric vehicle mobility
- Building a competitive advantage for Indonesia in the development of the regional EV battery chain
- Recycling mode selection and carbon emission reduction decisions for a multi-channel closed-loop supply chain of electric vehicle power battery under cap-and-trade policy
- A sustainable circular supply chain network design model for electric vehicle battery production using Internet of things and big data
- The CO<sub>2</sub> Impact of the 2020s Battery Quality Lithium Hydroxide Supply Chain
- Electrolyte and supply chain management: A bibliometric and systematic review
- Long-term Indonesia's Nickel Supply Chain Strategy for Lithium-Ion Battery as Energy Storage System
- Battery Nickel Bottlenecks
- Life-cycle analysis, by global region, of automotive lithium-ion nickel manganese cobalt batteries of varying nickel content
- Battery minerals from Finland: Improving the supply chain for the EU battery industry using a metallurgical approach
- Exploring recycling options in battery supply chains—a life cycle sustainability assessment
- Assessment of social sustainability hotspots in the supply chain of lithium-ion batteries
- Conflict minerals and battery materials supply chains: A mapping review of responsible sourcing initiatives
- Analysis of nickel sulphate datasets used in lithium-ion batteries
- Design of battery supply chains under consideration of environmental and socio-economic criteria
- Analysis of international nickel flow based on the industrial chain
- Value recovery from spent lithium-ion batteries: A review on technologies, environmental impacts, economics, and supply chain
- Dynamic evolution of the zinc-nickel battery industry and evidence from China
- A perspective on the sustainability of cathode materials used in lithium-ion batteries
- An improved resource midpoint characterization method for supply risk of resources: integrated assessment of Li-ion batteries
- Deep-sea nodules versus land ores: A comparative systems analysis of mining and processing wastes for battery-metal supply chains
- Industrial policy, trade, and clean energy supply chains
- The electric vehicle revolution: Critical material supply chains, trade and development
- Industrial policy for electric vehicle supply chains and the US-EU fight over the Inflation Reduction Act
- Strategic Battery Partnership: Reducing Foreign Dependence in the Electric Vehicle Supply Chain
- Electric vehicle battery secondary use under governments' subsidy: A closed-loop supply chain perspective
- Blockchain for battery supply chain monitoring and battery trading
- Trade structure and risk transmission in the international automotive Li-ion batteries trade
- The EV Revolution: Critical Material Supply Chains, Trade, and Development
- An Overview of the Lithium Supply Chain
- Comparison of lithium-ion battery supply chains—a life cycle sustainability assessment
- The ev transition: Key market and supply chain enablers
- The Lithium Supply Crunch Doesn't Have to Stall Electric Cars
- Hydrometallurgical Routes to Close the Loop of Electric Vehicle (EV) Lithium-Ion Batteries (LIBs) Value Chain: A Review
- Global warming potential of lithium-ion battery cell production: Determining influential primary and secondary raw material supply routes
- Lithium mining: How new production technologies could fuel the global EV revolution
- The cobalt and lithium global supply chains: status, risks and recommendations
- Sustainable value chain of retired lithium-ion batteries for electric vehicles
- Assessing the potential of quebec lithium industry: Mineral reserves, lithium-ion batteries production and greenhouse gas emissions
- Current and Future Global Lithium Production Till 2025
- Lithium-Ion Batteries Recycling Trends and Pathways: A Comparison
- Alternative battery chemistries and diversifying clean energy supply chains
- Lithium and cobalt
- A Study on the Cradle-to-Gate Environmental Impacts of Automotive Lithium-ion Batteries
- Determining requirements and challenges for a sustainable and circular electric vehicle battery supply chain: A mixed-methods approach
- Status and role in rechargeable lithium battery supply chain: importance of quantitative failure analysis
- Critical Factors to Consider in Purchasing for a Sustainable Inbound Supply Chain: A Perspective on Large Scale Lithium-ion Battery Manufacturing
- Assessing battery supply chain networks for low impact vehicles
- A comparative assessment of value chain criticality of lithium-ion battery cells
- Conflict minerals and battery materials supply chains: A mapping review of responsible sourcing initiatives
- Battery Critical Materials Supply Chain Challenges and Opportunities: Results of the 2020 Request for Information (RFI) and Workshop
- Key Strategic Issues in Supply Chain Domain Pertaining to Battery Industry.
- Identifying trends in battery technologies with regard to electric mobility: evidence from patenting activities along and across the battery value chain
- Electric Vehicle Battery Supply Chain and Critical Materials: A Brief Survey of State of the Art
- Estimating the environmental impacts of global lithium-ion battery supply chain: A temporal, geographical, and technological perspective
- A SWOT Analysis of the UK EV Battery Supply Chain
- The Electric Vehicle Supply Chain Ecosystem: Changing Roles of Automotive Suppliers
- Electric Vehicle Battery Supply Chain and Critical Materials: A Brief Survey of State of the Art
- Estimating the environmental impacts of global lithium-ion battery supply chain: A temporal, geographical, and technological perspective
- A SWOT Analysis of the UK EV Battery Supply Chain
- The Electric Vehicle Supply Chain Ecosystem: Changing Roles of Automotive Suppliers
- Building a North American electric vehicle supply chain
- Friend-shoring battery supply chains
- Estimating the environmental impacts of global lithium-ion battery supply chain: A temporal, geographical, and technological perspective
- Global Supply Chains of EV Batteries
- Building a Sustainable Electric Vehicle Battery Supply Chain
- INVESTIGATING THE U.S. BATTERY SUPPLY CHAIN AND ITS IMPACT ON ELECTRIC VEHICLE COSTS THROUGH 2032
- GLOBAL STATE OF Sustainable ELECTRIC VEHICLE BATTERIES
- Collection and recycling decisions for electric vehicle end-of-life power batteries in the context of carbon emissions reduction
- The Paradox of Green Growth: Challenges and Opportunities in Decarbonizing the Lithium-Ion Supply Chain
- Implications of the Electric Vehicle Manufacturers' Decision to Mass Adopt Lithium-Iron Phosphate Batteries
- Reducing new mining for electric vehicle battery metals: responsible sourcing through demand reduction strategies and recycling
- Automated assembly of Li-ion vehicle batteries: A feasibility study
- Field Study and Multimethod Analysis of an EV Battery System Disassembly
- COVID-19 disrupts battery materials and manufacture supply chains, but outlook remains strong
- An overview of global power lithium-ion batteries and associated critical metal recycling
- Does China's new energy vehicles supply chain stock market have risk spillovers? Evidence from raw material price effect on lithium batteries
- Deep-sea nodules versus land ores: A comparative systems analysis of mining and processing wastes for battery-metal supply chains
- A game theoretic approach for analyzing electric and gasoline-based vehicles' competition in a supply chain under government sustainability strategies: A case study of South Korea
- China's lithium supply chain: Security dynamics and policy countermeasures
- Exploring recycling options in battery supply chains – a life cycle sustainability assessment
- Lithium-Ion Battery Recycling in the Circular Economy: A Review
- Implications of circular production and consumption of electric vehicle batteries on resource sustainability: A system dynamics perspective
- On the influence of second use, future battery technologies, and battery lifetime on the maximum recycled content of future electric vehicle batteries in Europe
- Assessing battery supply chain networks for low impact vehicles
- Mapping a business opportunity in electric vehicle battery value chain: A multi-stakeholder framework to create a win-win-win situation
- An applied analysis of the recyclability of electric vehicle battery packs
- A Novel Prediction Process of the Remaining Useful Life of Electric Vehicle Battery Using Real-World Data
- Capturing the battery value-chain opportunity
- Digital Twin-Driven Framework for EV Batteries in Automobile Manufacturing
- Electric vehicle battery state changes and reverse logistics considerations
- Optimising the geospatial configuration of a future lithium ion battery recycling industry in the transition to electric vehicles and a circular economy<sup>5</sup>
- Can Cobalt Be Eliminated from Lithium-Ion Batteries?
- Sizing and Locating Planning of EV Centralized-Battery-Charging Station Considering Battery Logistics System
- Taming the Hydrax: Future of Lithium-Ion Supply Chain in an Era of Unprecedented Volatility
- Battery recycling and recycling alone will not ease the strain of the transition from future cobalt shortages
- Resilience assessment of the lithium supply chain in China under impact of new energy vehicles and supply interruption
- Optimal policy for the recycling of electric vehicle retired power batteries
- Data requirements and availabilities for a digital battery passport – A value chain actor perspective
- Trade structure and risk transmission in the international automotive Li-ion batteries trade
- The Supply Chain Diversification and India-South Korea Cooperation in a Contested East Asia in the Post-COVID-19 Era
- Toward Sustainable Reuse of Retired Lithium-Ion Batteries from Electric Vehicles
- Steering extended producer responsibility for electric vehicle batteries
- Traceability Management Strategy of the EV Power Battery Based on the Blockchain
- Electric vehicle lithium-ion battery recycled content standards for the US – targets, costs, and environmental impacts
- What is the contribution of different business processes to material circularity at company-level? A case study for electric vehicle batteries
- To shred or not to shred: A comparative techno-economic assessment of lithium ion battery hydrometallurgical recycling retaining value and improving its value in LIB supply chains
- Value recovery from spent lithium-ion batteries: A review on technologies, environmental impacts, economics, and supply chain
- Optimal choice of power battery joint recycling strategy for electric vehicle manufacturers under a deposit-refund system
- Collection mode choice of spent electric vehicle batteries: considering collection competition and third-party economies of scale
- Electric vehicle battery capacity allocation and recycling with downstream competition
- Tackling EV Battery Chemistry in View of Raw Material Supply Shortfalls
- Which is better? Business models of partial and cross ownership in an NEV supply chain
- Minerals availability and supply chain considerations for the use of Pd and Pdium in grid-scale redox flow batteries
- Circularity of Lithium-Ion Batteries in Electric Vehicles
- Assessment of end-of-life electric vehicle batteries in China: Future scenarios and economic benefits
- The Resilience of the Renewable Energy Electromobility Supply Chain: Review and Trends
- Industrial Policy, Trade, and Clean Energy Supply Chains
- Concurrent design of product and supply chain architectures for modularity and flexibility: process, methods, and application
- A Sustainable Framework for the Second-life Battery Ecosystem based on blockchain
- Comparative evaluation and policy analysis for recycling retired EV batteries with different collection modes
- HOW TECHNOLOGY, RECYCLING, AND POLICY CAN MITIGATE SUPPLY RISKS TO THE LONG-TERM TRANSITION TO ZERO-EMISSION VEHICLES
- Decarbonizing the automotive sector: a primary raw material perspective on targets and timescales
- The Emerging Electric Vehicle and Battery Industry in Indonesia: Actions around the Nickel Ore Export Ban and a SWOT Analysis
- The Emerging Electric Vehicle and Battery Industry in Indonesia: Actions around the Nickel Ore Export Ban and a SWOT Analysis
- Reverse Logistics Network Design of Electric Vehicle Batteries considering Recall Risk
- Battery capacity needed to power electric vehicles in India from 2020 to 2035
- The future of the automotive sector: Emerging battery value chains in Europe
- Radical innovations as supply chain disruptions? A paradox between change and stability
- End of Electric Vehicle Batteries: Reuse vs. Recycle
- Securing Decarbonized Road Transport – a Comparison of How EV Deployment Has Become a Critical Dimension of Battery Security Strategies for China, the EU, and the US
- A Review on Battery Market Trends, Second-Life Reuse, and Recycling
- Life cycle impact assessment of electric vehicle batteries in Europe
- Intelligent disassembly of electric-vehicle batteries: a forward-looking overview
- Research on decision optimization of new energy vehicle supply chain considering demand disruptions under dual credit policy
- Transition to electric vehicles in China: Implications for private motorization rate and battery market
- Mirroring in production? Early evidence from the scale-up of Battery Electric Vehicles (BEVs)
- Life-Cycle Assessment Considerations for Batteries and Battery Materials
- Operation Management of Multiregion Battery Swapping-Charging Networks for Electrified Public Transportation Systems
- Spatial modeling of a second-use strategy for electric vehicle batteries to improve disaster resilience and circular economy
- On the sustainability of lithium ion battery industry – A review and perspective
- Comparison of Electric Vehicle Lithium-Ion Battery Recycling Allocation Methods
- Manufacturing value chain for battery electric vehicles in Pakistan: An assessment of capabilities and transition pathways
- The End of Globalized Production? Supply-Chain Resilience, Technological Sovereignty, and Enduring Global Interdependences in the Post-Pandemic Era
- Environmental feasibility of secondary use of electric vehicle lithium-ion batteries in communication base stations
- Supply chain risks of critical metals: Sources, propagation, and responses
- Decentralized Planning of Lithium-Ion Battery Production and Recycling
- Improvements in electric vehicle battery technology influence vehicle lightweighting and material substitution decisions
- Rethinking Chinese supply resilience of critical metals in lithium-ion batteries
- Optimal pricing strategy in the closed-loop supply chain using game theory under government subsidy scenario: A case study
- Predictive model for energy consumption of battery electric vehicle with consideration of self-uncertainty route factors
- Perspectives on Cobalt Supply through 2030 in the Face of Changing Demand
- McKinsey Electric Vehicle Index: Europe cushions a global plunge in EV sales
- Lithium in International Law: Trade, Investment, and the Pursuit of Supply Chain Justice

Figure 5: Sources of academic papers.

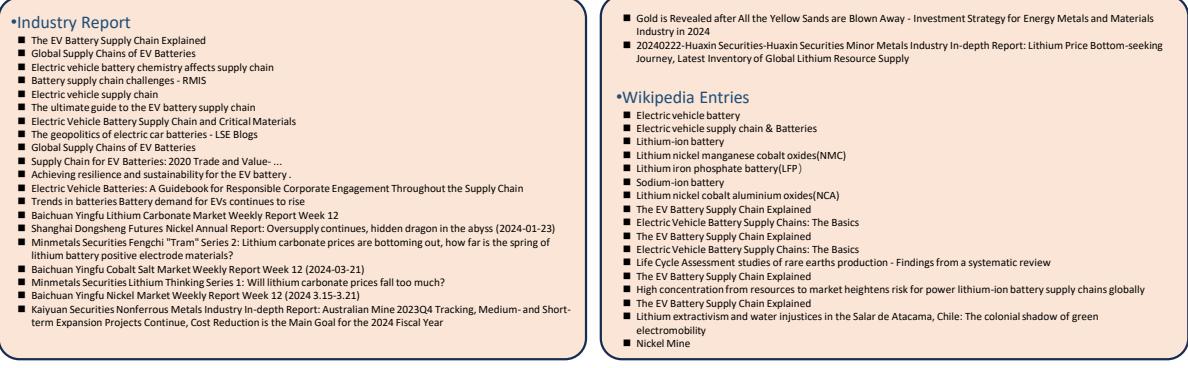


Figure 6: Sources of Wikipedia entries and industry reports.

Table 7: The number of event types included in each quarter in the news dataset.

Event Type	2022Q1	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2	2023Q3	2023Q4
Acquisition and Investment	0	5	4	0	2	2	2	1
Changes in Supply and Demand	6	5	4	3	3	1	2	3
Enterprise Issue	3	1	1	3	0	1	0	0
Economic Environment	3	3	5	6	4	2	6	1
Humanitarian and Ethical Crisis	1	5	3	1	2	0	2	3
Natural Disaster	6	8	6	5	6	6	6	6
Political Issue	3	14	18	16	14	10	15	10
EV Battery Technology Progress	1	1	1	2	4	3	2	0
Sign a Supply Agreement	0	3	1	1	2	2	1	4
Trade Policy	2	6	5	2	5	1	2	9
War and Conflict	3	6	12	10	5	7	7	7

Table 8: The number of times each country is mentioned in the news dataset in each quarter.

Country	2022Q1	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2	2023Q3	2023Q4
USA	6	9	18	15	17	12	9	10
China	1	10	12	6	12	4	4	9
EU	3	11	6	4	8	4	8	2
Japan	0	1	0	1	1	2	2	1
Russia	4	10	4	14	6	6	4	3
Other	20	42	38	33	31	23	35	40

Table 9: Statistics of the number of words and paragraphs in the dataset.

	Total Paragraphs	Total Words
Meta Data	3,022	152,489
Fusion Data	12,070	660,054

- A *description* provides a detailed 2-3 sentence textual explanation of the event.
- Participants include all sub-events ( $E_{\text{sub}}$ ) related to the main event, and a *subsubevent* ( $E_{\text{subsub}}$ ) can be used if an event is part of a sub-event within the hierarchy.

The suffix *P0.5* indicates the importance of a sub-event to its parent event. The *Gate* specifies the relationship between the main event and its sub-events:

- Use 'and' if no sub-events can be missing.
- Use 'or' if some sub-events can be missing.
- Use 'xor' if sub-events cannot exist simultaneously.

*Relations* describe the connections between events. For example, if *ev1.2* is caused by (happens after) *ev1.1*, it is expressed as '*ev1.1>ev1.2*'.

Our prompt includes demonstration and chain of thought (CoT) techniques:

- We manually annotated the hierarchical structure for one text in the schema learning dataset to use as an example in the prompt.
- We provided a step-by-step CoT, showing how  $E$  and  $E_{\text{sub}}$  in  $H$  were extracted from specific sentences in the schema learning dataset.

- **Acquisition and Investment**
  - Exclusive: Canada to invest C\$2 billion on mineral strategy for EV battery supply chain
  - China's EV battery materials industry set for \$11bn capacity buildup
  - Why carmakers are pouring billions into new electric vehicle battery factories
  - Private equity in talks with UK's BMI for EV battery exposure
  - CATL and Indonesia jointly build a nearly \$6 billion power battery industry chain project
  - Durham battery storage company raises \$100 million
  - Volkswagen announces \$20 billion effort to build its own EV batteries
  - Panasonic to open \$4B EV battery plant in Kansas
  - CATL announces construction of second European factory in Hungary
  - Tesla co-founder's startup gets \$2 billion to boost EV battery production
  - CATL and HGP establish partnership to jointly promote 5GWh battery energy storage application
  - GM and POSCO Future M Investing \$1bn in North American EV Battery Supply Chain
  - Companies invest in EV battery factories in Europe
  - Ecopromb, SK, Ford investing in Québec; building cathode plant to solidify EV supply chain in NA
  - Redwood Materials raises \$1B to expand circular battery supply chain in US
  - Nissan leads \$2.5 billion investment to build two more EVs in UK
  
- **Trade Policy**
  - As the US struggles to "green" supply chains, new EU battery regulation offers lessons
  - The New Climate Bill Demands All-American EV Batteries
  - New US Climate Bill Seeks to Onshore Electric Vehicle Supply Chain
  - The Inflation Reduction Act places a big bet on alternative mineral supply chains
  - The New Climate Bill Demands All-American EV Batteries
  - U.S. Push to Secure EV Battery Supply Chains and the Role of China
  - U.S. strikes at China with EV battery deal
  - EU Could End Reliance On Chinese Battery Supply Chain By 2030 Says T&E
  - Ford-CATL Partnership Illustrates the Challenge of Decoupling EV Supply Chains
  - Global EV battery supply chain puzzles over China graphite curbs
  - US-Canada Critical Mineral, EV Battery, and Semiconductor Cross-Border Supply Chain Issue
  - New US rules on Chinese batteries could push up price of electric cars
  - China restricts exports of graphite as it escalates a global tech war
  - China says Biden plan to shut it out of US battery supply chain violates WTO rules
  - US to limit Chinese firms, battery parts from winning EV tax credits
  - Senator asks Treasury to bar Chinese battery firms, minerals from US EV tax credits
  - UK Issues New Round of Targeted Sanctions Against Russia
  - Further Sanctions Against Russia Being Discussed by EU
  - Gas Supplies To Poland And Bulgaria To Be Cut Off By Russia
  - Russian Billionaire Shields Assets From European Sanctions
  - Singapore's National Dish Affected By The Malaysian Export Ban
  - Gulf States Sanction And Boycott India After Unwanted Remarks
  - Russia's Economy Will Be Hit By Further Sanctions
  - China Sanctions Pelosi, Halts Cooperation With The United States Over Pelosi's Taiwan Trip
  - China Vows To Take Countermeasures As The United States Approves \$1.1BN Arms Sales To Taiwan
  - Grain Export Deal Between Ukraine And Russia Brokered By United Nations Suspended By Russia
  - Grain Deal Extended By Russia And Ukraine Amid Disagreement
  - G7 Request For Black Sea Grain Deal To Be Extended
  - Russia Confirms It Will Not Renew Grain Deal With Ukraine
  - Ukraine Welcomes The Arrival Of First Grain Ships Using New Route
  - China Promise To Deepen Trade Ties With Vietnam
  - Eight North Korean Sanctioned By South Korea Over Arms Trade
  
- **EV Enterprise Related**
  - VW and Bosch to upscale EV battery output in Europe
  - Governor Ivey Joins Dura Automotive to Celebrate Grand Opening of High-Tech Factory in Muscle Shoals for EV Battery Enclosures
  - Auto Giants Race to Build U.S. EV Battery Assembly Plants
  - CATL's German factory obtains battery cell production license
  - Automakers race to build EV battery supply chains in North America
  - General Motors Fortifies EV Battery Supply-Chain Links
  - GM's North American battery supply chain is key to EV profits
  - CATL's German factory successfully achieves battery cell production
  - The South is building the most vibrant EV and battery hub in the US
  
- **Production Technology Progress**
  - The Transportation Transformation: Battery Research Today and Tomorrow
  - CATL releases Kirin battery with global highest integration
  - Ford releases new battery capacity plan, raw materials details to scale EVs
  - How the US plans to transform its lithium supply chain
  - Local, clean and circular supply chains: Panasonic advances EV battery tech
  - How Lithium Batteries Can Power the US Economy
  - Ford taps Michigan for new LFP battery plant; new battery chemistry offers customers value, durability, fast charging, creates 2,500 more new American jobs
  - Electric Vehicle Battery Manufacturing Capacity in North America in 2030 is Projected to be Nearly 20 Times Greater than in 2021
  - Ascend Elements Opens North America's Largest Electric Vehicle Battery Recycling Facility in Georgia
  - Study unveils policy insights for reshoring EV battery production
  - Batteries: EVs to use silicon, solid state for next-generation batteries
  - BMW powers Spartanburg with 'local' battery supply chain
  - New EV Battery Materials Will Beget New Dilemmas
  - Panasonic needs four more EV battery plants, executive says
  
- **Economic Environment**
  - Indonesia's Battery Industrial Strategy
  - DOE makes \$3.1B available for battery manufacturing incentives
  - Developing a resilient Canadian battery supply chain
  - Battery Policies and Incentives Database Contributes to U.S. Efforts To Build a Secure Electric Vehicle Battery Supply Chain
  - US increases production to catch China in global battery race
  - The CHIPS Act Is Essential. So Is a Resilient EV Battery Supply Chain
  - EV tax credits could stall out on lack of US battery supply
  - Electric Vehicle Battery Production May Lead To Coal Country's Return
  - DOE taps 20 companies to receive \$2.8B for battery manufacturing, minerals processing build-out
  - The future of vehicles is electric': Biden announces \$2.8B for battery supply chain
  - Canada Has an EV Edge, If It Acts Now
  - S.2.8B US EV supply-chain push appears to favor red states
  - Why Canada has the potential to become an EV battery supply chain powerhouse
  
- US battery supply chain investments reach US\$92 billion since Biden took office
- Biden's EV bet is a gamble on critical minerals
- Dead EV batteries turn to gold with US incentives
- UK Inflation Rate Reaches 40-Year High As Food Prices Surge
- Bank Of England To Get More Aggressive With 50 BPS Hike Later In The Week
- Huge Tax Cuts Are Being Questioned By Investors As Pounds Sinks
- Biden Threatens Windfall Tax As He Accuses Oil Companies Of War Profiteering
- Inflation In The United Kingdom Jumps To 41-Year High Of 11.1%
- China Sets 5% As Their Economic Growth Target For 2023
- Ukraine Secures First IMF Loan To A Country At War
- Largest Oil Refinery In Africa Launched By Aliko Dangote
- China's Economy Experienced A Growth Of 6.3% In The Second Quarter
- China To Kickstart Economy, After Plans To Improve Internal Migration
- A 40% Windfall Tax Was Approved By Italian Government As A Result Of Soaring Profits
- France Plans To End The Use Of Fossil Fuels By 2030
- South Africa Gets \$1 Billion Loan From World Bank To Tackle Power Crisis

Figure 7: Categories and examples of news articles in the dataset.



Figure 8: Distribution of sources in the news dataset.

According to the provided paragraphs:

```
### {}_Paragraphs_provided ###

extract a detailed hierarchical structure related to the EV battery supply chain.
The hierarchical structure should include the following levels:
- **Event**: Anything that happens related to the EV battery supply chain.
- **Event ID**: A unique identifier for each event.
- **Description**: A detailed 2-3 sentence explanation of the event.
- **Participants**: All sub-events related to this event and their importance, the importance needs
    to be set as 0 ~ 1, the higher the more important.
- **Gate**: The relationship between an event and its sub-events:
    - Use **'and'** if no sub-events can be missing.
    - Use **'or'** if some sub-events can be missing.
    - Use **'xor'** if sub-events cannot exist simultaneously.
- **Relations**: The event-event relations (e.g., ev1.1>ev1.2, which means ev1.2 happens after ev1.1)

If any level is empty, set its value to 'xxxx'.

Strictly use the exact following format for each event:
```
Event N
event: [Event Name]
event_id: evN
description: [Detailed Description]
participants: [Subevent 1] evN.1_P[Importance], [Subevent 2] evN.2_P[Importance], ...
Gate: [Gate]
Relations: [Event Relations]

Subevent N.1
subevent: [Subevent Name]
event_id: evN.1
description: [Detailed Description]
participants: [Subsubevent 1] evN.1.1_P[Importance], ...
Gate: [Gate]
Relations: [Event Relations]

Subsubevent N.1.1
subsubevent: [Subsubevent Name]
event_id: evN.1.1
description: [Detailed Description]
participants: [Subsubsubevent 1] evN.1.1.1_P[Importance], ...
Gate: [Gate]
Relations: [Event Relations]
```
``
```

Figure 9: Example of hierarchical structure extraction. (Part 1)

The prompt given to the LLMs is detailed and specific, ensuring that the models understand the exact format and type of information we are extracting. By integrating demonstration and CoT techniques, our prompt provides clear guidance to the LLMs, improving the accuracy and relevance of the extracted structures. Below is an example of the prompt used in Fig. 9.

To validate our approach, we tested the prompt with various texts from the schema learning dataset. The hierarchical structures extracted were compared with manually annotated structures to ensure accuracy and consistency. This process ensured that the LLMs reliably produced high-quality hierarchical structures that aligned with expert knowledge in

the EV battery supply chain domain.

## D Schema Generation & Merging

With human-in-the-loop schema induction, our schema learning dataset generated 125 individual schemas  $\langle S_1, S_2, \dots, S_{125} \rangle$ . To create a single comprehensive schema, it is essential to merge all individual schemas into a final schema ( $S_{final}$ ). The process of merging schema format files involves systematically integrating multiple schemas into a cohesive schema. The key components include *context*, *id*, *events*, and *relations*. These components determine the information in each event and its correlation with other events, hence the merging process must address all of them.

```
Use the provided example for guidance:
```

```
### Example:
```

```
**Input Paragraph**:
```

```
““  
Three main methods are used in lithium-ion recycling: pyrometallurgical, hydrometallurgical, bioleaching, and direct recycling. The battery is melted in a hot furnace to recover some of the cathode metal in pyrometallurgy. Pyrometallurgy employs extreme heat to transform metal oxides into cobalt, copper, iron, and nickel alloys. Although it has a straightforward process and a reasonably mature technology, the main drawbacks are its high cost and high environmental pollution. Hydrometallurgy is a metal recovery method involving aqueous solutions to perform leaching processes to precipitate a particular metal. In hydrometallurgy, specialized solution reagents are primarily used to leach the targeted metals out from the cathode substance. Although it is a highly effective and power-efficient method, its drawbacks include a lengthy production time and a complicated process. Combinations of both pyrometallurgy and hydrometallurgy are also used due to their advantages in sorting starting materials for cells. The bioleaching technique uses bacteria to retrieve precious metals, but it is challenging because the bacteria need a substantial amount of time to grow and are easily susceptible to contamination.  
““
```

```
**Extracted Hierarchical Structure**:
```

```
““  
Event 1  
event: lithium-ion recycling  
event_id: ev1  
description: Methods for recycling lithium-ion batteries including pyrometallurgical, hydrometallurgical, bioleaching, and direct recycling.  
participants: pyrometallurgical ev1.1_P1, hydrometallurgical ev1.2_P1, bioleaching ev1.3_P1  
Gate: or  
Relations: ev1.1>ev1.3, ev1.2>ev1.3  
  
Subevent 1.1  
subevent: pyrometallurgical  
event_id: ev1.1  
description: Employs extreme heat to transform metal oxides into cobalt, copper, iron, and nickel alloys.  
participants: metal oxides ev1.1.1_P1, cobalt ev1.1.2_P0.5, copper ev1.1.3_P0.5, iron ev1.1.4_P0.5, nickel alloys ev1.1.5_P0.5  
Gate: and  
Relations: ev1.1.1>ev1.1.2, ev1.1.1>ev1.1.3, ev1.1.1>ev1.1.4, ev1.1.1>ev1.1.5  
  
Subevent 1.2  
subevent: hydrometallurgy  
event_id: ev1.2  
description: Uses aqueous solutions to leach targeted metals out from the cathode substance.  
participants: xxxx  
Gate: xxxx  
Relations: xxxx  
  
Subevent 1.3  
subevent: bioleaching  
event_id: ev1.3  
description: Uses bacteria to retrieve precious metals.  
participants: xxxx  
Gate: xxxx  
Relations: xxxx  
““
```

Think about this extracted structure step by step:

Starting with the first sentence in the paragraph 'Three main methods are used in lithium-ion recycling: pyrometallurgical, hydrometallurgical, bioleaching, and direct recycling.' From this sentence, we learn that 'pyrometallurgical', 'hydrometallurgical', 'bioleaching, and direct recycling' are three methods of 'lithium-ion recycling', so select 'lithium-ion recycling' as the event, and the three methods as subevents and participants of 'lithium-ion recycling'.

Figure 9: Example of hierarchical structure extraction. (Part 2)

---

**Algorithm 2** Schemas Merging Pseudocode

---

```
1: Input: List of schemas
2: Output: Merged schema
3: 1. Merge contexts from all schemas:
4: for all schema in schemas do
5:   for all context in schema["@context"] do
6:     if context not in merged_contexts then
7:       Add context to merged_contexts
8:     end if
9:   end for
10: end for
11: 2. Merge events from all schemas by event name:
12: for all schema in schemas do
13:   for all event in schema["events"] do
14:     event_name = event["name"]
15:     if event_name not in merged_events then
16:       merged_events[event_name] = event
17:     else
18:       merged_events[event_name] = merge_event_details(merged_events[event_name], event)
19:     end if
20:   end for
21: end for
22: 3. Merge relations / update event IDs by event names:
23: for all schema in schemas do
24:   for all relation in schema["relations"] do
25:     subject_name = GET event name by event ID relation["relationSubject"]
26:     object_name = GET event name by event ID relation["relationObject"]
27:     if subject_name in name_to_id and object_name in name_to_id then
28:       relation["relationSubject"] = name_to_id[subject_name]
29:       relation["relationObject"] = name_to_id[object_name]
30:       if relation not in merged_relations then
31:         Add relation to merged_relations
32:       end if
33:     end if
34:   end for
35: end for
36: 4. Final Schema:
37: The final merged schema includes all merged contexts, events, and relations, and is saved for evaluation.
```

---

To begin the merging process, we first aggregate the *context* data from all schemas. Each *context* is added to a *merged\_contexts\_list*, ensuring that duplicate contexts are avoided. This step is crucial to maintain a unified context for the merged schema. Next, we proceed to merge events from all schemas. Using the event name as the identifier, we check if the event already exists in the *merged\_events\_list*. If the event exists, its details are merged with the existing event; otherwise, the event is added directly to the list. This ensures that all events are comprehensively integrated without duplication.

Following the merging of events, we then merge relations and update event IDs. This involves retrieving the event names from the event IDs for *re-*

*lationSubject* and *relationObject* and updating the relations accordingly. It is important to ensure that both *subject\_name* and *object\_name* are present in the *name\_to\_id* dictionary, which stores event names and their related event IDs. The updated relation is added to the *merged\_relations* list if it is not already present, ensuring all connections are accurately maintained.

Finally, the comprehensive merged schema ( $S_{final}$ ) is created by including all merged *contexts*, *events*, and *relations*. The detailed pseudocode for merging schemas is shown in Algorithm 2. This algorithm ensures that all relevant information is retained and accurately integrated, resulting in a comprehensive schema that encapsulates the full breadth of the data from the schema learning dataset. The final schema ( $S_{final}$ ) enables accurate and efficient knowledge extraction and organization, enhancing the utility of the dataset for downstream tasks such as event prediction and analysis.

## E Schema Management System

The schema management interface (Fig. 10) facilitates the visualization, editing, and management of schemas. It includes the following modules:

### E.1 Schema Viewer

The schema viewer is crucial for visualizing schemas, providing an intuitive representation of events. It organizes events into a left-to-right tree structure, highlighting parent-child relationships. Within this structure, before-after relationships among child nodes are indicated through arrows and vertical ordering. Users can expand event nodes to reveal details such as descriptions, importance levels, and participant roles.

Key features of the schema viewer include:

- **Interactive Exploration:** Users can click on nodes to expand or collapse details about events and sub-events.
- **Contextual Information:** Hovering over a node displays additional context and metadata associated with the event.
- **Dynamic Layout:** The tree structure dynamically adjusts to accommodate the addition or removal of nodes, maintaining a clear and organized visual representation.
- **Collapsible Subtrees:** Users can collapse and expand subtrees to manage large schemas.

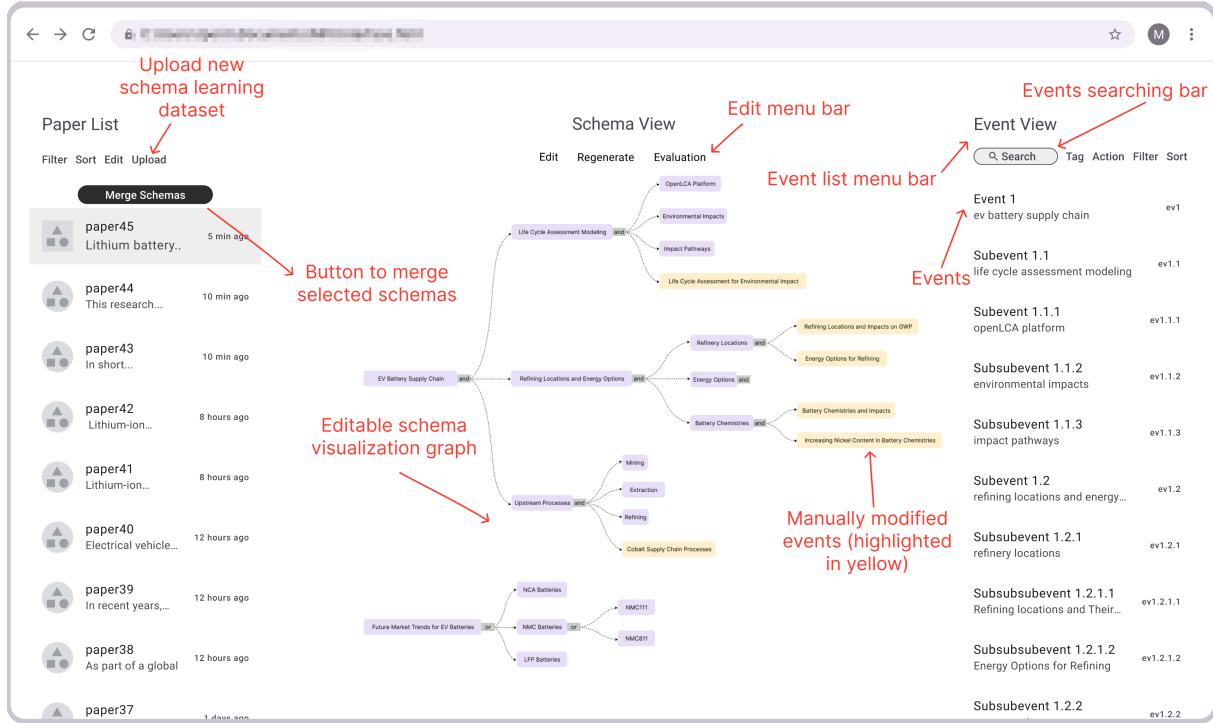


Figure 10: User interface for our schema management system.

- **Search Functionality:** A search bar allows users to quickly locate specific events or entities within the schema.
- **Real-Time Data Binding:** The viewer updates in real-time as changes are made, ensuring the displayed schema is always current.
- **Highlighting and Filtering:** Users can highlight specific paths or filter events based on criteria such as importance or type.

## E.2 Schema Editor

The schema editor allows users to interactively modify schemas. Users can add, edit, and delete events, sub-events, and relationships within the schema. Key functionalities include:

- **Drag-and-Drop Interface:** Users can drag and drop nodes to reassigned parent-child relationships or reorder events.
- **Form-Based Editing:** Clicking on a node opens a form where users can edit event details, such as descriptions, importance levels, and participant roles.
- **Validation Checks:** The editor performs real-time validation to ensure that all changes adhere to the schema format and constraints.
- **Undo/Redo Features:** Users can easily undo or redo changes to maintain the integrity of the schema editing process.

- **Schema Versioning:** The editor maintains different versions of schemas, allowing users to track changes over time and revert to previous versions if necessary.
- **Bulk Operations:** Users can perform bulk operations such as adding multiple events or updating several nodes at once.
- **Conflict Resolution:** The editor provides tools to resolve conflicts when multiple users make changes simultaneously.

## E.3 Frontend Architecture

The frontend of system is implemented as a single-page web application using React and TypeScript. This setup connects to an API server that provides application logic and access to a centralized schema database. The use of a browser-based application offers several advantages, including no need for user installations, centralized data management, and extensive functionality through JavaScript libraries. Key components include:

- **React<sup>9</sup>:** A JavaScript library for building user interfaces, providing the foundation for the application's dynamic and responsive design.
- **TypeScript<sup>10</sup>:** A statically typed superset of JavaScript, enhancing code reliability and

<sup>9</sup><https://react.dev/>

<sup>10</sup><https://www.typescriptlang.org/>

maintainability.

- **GoJS**<sup>11</sup>: A JavaScript library for creating interactive diagrams, enabling robust schema visualization.
- **API Integration:** The frontend communicates with the backend through API calls, fetching and submitting schema data.
- **Responsive Design:** The application is optimized for various screen sizes and devices, ensuring usability across different platforms.
- **State Management:** The application uses state management libraries such as Redux to manage and synchronize the state of the schema data across different components.
- **Performance Optimization:** Techniques such as code splitting and lazy loading are employed to ensure fast load times and smooth interactions.

The client-side application requests Schema Definition Files (SDF) from the API server and displays them to users. Edits to the SDF are maintained locally until the user saves the changes, synchronizing the server-side copy with the client's modifications. A simple locking mechanism is employed to prevent simultaneous edits by multiple users on the same schema, ensuring data integrity.

#### E.4 Backend Architecture

The backend of the interface is developed in Python, leveraging the Falcon web server framework, served by Gunicorn and nginx, and supported by a SQLite database. The backend is designed to be lightweight, minimalist, and easy to comprehend. Most functionalities are concentrated in the frontend to maintain responsiveness and interactivity, allowing the backend to focus primarily on data management. Python's versatility and popularity make it a suitable choice for the dynamic requirements of the system. Static typing in Python is enforced using Mypy<sup>12</sup> to facilitate development and reduce trivial bugs. Key components include:

- **Falcon**<sup>13</sup>: A minimalist web framework for building high-performance APIs, facilitating efficient communication between the frontend and backend.
- **Gunicorn**<sup>14</sup>: A Python WSGI HTTP server for running web applications, ensuring robust

and scalable performance.

- **nginx**<sup>15</sup>: A high-performance web server and reverse proxy, providing load balancing and enhancing security.
- **SQLite**<sup>16</sup>: A lightweight, disk-based database, chosen for its simplicity and reliability.
- **RESTful API**<sup>17</sup>: The backend exposes a RESTful API for the frontend to interact with schema data, supporting CRUD operations.
- **Security Features:** Implementations such as HTTPS, authentication, and authorization to ensure data privacy and integrity.
- **Scalability:** The architecture is designed to scale horizontally, with load balancers and database replication as needed.

#### E.5 AI-Driven Suggestions

The interface incorporates AI-driven suggestions to assist users in schema creation and modification. Large Language Models (LLMs) analyze existing schemas and user inputs to provide recommendations for schema elements, relationships, and structures. These suggestions are presented in real-time, enhancing user productivity and ensuring the creation of accurate and comprehensive schemas.

Key features of AI-driven suggestions include:

- **Contextual Recommendations:** The system provides context-aware suggestions based on the current schema and user actions.
- **Smart Auto-Completion:** As users type or modify schema elements, the interface offers auto-completion options to expedite the editing process.
- **Error Detection:** The AI models detect potential errors or inconsistencies in the schema and suggest corrections.
- **Learning from User Feedback:** The AI models improve over time by learning from user feedback and interactions, refining their suggestions and increasing accuracy.
- **Interactive Tutorials:** The interface includes tutorials and guidance to help users understand and leverage AI-driven suggestions effectively.

---

<sup>11</sup><https://gojs.net/latest/index.html>

<sup>12</sup><https://mypy-lang.org/>

<sup>13</sup><https://falcon.readthedocs.io/>

<sup>14</sup><https://gunicorn.org/>

<sup>15</sup><https://nginx.org/en/>

<sup>16</sup><https://www.sqlite.org/>

<sup>17</sup><https://restfulapi.net/>

## F Details of Event Extraction

### F.1 Event Span Identification

Event span identification involves locating and marking the spans of events within input text. We use two models for this task:

**Base Model:** This model is a fine-tuned version of the RoBERTa-large language model (Liu et al., 2019), trained on an internally annotated dataset. The task is formulated as sequence tagging, where the model identifies the start and end positions of event spans. For instance, in the context of supply chain disruptions, the model identifies spans corresponding to events like factory shutdowns, transport delays, or material shortages. This aligns with the cross-sentence event detection described in the main text:

$$\text{EventDetect}_{\text{multi-sentence}}(\mathbf{T}) \rightarrow \mathbf{E}_C \quad (16)$$

where  $\mathbf{T}$  represents the input text and  $\mathbf{E}_C$  the detected events. The model uses contextual information from neighboring sentences to accurately detect event boundaries, ensuring that even complex events spanning multiple sentences are correctly identified.

**Guided Model:** Inspired by Wang et al. (2021), this model uses a query-based approach to focus on schema-related events. The process involves two stages as follows:

1. **Discriminator Stage:** Queries representing event types are paired with sentences to predict if the query corresponds to an event type mentioned in the sentence. For example, queries include "factory shutdown due to labor strike" or "delay in shipping materials." This stage helps in filtering sentences that are likely to contain relevant events.
2. **Span Extraction Stage:** Sentences identified in the discriminator stage are further processed to extract event spans using sequence tagging. This ensures that the extracted spans are relevant to the supply chain context. By using sequence tagging, the model accurately marks the start and end points of events within the identified sentences.

This approach supports the cross-sentence event detection described in the main text, enriching event spans with relevant context and ensuring high precision in event identification.

### F.2 Event Argument Extraction

Event argument extraction involves identifying the roles and participants associated with events. This task is framed as extractive question answering, where the model extracts argument spans from the text based on role-specific questions. We fine-tune RoBERTa-large (Liu et al., 2019) on our internally annotated dataset with a sequence tagging loss function. For supply chain disruptions, arguments might include the specific factories, transportation modes, or materials directly affected by the event.

The extraction process is as follows:

1. **Role-Specific Questions:** The model is trained to answer questions like "Which factory was shut down?" or "What material was delayed?" This method ensures that the arguments are specific and relevant.
2. **Contextual Embeddings:** This step is enriched by contextual embeddings generated by BERT:

$$\text{BERT}_{\text{context}}(\mathbf{E}_C) \rightarrow \mathbf{C}_E \quad (17)$$

generating contextual embeddings  $\mathbf{C}_E$ . These embeddings provide rich semantic information, enabling the model to better understand the context and improve the accuracy and relevance of the extracted arguments.

### F.3 Time Expression Linking & Normalization

Time expression linking connects time expressions to their corresponding events. Similar to argument extraction, this task uses extractive question answering to find start and end times for events. We fine-tune RoBERTa-base using the TempEval3 dataset (UzZaman et al., 2012).

The process includes:

1. **Extraction:** The model identifies time expressions within the text and links them to the corresponding events, ensuring that the timeline of events is accurately captured.
2. **Normalization:** Identified time expressions are then normalized into standard formats using SUTime (Chang and Manning, 2012) and HeidelTime (Strötgen and Gertz, 2013). For example, expressions like "next Monday" are converted into specific dates.

For supply chain disruptions, this ensures that timelines for events like "shipment delayed from March 15 to March 20" are accurately captured. This integrates into the event parameter extraction process, ensuring coherence and consistency.

#### F.4 Event Temporal Ordering

Event temporal ordering determines the chronological sequence of events. We frame this task as extractive question answering to address label imbalance issues, fine-tuning ROBERTa-large (Liu et al., 2019) with a sequence tagging loss.

Steps include:

1. **Pairwise Temporal Relations:** The model identifies pairwise temporal relations between events, such as "Event A happened before Event B."
2. **Consistency Checking:** Pairwise temporal relations are processed using Integer Linear Programming (ILP) (Schrijver, 1998) to ensure consistency and resolve any conflicts. This method helps in constructing a coherent timeline of events.

This is crucial for understanding the sequence of disruptions in supply chains, such as how a factory shutdown leads to delayed shipments. This aligns with the logical constraints and argument coreference to maintain event relationships modeled by GCNs:

$$\text{LogicCoref}(\mathbf{P}_C) \rightarrow \mathbf{P}_F \quad (18)$$

#### F.5 Coreference Resolution

We perform both within-document and cross-document coreference resolution using models finetuned on datasets like OntoNotes 5.0 (Pradhan et al., 2013).

The resolution process involves:

1. **Entity Clustering:** Entity and event coreference clusters are identified and linked to ensure consistency across documents. This helps in tracking the same entities and events mentioned in different parts of the text.
2. **Cross-Document Linking:** Linking entities and events across multiple documents ensures that all references to a specific factory, supplier, or shipment are recognized as the same entity.

This is critical for tracking entities like factories, suppliers, and shipments across multiple reports of supply chain disruptions. This supports the coreference resolution and event linking described in the main text:

$$\text{CorefLink}(\mathbf{E}_C) \rightarrow \mathbf{E}_L \quad (19)$$

yielding linked events  $\mathbf{E}_L$ .

#### F.6 Graph Convolutional Networks (GCNs) for Event Relationship Modeling

We leverage Graph Convolutional Networks (GCNs) to model complex event relationships and assess each event's impact. This involves constructing a graph where events are nodes and their interactions are edges.

Steps include:

1. **Node Importance Calculation:** Each node's importance is calculated using centrality measures, such as degree centrality, betweenness centrality, and eigenvector centrality. These measures help in understanding the influence of each event within the network.
2. **Edge Impact Calculation:** Edges represent the magnitude of impact, quantified by measures such as event severity and frequency of occurrence.

The impact score is then calculated as:

$$\text{ImpactScore}(e_i) = \text{Centrality}(e_i) + \text{Magnitude}(e_i) \quad (20)$$

where  $\text{Centrality}(e_i)$  reflects the event's importance within the network, and  $\text{Magnitude}(e_i)$  quantifies the event's impact intensity.

#### F.7 Logical Constraints and Argument Coreference

To ensure the robustness of our event extraction pipeline, we apply logical constraints and argument coreference resolution.

This involves multiple steps to refine the extracted event parameters and ensure logical consistency:

##### Logical Constraints Application:

1. **Defining Logical Rules:** We define a set of logical rules to maintain consistency within the extracted events. These rules include:
  - *Temporal constraints:* An event must occur before another if there is a chronological dependency.

- *Causal relationships*: If Event A causes Event B, then Event A must be identified as a precursor to Event B.
2. **Implementation:** The defined logical rules are implemented using a logic-based reasoning system that checks for any violations and rectifies them. For instance, if an event is detected as occurring before its cause, the system flags this inconsistency and corrects the sequence.

#### **Argument Coreference Resolution:**

1. **Coreference Detection:** We identify coreferences within and across documents. This involves detecting instances where different expressions refer to the same entity or event.
  - *Within-Document Coreference*: Ensures that all mentions of an entity within a single document are linked.
  - *Cross-Document Coreference*: Links mentions of the same entity or event across multiple documents to ensure global consistency.

#### **2. Refinement Process:**

- *Cluster Formation*: Entities and events identified as coreferent are grouped into clusters.
- *Coreference Chains*: We create chains of coreferent mentions, which are used to refine event parameters and ensure that all related mentions are consistently linked.
- *Manual Verification*: After automatic coreference resolution, manual verification is performed by domain experts to ensure accuracy and address any ambiguities.

#### **Combining Logical Constraints & Coreference:**

1. **Integration:** The logical constraints and coreference resolution processes are integrated to produce a coherent and logically consistent set of event parameters:

$$\text{LogicCoref}(\mathbf{P}_C) \rightarrow \mathbf{P}_F \quad (21)$$

2. **Validation:** The final set of event parameters  $\mathbf{P}_F$  undergoes a validation process to ensure

that all logical rules and coreference chains are satisfied. This step is crucial for maintaining the integrity of the event extraction pipeline.

3. **Feedback Loop:** A continuous feedback loop is established where the output is reviewed and refined based on new data and expert feedback. This iterative process helps in improving the model's performance over time.

By applying these detailed logical constraints and advanced coreference resolution techniques, we ensure that the event extraction pipeline produces high-quality, reliable, and contextually accurate event data, which is essential for robust supply chain disruption analysis.

## **G Details of Event Matching & Instantiation**

Event matching and instantiation involve aligning a schema from the schema library with events extracted by the schema extraction component, specifically for predicting supply chain disruptions. This process begins by instantiating one of the  $E_{\text{schema}}$  from the integrated library or selecting the extracted event  $E_{\text{ext}}$  that best matches the schema event  $E_{\text{schema}}$ . Subsequently, the task entails matching events in the  $E_{\text{schema}}$  with their corresponding events in  $E_{\text{ext}}$  extracted from the news dataset. Events in both  $E_{\text{ext}}$  and  $E_{\text{schema}}$  are organized in a highly structured manner, with parent events divided into child events. Events also contain temporal information, indicating that some events must precede others. Logical relationships are also defined: AND-gates connect all necessary child events for a parent event, OR-gates connect one or more needed child events, and XOR-gates indicate that only one child event can be present.

For example, a document about a raw material shortage in the EV battery supply chain might align with a "Supply Chain Disruption" schema in the schema library. Following the instantiation, a "notify suppliers" event in the schema might match with a graph  $G$  event describing a notification sent to cobalt suppliers. The "suppliers" participant of the schema event might match with the "cobalt suppliers" participant of the extracted event.

### **G.1 Matching Process & Techniques**

Our approach to event matching and instantiation involves several key steps and techniques to en-

sure accurate alignment between schema events and extracted events. This is particularly critical in the context of predicting supply chain disruptions, where precise event matching can provide actionable insights.

### G.1.1 Similarity Calculation

To determine the similarity between schema events and extracted events, we calculate a similarity score based on semantic and structural similarities. Semantic similarity (SemSim) is computed using sentence transformers to encode the semantic content of events. Structural similarity (StrSim) takes into account the hierarchical and temporal relationships between events.

**Semantic Similarity:** We use a sentence transformer model to encode events into semantic vectors. The cosine similarity between these vectors provides a measure of how semantically similar two events are:

$$\text{SemSim}(E_{\text{ext}}, E_{\text{schema}}) = \frac{\mathbf{v}_{\text{ext}} \cdot \mathbf{v}_{\text{schema}}}{\|\mathbf{v}_{\text{ext}}\| \|\mathbf{v}_{\text{schema}}\|} \quad (22)$$

where  $\mathbf{v}_{\text{ext}}$  and  $\mathbf{v}_{\text{schema}}$  are BERT embeddings of extracted and schema events.

**Structural Similarity:** We consider the context of events within their respective hierarchies. For example, an event's predecessors and successors, its parent event, and its child events all contribute to its structural context. Events with similar structures in both schema and extracted graphs are more likely to match:

$$\text{StrSim}(E_{\text{ext}}, E_{\text{schema}}) = \frac{|\mathbf{P}_{\text{ext}} \cap \mathbf{P}_{\text{schema}}|}{|\mathbf{P}_{\text{ext}} \cup \mathbf{P}_{\text{schema}}|} \quad (23)$$

where  $\mathbf{P}_{\text{ext}}$  and  $\mathbf{P}_{\text{schema}}$  are the parameter sets for the extracted and schema events.

### G.1.2 Event Matching

Once the similarity scores are calculated, we match each extracted event  $E_{\text{ext}}$  with the schema event  $E_{\text{schema}}$  that has the highest similarity score. This involves instantiating the schema event with information from the extracted event, ensuring that all relevant details and relationships are preserved.

**Example:** Consider a schema event "notify suppliers" in the context of a raw material shortage. An extracted event describing an email notification to cobalt suppliers would match this schema event if the similarity score is high. The instantiation process involves mapping the "suppliers" participant

in the schema to the "cobalt suppliers" entity in the extracted event:

$$\text{Instantiate}(E_{\text{matched}}, \mathbf{S}_{\text{schema}}) \rightarrow \mathbf{E}_{\text{inst}} \quad (24)$$

where  $\mathbf{E}_{\text{inst}}$  is the instantiated event enriched with attributes from the schema.

### G.1.3 Consistency Checks

After matching events, we perform consistency checks to ensure that the instantiated schema adheres to logical and temporal constraints. This includes verifying that:

- All necessary child events are present (AND-gates).
- At least one required child event is present (OR-gates).
- Only one of the mutually exclusive child events is present (XOR-gates).

These checks ensure that the instantiated schema is logically coherent and temporally consistent:

$$\text{ConsistencyCheck}(\mathbf{E}_{\text{inst}}, \mathbf{S}_{\text{schema}}) \quad (25)$$

## G.2 Continuous Improvement

To enhance the accuracy and robustness of our matching and instantiation process, we incorporate continuous improvement through manual review and feedback from domain experts. This involves:

- Validating the instantiated events with domain experts to ensure they accurately reflect real-world scenarios.
- Refining our models based on feedback, adjusting similarity metrics, and improving our semantic and structural encoding techniques.
- Iteratively updating our schema library and extraction models to incorporate new insights and improve performance.

By leveraging domain expertise and feedback, we continually refine our event matching and instantiation process, ensuring it remains effective and relevant for predicting and analyzing supply chain disruptions.

## H Details of Disruption Prediction

Given an instantiated event graph  $\mathbf{G}_{\text{inst}} = (N, E)$ , where  $N$  represents event nodes (e.g., specific supply chain activities) and  $E$  denotes event-event

---

**Algorithm 3** Event Matching and Instantiation

---

```

1: Input: Extracted events  $E_{\text{ext}}$ , schema library events
    $E_{\text{schema}}$ 
2: Output: Instantiated events  $E_{\text{inst}}$ 
3: Calculate Similarity           ▷ Compute similarities
4: for each  $E_{\text{ext}}$  in  $E_{\text{ext}}$  do
5:   for each  $E_{\text{schema}}$  in  $E_{\text{schema}}$  do
6:      $\text{Sim}(E_{\text{ext}}, E_{\text{schema}}) \leftarrow \alpha \cdot \text{SemSim}(E_{\text{ext}}, E_{\text{schema}}) +$ 
       $\beta \cdot \text{StrSim}(E_{\text{ext}}, E_{\text{schema}})$ 
7:   end for
8: end for
9: Match Events ▷ Match extracted events to schema events
10: for each  $E_{\text{ext}}$  in  $E_{\text{ext}}$  do
11:    $E_{\text{matched}} \leftarrow \arg \max_{E_{\text{schema}}} \text{Sim}(E_{\text{ext}}, E_{\text{schema}})$ 
12:    $E_{\text{inst}} \leftarrow \text{Instantiate}(E_{\text{matched}}, S_{\text{schema}})$ 
13:   Perform ConsistencyCheck( $E_{\text{inst}}, S_{\text{schema}}$ )
14: end for
15: Continuous Improvement           ▷ Manual review and
   feedback
16: for each  $E_{\text{inst}}$  do
17:   UpdatedModels  $\leftarrow \text{ValidateRefine}(E_{\text{inst}})$ 
18: end for
19: Return: Instantiated events  $E_{\text{inst}}$ 

```

---

temporal links (e.g., dependencies or sequences of activities), the goal is to classify whether unmatched schema events (nodes) could potentially occur within this graph.

Formally, let  $I$  be the set of matched schema events within the graph. The task involves classifying each node in the remaining schema event nodes, represented by  $N \setminus I$ , as a missing event (positive or negative) given the instantiated graph.

To address the limitations of existing methods, we developed a novel approach that leverages the structural information within the schema graph and incorporates logic gates and hierarchies. Our approach consists of three stages: (1) schema-guided prediction, (2) constrained prediction, and (3) argument coreference.

### H.1 Schema-Guided Prediction

In this stage, we utilize a trained graph neural network specifically designed for schema graphs to score and select unmatched events in the instantiated graph. Key steps include:

- **Graph Neural Network:** A GCN is trained on schema graphs to learn representations of nodes and edges. The propagation rule is given by:

$$\mathbf{H}^{(l+1)} = \sigma(\mathbf{A}\mathbf{H}^{(l)}\mathbf{W}^{(l)}) \quad (26)$$

where  $\mathbf{H}^{(l)}$  is the hidden state at layer  $l$ ,  $\mathbf{A}$  is the adjacency matrix,  $\mathbf{W}^{(l)}$  is the weight matrix, and  $\sigma$  is a non-linear activation function.

- **Node Scoring:** Using the learned representations, the GCN scores and selects unmatched events in the instantiated graph.

- **Prediction Output:** The first-stage prediction output consists of the most likely missing events.

### H.2 Constrained Prediction

This stage applies logical constraints and hierarchical relations to refine the initial predictions from the schema-guided prediction stage. Key steps include:

- **Logical Constraints:** We refine initial predictions ( $\hat{y}$ ) to produce final predictions ( $\hat{y}'$ ) that adhere to known rules:

$$\begin{aligned} \hat{y}' &= \arg \min_{\hat{y}' \in \mathcal{Y}} \text{Constrain}(\hat{y}) \\ \text{subject to } \mathcal{C}(\hat{y}') &= \text{true} \end{aligned} \quad (27)$$

where  $\mathcal{C}$  represents constraint sets. For example, a constraint might ensure that a major supplier's disruption increases risk for dependent manufacturers.

- **Hierarchical Relationships:**

- **Child-to-Parent Propagation:** If a child event node is predicted or matched, its parent node is also predicted.
- **AND-Siblings Propagation:** If a predicted node has AND-sibling nodes, all its siblings are also predicted.
- **Iterative Refinement:** The constrained prediction approach is applied iteratively until no further nodes can be predicted.

### H.3 Argument Coreference

In this phase, we utilize coreference entity links and instantiated entities to generate predictions for the arguments associated with the predicted events. Key steps include:

- **Coreference Links:** Coreference entity links specified in the schema are used to ensure consistency among entity mentions:

$$\begin{aligned} R_{ij} &= \arg \max_{E_i, E_j \in \mathcal{E}} \text{Coref}(E_i, E_j) \\ \text{subject to } \text{Coref}(E_i, E_j) &= \text{true} \end{aligned} \quad (28)$$

where  $(E_i, E_j)$  denotes each event pair and  $R_{ij}$  represents their relation.

- **Instantiated Entities:** Instantiated entities from the previous stages are leveraged to generate arguments for predicted events.
- **Final Output:** This stage produces the final prediction output, including both events and their arguments.

## I Experiment Details

### I.1 Experiment Setup

In Experiment 4.1, we evaluate the efficacy of three distinct Large Language Models (LLMs) in extracting hierarchical structures from our schema learning dataset. Leveraging domain expert knowledge, we annotate individual schemas for each article in our academic corpus using our proprietary system viewer and editor. We then employ the methodology outlined in Appx. D to synthesize these schemas into an integrated library. This combination of individual schemas and the integrated library serves as the ground truth for our hierarchical information extraction phase. Our schema learning performance evaluation consists of two key components. First, we compare the hierarchical information extracted by the three LLMs against our established ground truth. Second, we assess the consistency, accuracy, and completeness of the hierarchical structures derived from the textual content of each article in the schema learning dataset, with domain experts actively participating in this evaluation process.

In Experiment 4.2, we apply a similar annotation methodology to our news dataset as used for the schema learning dataset. However, annotating the news dataset presents unique challenges, as news reports typically do not explicitly elucidate the connections between events. Instead, they often employ speculative language to describe event interrelations. To ensure annotation accuracy, we heavily rely on domain knowledge derived from scenario documents throughout the annotation process. Subsequently, we utilize the ground truth extracted from these reports to evaluate our system’s performance in predicting news report outcomes.

### I.2 Evaluation Metrics

**Subjective Schema Learning.** For subjective schema evaluation, we ensure that the event schemas generated from each paper and news report are consistent, accurate, and complete. The schema derived from academic papers demon-

strates a logical hierarchical structure, while the schema produced from news reports presents a well-defined temporal sequence. Experts manually review the schemas to verify these attributes, providing qualitative feedback on the logical coherence and comprehensiveness of the extracted structures. Each schema is rated on a scale from 1 to 5, where 1 indicates poor quality and 5 indicates excellent quality.

**Objective Disruption Detection.** We compare the instantiated schemas learned by our system with manually annotated ground truth to assess the degree of overlap. This comparison uses an evaluation metric similar to Smatch (Cai and Knight, 2013), which involves breaking down both our schema and the ground truth into quadruples of the form *relation(event1, event2, importance)*. For instance, the *event* of *Raw Material Mining* includes the *subevent* of *Lithium Mining* with an associated importance value, represented by the quadruple *subevent(raw material mining, lithium mining, importance)*. Other relations include participants, gates, sequential events, etc.

To evaluate the results, we 1) map the events in the learned schema  $S_l$  to those in the ground-truth schema  $S_{gt}$ , 2) establish a one-to-one mapping of quadruples between the learned schema  $S_l$  and the ground-truth schema  $S_{gt}$ , 3) calculate Precision, Recall, and F-score as follows:

$$\text{Precision} = \frac{\text{number of matched quadruples in } S_l}{\text{total quadruples in } S_l} \quad (29)$$

$$\text{Recall} = \frac{\text{number of matched quadruples in } S_l}{\text{total quadruples in } S_{gt}} \quad (30)$$

$$\text{F-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (31)$$

## J Disruption Prediction Case Studies

### J.1 Case 1: Lithium Supply Chain Disruption in Early 2023

In early 2023, significant disruptions in the lithium supply chain were caused by escalating geopolitical tensions between Australia and China. As Australia is one of the world’s largest suppliers of lithium, political factors heavily influenced its export policies, severely impacting the global supply chain for EV batteries, which rely heavily on lithium.

**System Prediction:** Our system accurately predicted the potential supply disruption by analyzing various news reports on geopolitical developments and export data. The system monitored news related to geopolitical tensions between Australia and China, analyzed export data indicating changes in Australia's lithium export policies, and integrated insights from scenario documents highlighting the dependence of the EV battery supply chain on Australian lithium exports.

**Outcome:** The system flagged the risk of Australia's export restrictions to China, providing early warnings of potential disruptions in the EV battery supply chain. This allowed stakeholders to proactively seek alternative sources and mitigate the impact on production.

### J.2 Case 2: Nickel and Cobalt Supply Issues in March 2023

In March 2023, a major disruption in the global supply chain occurred due to large-scale worker strikes and regulatory changes in the Democratic Republic of Congo (DRC), a primary supplier of cobalt. Cobalt is crucial for EV batteries, and the disruption had a significant negative impact on the global supply chain.

**System Prediction:** Our system successfully forecasted the potential supply chain interruptions by analyzing news reports on strike activities and updates on government regulations in the DRC. It also assessed historical data on cobalt supply and demand to identify vulnerabilities and integrated expert feedback on the impact of labor strikes and regulatory changes on cobalt production.

**Outcome:** The system provided early warnings about the potential disruptions, enabling companies to adjust their supply chain strategies. This included diversifying sources of cobalt and increasing inventories to buffer against supply shortages.

### J.3 Case 3: Impact of the Inflation Reduction Act in August 2023

In August 2023, the United States passed the Inflation Reduction Act, which included significant incentives for domestic EV battery production. This led to a rapid increase in investments but also highlighted potential material shortages, causing disruptions in the EV battery supply chain.

**System Prediction:** Our system predicted the possibility of short-term material shortages by analyzing the market response data to the Inflation Re-

duction Act, monitoring global distribution reports of EV battery materials, and assessing the impact of increased domestic production incentives on the supply and demand balance.

**Outcome:** The system identified the risks posed by the sudden increase in demand for battery materials, providing early warnings to stakeholders. This allowed them to take proactive measures such as securing long-term supply contracts and exploring alternative materials to mitigate potential shortages.

## K SHIELD's User Interface

The SHIELD user interface is designed to be intuitive and user-friendly, facilitating the efficient upload and analysis of news reports. It comprises two main sections: the news report upload and the disruption analysis results.

**News Report Upload.** On the right side of the interface, users can upload their collected news report texts. It includes a text box for input and a submission button to upload the report (see Fig. 11a). Key features include:

- **Upload Box:** Allows users to paste or type their news report texts.
- **Submit Button:** Initiates the analysis process once the report is uploaded.
- **Uploaded Reports List:** Displays previously uploaded news reports, enabling users to review and compare past submissions easily.

**Disruption Analysis Results.** After submitting a news report, users can view the real-time results of the disruption analysis on the left side of the interface (see Fig. 11b). The comprehensive overview of the analysis include:

- **Generated Schema:** Displays the hierarchical of events identified in the news report.
- **Events List:** Lists all detected events and their details, allowing to see which events were identified and how they are connected.
- **Evaluation Score:** Shows the real-time evaluation score, assessed against the schema library for accuracy and completeness.
- **Schema Editing:** Allows to edit the generated schema. Users can make changes to the structure, relationships, and details of events.
- **Regenerate Evaluation:** Users can choose to regenerate the evaluation score based on the edited schema, ensuring that the modifications are reflected in the updated score.

## L Author Contributions

### Schema Learning Dataset:

- Yuzhi Hu: Created the scenario document.
- Yifei Dong: Collected paper lists and Wiki-data lists.
- Aike Shi: Collected Wikipedia lists and weekly report lists.
- Yifei Dong, Wei Liu, and Aike Shi: Extracted paragraphs from collected articles.

### Supply Chain News Dataset:

- Yuzhi Hu: Conducted data crawling and classification.
- Yuzhi Hu, Yifei Dong, Wei Liu, and Aike Shi: Labeled ground truth for the news dataset.

### Schema Learning:

- Yifei Dong: Designed and modified prompts for generating structured information, designed the format for structured information, wrote scripts for converting structured information to SDF, and for generating structured information using Llama3 (partially contributed by Aike Shi).
- Yifei Dong, Aike Shi, and Wei Liu: Generated structured information using GPT-4o with zero-shot learning.

### Human Curation:

- Yifei Dong, Wei Liu, Aike Shi, and Yuzhi Hu: Curated LLM-generated SDFs.
- Yifei Dong and Aike Shi: Wrote scripts for schema merging.

### System Construction:

- Zhi-Qi Cheng: Provided guidance for system implementation, designed the system prototype, and performed system implementation.
- Yifei Dong: Performed system implementation and debugging (with minor contributions from Aike Shi).

### Evaluation:

- Wei Liu and Yifei Dong: Designed evaluation metrics.
- Aike Shi and Yifei Dong: Wrote evaluation scripts.

### Frontend Interface:

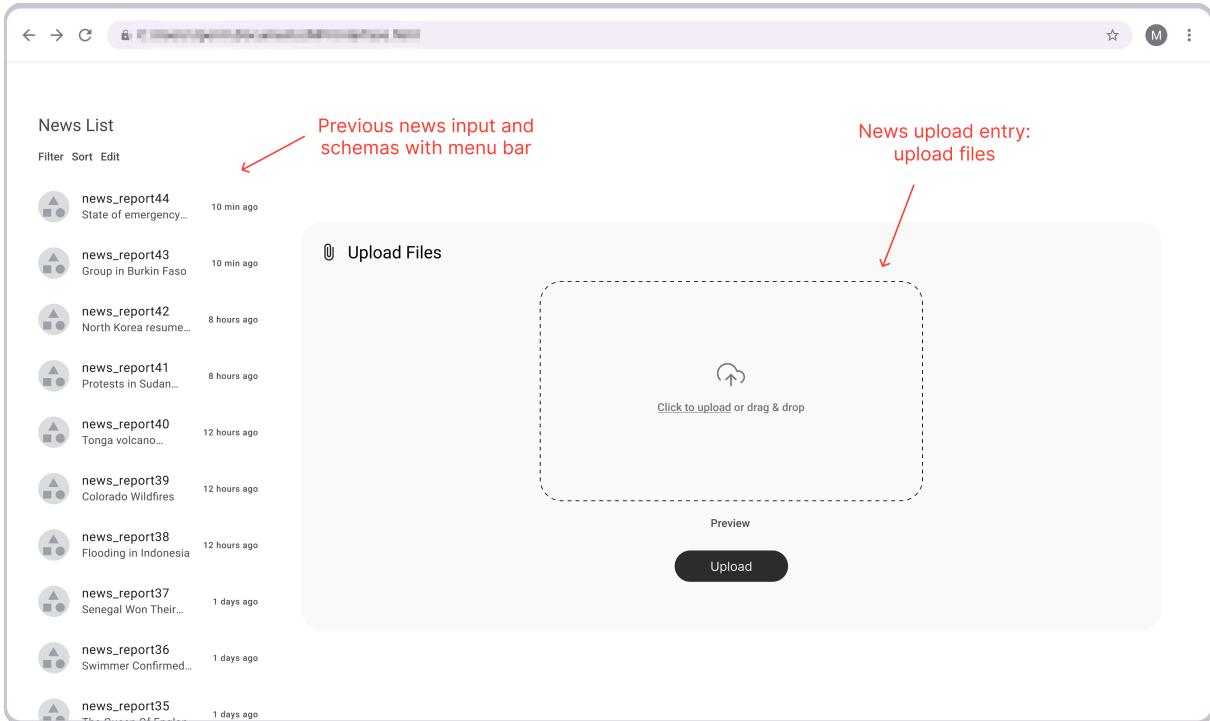
- Yifei Dong: Provided the design.
- Aike Shi and Wei Liu: Implemented interface.
- Wei Liu: Optimized the interface.

### Paper:

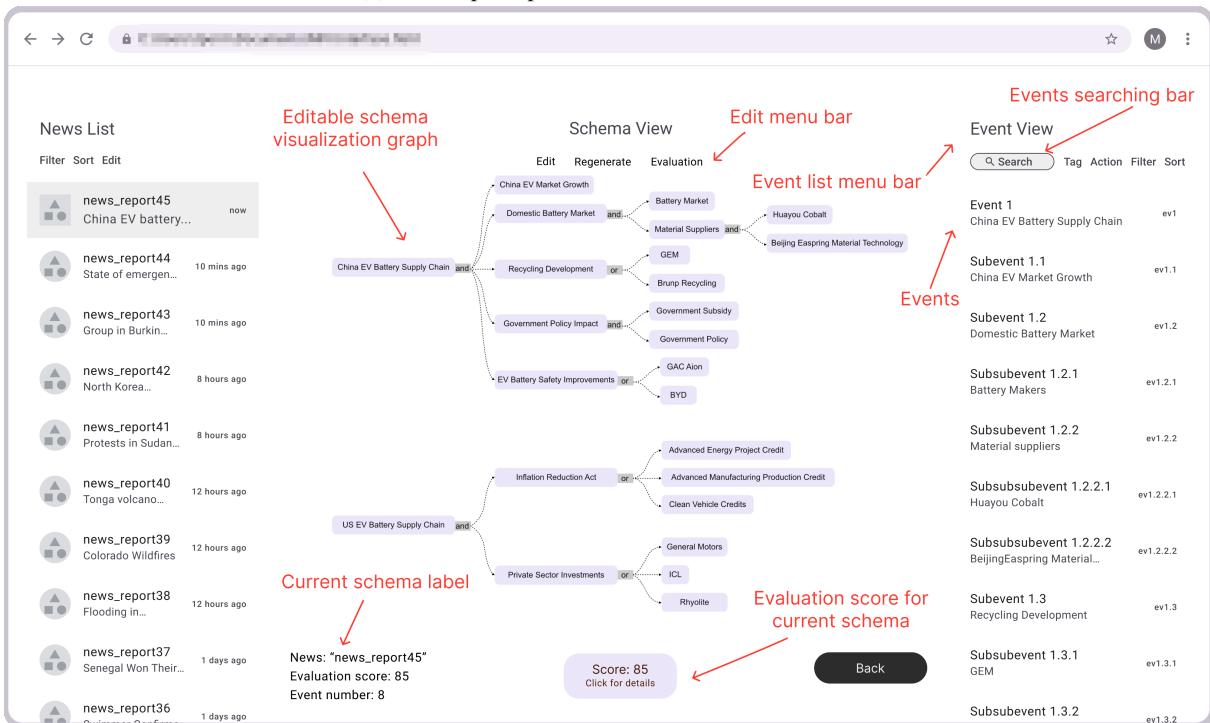
- Yifei Dong: Contributed to all sections except the news dataset part.
- Yuzhi Hu: Contributed to the news dataset and schema dataset parts and created and optimized Figs. 5, 6, 7, and 8.
- Wei Liu, Yifei Dong, and Aike Shi: Contributed to Figs. 1, 2, 3, 4, 9, 10, and 11.
  - Yifei Dong and Wei Liu designed all figures.
  - Wei Liu created and optimized all figures.
  - Aike Shi designed the user interface elements for Figs. 4, 10, and 11.
- Jason O'Connor: Contributed to paper revisions and suggestions.
- Zhi-Qi Cheng: Organized and rewrote the paper, and supervised the modification of all figures.

### Feedback and Guidance:

- Jason O'Connor: Provided feedback and project guidance from a supply chain expert perspective.
- Kate Whitefoot and Alexander Hauptmann: Provide guidance and supervision for the entire project.



(a) News report upload section of the user interface.



(b) Visualization and editing of the final prediction results.

Figure 11: User interface for the disruption prediction analysis in SHIELD.