

# Enhanced Cross-Domain Named Entity Recognition

Subjects: Computer Science, Artificial Intelligence

Contributor: Xiaoming Liu , Mengyuan Cao , Guan Yang , Jie Liu , Yang Liu , Hang Wang

Cross-domain named entity recognition (NER) is a crucial task in various practical applications, particularly when faced with the challenge of limited data availability in target domains. Existing methodologies primarily depend on feature representation or model parameter sharing mechanisms to enable the transfer of entity recognition capabilities across domains.

cross-domain named entity recognition

causally invariant knowledge

## 1. Introduction

Named entity recognition (NER) is a fundamental task in natural language processing (NLP), aimed at identifying entities with specific semantic meanings from text, such as names of people, locations, organizations, and institutions. It plays a significant role in knowledge graphs, information extraction, and text understanding [1][2][3]. In practical applications, the considerable variance in text genres and terminologies across diverse domains presents a substantial challenge, frequently leading to a scarcity of annotated data within specific target domains. Consequently, the adaptation of named entity recognition (NER) models for cross-domain scenarios, specifically cross-domain named entity recognition (CD-NER), has garnered significant research attention in recent years. This is particularly relevant in resource-constrained environments where the availability of labeled data is limited [4].

The current research on CD-NER has primarily focused on three distinct strategies. First, some researchers [5][6] have explored multi-task joint learning approaches, enhancing cross-domain entity recognition by simultaneously training models on both source and target domains to obtain refined feature representations across tasks. Second, a group of scholars [7][8] have proposed innovative model architectures aimed at understanding the complex semantic dynamics between domains, thus improving cross-domain performance. Third, another set of researchers [9][10] have leveraged pre-trained language models (PLMs) to develop models in data-rich domains, establishing robust source domain models. They have further improved cross-domain performance by transferring feature knowledge from the source domain to the target domain through fine-tuning and domain parameter sharing techniques. A notable example of current state-of-the-art CD-NER models is Cp-NER [10], which utilizes a frozen PLM while employing collaborative domain prefix adjustments to enhance the PLM, obtaining a significant improvement in cross-domain performance, as demonstrated by its superior performance on the CrossNER benchmark. However, it is important to note that existing methodologies often depend on inter-domain generalized knowledge for cross-domain transfer, which may inadvertently introduce out-of-domain knowledge that may not align with the specific task requirements during transfer. This observation underscores the need for a more

informed approach to CD-NER, a challenge researchers' proposed Causal Structure Alignment-based Cross-Domain Named Entity Recognition (CSA-NER) model aims to address.

To effectively harness domain-invariant knowledge, researchers' CSA-NER model employs a strategy that extracts causal invariant knowledge between domains. This is achieved by constraining domain-invariant knowledge through causal learning, ultimately enhancing the performance of the target domain. Specifically, **Figure 1** illustrates the acquisition of cross-domain causal invariant knowledge from similar syntactic structures in contexts and entities, where an ellipsis in the target domain denotes the omitted text "good way to". This process requires causal inference to learn causal relationships between entities and hidden syntactic structures. Subsequently, causal invariant knowledge hidden in syntactic structures and entities is extracted by aligning similar causal structures using GOT. This approach serves to alleviate the impact of out-of-domain knowledge on the task within the target domain. In various scientific domains, the concept of causal invariance has been extensively explored. For instance, Chevalley [11] designs a unified invariant learning framework that expertly utilizes distribution matching to enrich the acquisition of causal invariant knowledge, leading to a noteworthy enhancement in the model's performance. Chen [12] introduced causally inspired invariant graph learning to discern and leverage causally invariant knowledge pertaining to graph data. This is achieved by constructing causal graphs to represent shifts in the distribution of graphs, enabling the model to concentrate solely on the subgraphs that encapsulate the most pertinent information about the underlying causes of the labels. Furthermore, Arjovsky [13] argued that there is no causal relationship between the spurious correlation resulting from the transfer from the source domain to the target domain and the prediction target, and proposed an invariance risk minimization algorithm to mitigate the model's over-reliance on data bias by using causality tools to characterize the spurious correlation and invariance in the data.

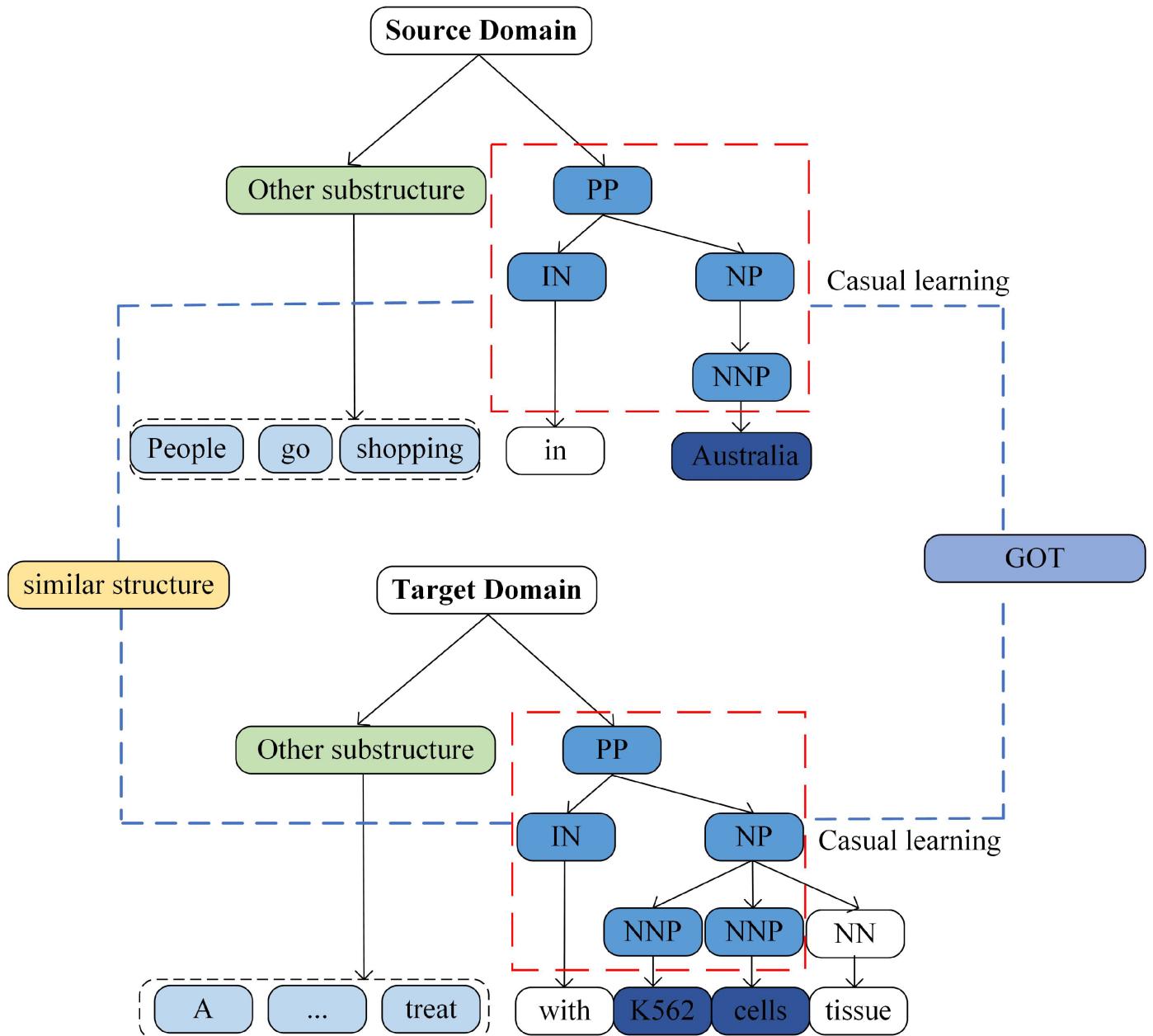


Figure 1. Cross-domain causal similarity structure.

## 2. Enhanced Cross-Domain Named Entity Recognition

### 2.1. Cross-Domain Named Entity Recognition

Cross-domain named entity recognition, which aims to utilize knowledge learned from resource-rich source domains to improve entity recognition in target domains, has received increasing research attention because it can alleviate the problems of data dependency and insufficient training data. Zhang [5] proposed a Multi-Cell Composition LSTM structure that models each entity type as a separate cell state, thus solving the problems of data annotation scarcity and entity ambiguity. These methods need to be trained on a large amount of source domain data to adapt to each domain, making them time consuming and inefficient. Hu [8] proposed a new auto-

regressive modeling framework that exploits semantic relationships between domains to migrate semantic features with the same label in the source domain to the target domain to jointly predict entity labels. Zheng [9] constructed a labeled graph by pre-training a language model and solved the cross-domain label semantic feature mismatch problem by dynamic graph matching. Chen [10] utilized frozen PLMs and conducted collaborative domain-prefix tuning to stimulate the potential of PLMs to handle NER tasks across various domains. In contrast, previous methods based on the transfer of semantic feature knowledge do not solve the negative transfer problem well and thus fail to produce more stable predictions by exploiting the causally invariant knowledge present in the source domain.

## 2.2. Few-Shot Named Entity Recognition

Few-shot named entity recognition (FS-NER) aims to identify new classes in resource-poor scenarios and also highlights good cross-domain capabilities. Fritzler [14] used prototype networks to achieve entity recognition for few-shot. Tong [15] proposed mining undefined classes to improve the robustness of the model and thus better adapt to few-shot learning. Cui [16] combined prompted learning templates and BART models for guided entity recognition to improve model performance and cross-domain applications. The authors of [17] do not even need a richly resourced source domain to accomplish small-sample learning without template tuning using prompted learning. The authors of [9] improve domain adaptation in low-resource domains by extracting semantic information of labels in resource-rich source domains. Although the above methods have been significantly improved in small-sample learning, they only improve the model domain adaptation [18] and generalization ability through few-shot training, but do not take into account the fact that the migrated causally invariant knowledge plays a key role in the downstream task.

## 2.3. Causal Invariant Learning

Causal invariant learning is a common solution for domain adaptation and domain generalization in solving cross-domain migration problems, where domain generalization is crucial for learning causal invariant knowledge in the domain. For example, Li [19] introduces a method called Distortion Invariant representation Learning (DIL) to enhance the generalization ability of deep neural networks in image restoration by addressing various types and degrees of image degradation from a causal perspective. Rojas-Carulla [20] proposed a transfer learning method based on causal modeling, which aims to find predictors that lead to invariant conditions through tasks with known underlying causal structure and tasks involving interventions on variables other than the target variable. Yang [21] proposed a causal self-encoder that learns causal representations by integrating them into a unified model using self-encoder and causal structure learning in the source domain, and utilizes this medium causal representation in the target domain for prediction. However, the method lacks the extraction and utilization of causal invariant knowledge.

## References

1. Ehrmann, M.; Hamdi, A.; Pontes, E.L. Named entity recognition and classification in historical documents: A survey. *ACM Comput. Surv.* 2023, 56, 1–47.
2. Ahmad, P.N.; Shah, A.M.; Lee, K. A Review on Electronic Health Record Text-Mining for Biomedical Name Entity Recognition in Healthcare Domain. *Healthcare* 2023, 11, 1268.
3. Tsai, C.-M. Stylometric Fake News Detection Based on Natural Language Processing Using Named Entity Recognition: In-Domain and Cross-Domain Analysis. *Electronics* 2023, 12, 3676.
4. Liu, Z.; Xu, Y.; Yu, T. Crossner: Evaluating cross-domain named entity recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 13452–13460.
5. Chen, J.; Zhang, Y. Multi-cell compositional LSTM for NER domain adaptation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 5906–5917.
6. Tan, Z.; Chen, Y.; Liang, Z. Named Entity Recognition for Few-Shot Power Dispatch Based on Multi-Task. *Electronics* 2023, 12, 3476.
7. Liu, Z.; Winata, G.I.; Xu, P. Coach: A Coarse-to-Fine Approach for Cross-domain Slot Filling. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 19–25.
8. Hu, J.; Zhao, H.; Guo, D.; Wan, X.; Chang, T. A label-aware autoregressive framework for cross-domain NER. In Proceedings of the Findings of the Association for Computational Linguistics: NAACL 2022, Seattle, WA, USA, 10–15 July 2022; pp. 2222–2232.
9. Zheng, J.; Chen, H.; Ma, Q. Cross-domain named entity recognition via graph matching. In Proceedings of the Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, 22–27 May 2022; pp. 2670–2680.
10. Chen, X.; Li, L.; Fei, Q.; Zhang, N.; Tan, C.; Jiang, Y.; Chen, H. One Model for All Domains: Collaborative Domain-Prefix Tuning for Cross-Domain NER. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23), Macao, China, 19–25 August 2023; Volume 2301, p. 10410.
11. Chevalley, M.; Bunne, C.; Krause, A.; Bauer, S. Invariant causal mechanisms through distribution matching. *arXiv* 2022, arXiv:2206.11646.
12. Chen, Y.; Zhang, Y.; Bian, Y.; Yang, H.; Ma, K.; Xie, B.; Liu, T.; Han, B.; Cheng, J. Learning causally invariant representations for out-of-distribution generalization on graphs. *Adv. Neural Inf. Process. Syst.* 2022, 35, 22131–22148.
13. Arjovsky, M.; Bottou, L.; Gulrajani, I.; Lopez-Paz, D. Invariant Risk Minimization. *arXiv* 2019, arXiv:1907.02893.

14. Fritzler, A.; Logacheva, V.; Kretov, M. Few-shot classification in named entity recognition task. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, Limassol, Cyprus, 8–12 April 2019; pp. 993–1000.
15. Tong, M.; Wang, S.; Xu, B.; Cao, Y.; Liu, M.; Hou, L.; Li, J. Learning from Miscellaneous Other-Class Words for Few-Shot Named Entity Recognition; Association for Computational Linguistics (ACL): Cedarville, OH, USA, 2021; pp. 6236–6247.
16. Cui, L.; Wu, Y.; Liu, J.; Yang, S.; Zhang, Y. Template-Based Named Entity Recognition Using BART. In Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Online, 1–6 August 2021; pp. 1835–1845.
17. Ma, R.; Zhou, X.; Gui, T.; Tan, Y.; Li, L.; Zhang, Q. Template-free Prompt Tuning for Few-shot NER. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 10–15 July 2022; Volume 2109, p. 13532.
18. Lu, W.; Wang, J.; Li, H.; Chen, Y.; Xie, X. Domain-invariant Feature Exploration for Domain Generalization. *Trans. Mach. Learn. Res.* 2022, 2835–8856.
19. Li, X.; Li, B.; Jin, X.; Lan, C.; Chen, Z. Learning Distortion Invariant Representation for Image Restoration from A Causality Perspective. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 1714–1724.
20. Rojas-Carulla, M.; Schölkopf, B.; Turner, R.; Peters, J. Invariant models for causal transfer learning. *J. Mach. Learn. Res.* 2018, 19, 1309–1342.
21. Yang, S.; Yu, K.; Cao, F.; Liu, L.; Wang, H.; Li, J. Learning causal representations for robust domain adaptation. *IEEE Trans. Knowl. Data Eng.* 2021, 35, 2750–2764.

Retrieved from <https://encyclopedia.pub/entry/history/show/121411>