

# Relating Relations: Meta-Relation Extraction from Online Health Forum Posts

Daniel Stickley

University of Brighton

UK

d.stickley@brighton.ac.uk

## Abstract

Relation extraction is a key task in knowledge extraction, and is commonly defined as the task of identifying relations that hold between entities in text. This thesis proposal addresses the specific task of identifying meta-relations, a higher order family of relations naturally construed as holding between other relations which includes temporal, comparative, and causal relations. More specifically, we aim to develop theoretical underpinnings and practical solutions for the challenges of (1) incorporating meta-relations into conceptualisations and annotation schemes for (lower-order) relations and named entities, (2) obtaining annotations for them with tolerable cognitive load on annotators, (3) creating models capable of reliably extracting meta-relations, and related to that (4) addressing the limited-data problem exacerbated by the introduction of meta-relations into the learning task. We explore recent works in relation extraction and discuss our plans to formally conceptualise meta-relations for the domain of user-generated health texts, and create a new dataset, annotation scheme and models for meta-relation extraction.

## 1 Introduction

The vast amounts of information now stored digitally in various forms e.g. social media and online forum posts, electronic health records, academic papers, etc., represent large amounts of unstructured data that contain useful information for a variety of purposes. Due to the scale of this data, automatic Knowledge Extraction (KE) holds the promise of allowing the automatic extraction of machine interpretable knowledge from these unstructured sources. Relation Extraction (RE) identifies relations that hold (typically) between entities in text; it is often an important and challenging sub-task in KE - and is the primary focus of our thesis.

Although there has been much work on RE and KE (Han et al., 2020; Manchanda and Phansalkar, 2019; Ma et al., 2019; Belz et al., 2019) there still remain considerable problems, including the high cost of creating high-quality annotated data, adequate conceptualisation of higher-order relationships that link events and/or cross sentence boundaries, and extraction from messy user-generated text.

In this research proposal we explore recent work in the field of RE, and identify some of the limitations that current research still faces. Next, we outline the specific problems that our research will address, before presenting a first outline of our approach based on meta-relations. We demonstrate the role of meta-relations with a worked example and outline how they will allow us to extract more richly structured data that more adequately captures the relationships between different types of information, as well as lowering the cognitive load on annotators. We then go on to discuss the data and methods we will use in our research, how these will progress our research, address our objectives, and conclude with a summary of the expected contributions of our thesis.

## 2 Research Context

This thesis will focus on relation extraction from natural language - with a particular focus on user generated text such as in online forum and social media posts. The task of relation extraction is typically approached by identifying named entities in the text, then identifying relations that hold between these entities. RE is often a very important component of KE systems, where these entities and relations are then used to populate a knowledge graph. This section will discuss recent RE approaches in order to establish the context our research will belong to.

Relation extraction is a task that has seen much attention over the years, as Han et al.’s survey (2020) shows. Previous work has used statistical approaches, feature based methods, and graph-based approaches, but more recently neural machine learning models that use word embeddings as inputs have become the norm. Looking at shared tasks in RE, we can see particular language models and neural architectures that have been widely adopted and adapted for RE, such as the BERT language model (Devlin et al., 2018). A widely used architecture for RE is the Recurrent Neural Network (RNN), such as the Bi-LSTM (Huang et al., 2015) - a very popular model that has seen wide adoption and further developments such as using attention to enhance performance (Geng et al., 2020). BERT and Bi-LSTM based methods are the top performing models in various shared tasks, such as RuREBus (Ivanin et al., 2020), CCKS (Wang et al., 2019a), FewRel2.0 (Gao et al., 2019), and BioNLP 2019’s AGAC and BB tasks (Wang et al., 2019b; Bossy et al., 2019). Han et al. (2020) also identify various challenges still faced by RE systems: current RE models often work in a simplified setting, but struggle with more complex tasks such as inter-sentence relations, and in few shot learning and open domain scenarios. Despite recent works attempting to overcome these challenges (Christopoulou et al., 2019; Gao et al., 2019; Cui et al., 2018; Wu et al., 2019) they can still be considered largely unsolved.

One challenge commonly faced in RE research is acquiring high quality labelled data to train and evaluate supervised models on. State-of-the-art models require a large amount of adequately varied, labelled data in order to be successfully trained, however RE is usually domain specific and so for each new task it can be difficult to acquire or create high quality gold standard training data. This is one of the main motivations for few-shot and cross-domain RE (Gao et al., 2019), as they both focus on learning with small amounts of labelled data. Belz et al. (2019) discuss the complexities that must be considered when creating labelled data for RE, and the difficulties that are encountered. They explore issues such as the high cognitive load that annotators can encounter with complex annotation schemes, and the resource intensive nature of annotating data with experts. In addition to few-shot learning and cross domain adaptation, there have been other methods attempting to mitigate the lack

of data, such as initially training on coarse-grained labels (which are much quicker and cheaper to produce), then adopting a label-as-you-go approach whereby users create fine-grained labels as the system is used (Manchanda and Phansalkar, 2019).

Distant supervision has been used to increase usable data, and is applied to RE by using the basic assumption *”If two entities participate in a relation, any sentence that contains those two entities might express that relation”* (Mintz et al., 2009) to create weakly labelled data. This understandably leads to noisy data, with false positives and incomplete labels, however it has still shown to work reasonably well in practice (Smirnova and Cudré-Mauroux, 2018). Several approaches have attempted to address the shortcomings of distant supervision methods in RE, such as revising the basic assumption (that all sentences containing two entities with a known relation express that relation) to reduce the noise of the data (Riedel et al., 2010), creating undirected graphical models for distant supervision (Hoffmann et al., 2011; Surdeanu et al., 2012), allowing more than one relation to be predicted for two entities.

These models have since been further developed to introduce negative examples (of unrelated entity pairs) with an additional layer that denotes whether a relation holds for given entities (Min et al., 2013), leading to an increase in precision by between 1-4% (Smirnova and Cudré-Mauroux, 2018). Ritter et al. (2013) add a variable when aligning the text with the knowledge base, denoting whether a relation fact is present in the text. The model penalises disagreement between this variable and the variable that indicates whether the relation is present in the knowledge base. This allows us to encode certain intuitions about how likely certain relations are to be present or absent from the text, helping to reduce the issue of missing labels. Despite the continued development of distantly supervised approaches, noisy and incomplete data remains a problem that limits the performance of such systems (Smirnova and Cudré-Mauroux, 2018).

Current RE methods work best for relations that can be construed as holding between two named entities in text, e.g. [Margaret Hamilton, *worked\_at*, NASA]. There has been less progress on relations that are inherently higher order, such as temporal relations which hold between events, e.g. *BEFORE*[[Margaret Hamilton, *worked\_at*, NASA], [Margaret Hamilton, *founded*, Hamilton

Technologies]], or causal relations, which can hold between events, entities, and relations. Both temporal and causal RE are typically treated as separate tasks from general RE, which means they cannot take advantage of the mutually constraining nature of the general RE. A recent survey of temporal relation extraction (Alfattni et al., 2020) shows that extracting and classifying time expressions e.g. ‘last week’ and events e.g. ‘pain has increased’ are close to solved problems, but there still is much room for improvement in extracting relations that involve such time expressions. Causal relation extraction has seen much attention, with a recent focus on deep neural network approaches, however a recent survey shows that causal RE is also an open problem (Yang et al., 2021). Developing RE methods to better extract higher-order relations such as temporal and causal relations is still an important challenge for research to address.

From our review of the relation extraction literature, of which selected highlights were explored in this section, we have identified several key areas to address in our thesis. We will focus on the task of extracting a higher-order family of relations - specifically temporal, comparative, and causal - naturally construed as holding between other relations. We will develop theoretical underpinnings of and practical solutions for the challenges of (1) incorporating meta-relations into RE task conceptualisations, and annotation schemes for lower-order relations and named entities, (2) obtaining annotations for them with tolerable cognitive load on annotators, (3) creating models capable of reliably extracting meta-relations, and related to that (4) how to overcome the limited-data problem exacerbated by the introduction of meta-relations into the learning task.

### 3 Meta-Relations

The focus of the proposed thesis is a class of relations that are inherently *meta*, because they hold between (structured) relations, rather than monadic entities. This class includes temporal relations such as BEFORE and AFTER, comparative relations such as LONGER and SHORTER, and causal relations. As a starting point, we propose to incorporate meta-relations into the RE processes as follows, aiming to explore both sequential and joint modelling of the subtasks:

1. Named entity recognition (NER) is used to identify relevant entities in the text;

2. Relation extraction is performed, linking pairs of entities with a relation;
3. Meta-relation extraction is performed on the lower-level relations extracted in the previous step, linking pairs of relations with a meta-relation.

Let us break this down with a worked example - taken from the dataset presented by Belz et al. (2019), and using their annotation scheme:

I took Wellbutrin for a short period of time but it make me sick so I was given Zoloft

This sentence contains information that the user took - but subsequently stopped taking<sup>1</sup> - Wellbutrin, it made them feel sick, and they were then given Zoloft to take. As a human reader the temporal order in this information is easy to discern, however in order to extract this information in a machine readable format e.g. as relational triples, there are several steps that must be completed. First, named entities and other lowest-level units of information must be identified; in our case these would be drugs {*Wellbutrin, it, Zoloft*}, drug modification actions {*took, was given*}, and drug effects {*make me sick*}. Next there may be an entity linking step where the identified entities are linked to nodes in a knowledge graph. This is followed by the RE step, where the relations between pairs of entities will be identified: modification/drug relations will denote the user started or stopped taking a drug, and drug/effect relations will show effects associated with a drug.

With typical RE, this is where the process would end and the (unordered) extracted information would be: Wellbutrin started being taken, Wellbutrin stopped being taken, Wellbutrin caused nausea, and Zoloft started being taken. This information is useful but lacks important information such as the order of events and causality. Each relation is treated separately, and from this relation information we cannot infer the order of events as these same relations could be present in another order of events. e.g. if the user was taking Zoloft and Wellbutrin together, then they stopped taking Wellbutrin after feeling sick. This demonstrates that the temporal order of events is a crucial aspect to properly understanding the content in a sentence/document.

<sup>1</sup>We can make this assumption as these are both antidepressant medications, even though this may not be immediately clear solely from this text span.

It is crucial in a number of real world tasks such as automated Yellow Card filling,<sup>2</sup> for reporting unknown adverse side effects from drugs, where the temporal order of events - which drugs a person was taking, at what time, and in what order - must be reported as all of these factors may play a role in the adverse reaction (Belz et al., 2019).

Figure 1 shows how temporal meta-relations could be integrated with other extracted relations. In doing this we extract an unambiguous temporal order that shows Wellbutrin was taken, then the user felt sick, then they stopped taking Wellbutrin and started taking Zoloft. Temporal meta-relations are only one example, there are other types of meta-relations that could provide additional key pieces of information such as causal meta-relations. In our example in figure 1, we can see two event relations from Blez et al.'s (2019) annotation scheme: DRUG\_MODIF[TYPE=START] and DRUG\_EFFECT. Here we could use a causal meta-relation to show explicitly that starting to take the labelled drug (in this case Wellbutrin) is what caused the labelled effect (nausea).

## 4 Problem Addressed, Data, Methods

### 4.1 Problem addressed

Firstly, we will create a formal conceptualisation of the meta-relation extraction task and a new annotation scheme that incorporates meta-relations. We will use this to create a new large scale annotated dataset. As demonstrated by Belz et al. (2019), task conceptualisation is a complex and important task itself, and a crucial component in creating an annotation scheme and dataset, and in defining learning tasks. Therefore, creating a formal conceptualisation for meta-relation extraction will be an important first step in our research.

Meta-relations allow us to create methods for higher-order relation extraction, and to create more complete RE systems that develop a better understanding of the text, by extracting more richly structured information that more adequately captures the relationships between different types of information, as well as lowering the cognitive load on annotators. To achieve this, it will be important to address the different types of meta-relations and information we wish to extract when conceptualising the meta-relation task in the first instance.

<sup>2</sup>For information about the Yellow Card Scheme see <https://yellowcard.mhra.gov.uk>

So far our objectives will further the creation and performance of directly supervised methods using gold standard labelled data, however in order to make our novel conceptualisation of meta-relation extraction more generally applicable we will also develop less supervised methods. Approaches using distantly supervised and data augmentation methods allow the extension of existing datasets with meta-relations, whereas other methods such as bootstrapping and transfer learning enable the adoption of meta-relation extraction for neighbouring domains when data is limited. Utilising methods such as these will allow new meta-relation extraction models to be created with a reduced need for the resource intensive process of creating new data and models from scratch.

### 4.2 Data

We plan to use two existing data resources which we will adapt for our purposes. Belz et al. (2019) have collected 148,575 posts from online drug forums, and so far have annotated 2,000 posts for RE with the first phase of their annotation scheme described in their paper. We plan to adapt this annotation scheme, extending it to incorporate typed relations and meta-relations. The development of this annotation scheme will be a nontrivial task as it is important to properly consider how best to annotate the data to produce high quality labels whilst minimising cognitive load on workers during the annotation process.

Belz et al. describe the difficulty of maintaining a manageable cognitive strain on annotators. The goal of annotating the data is to produce labels that convey a complex level of information for the machine learning model to learn from, however if the labelling task is too difficult or convoluted it can lead to issues with slow turnaround time and low inter-annotator agreement, as well as annotators abandoning their work. Because of these potential issues, it is important that we consider the cognitive load annotators will experience when creating the annotation pipeline, for example by splitting annotation up into several simpler phases, and facilitating workshops to ensure the workers fully understand the annotation task and are able to use the required labelling tools.

The second resource we have identified that will be useful in this work is the MedNorm corpus and embeddings, created for cross-terminology medical concept normalisation (Belousov et al., 2019).

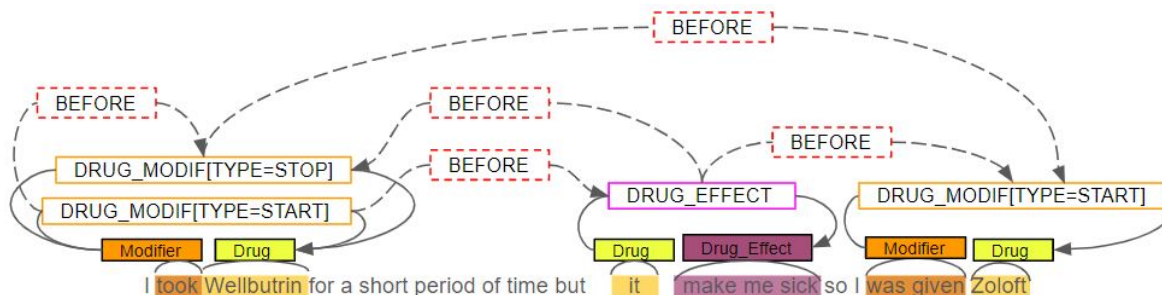


Figure 1: An example sentence taken from the dataset presented by Belz et al. (2019), annotated with meta-relations. Filled coloured boxes represent entities and lowest-level units of information, solid boxes and arrows represent lower order relations, dashed boxes with dashed arrows represent meta-relations.

This corpus is made up of five other datasets for medical concept normalisation, merged and annotated with a more uniform approach using two well known biomedical vocabularies and lexicons - SNOMED-CT and MedDRA (Stearns et al., 2001; Brown et al., 1999). This corpus will prove useful in developing RE systems for the medical domain, allowing us to leverage biomedical vocabularies and develop models that extract useful data for downstream tasks such as knowledge graph population.

### 4.3 Methods

As stated earlier, our research at the start aims to create a formal conceptualisation of meta-relation extraction, followed by creating an annotation scheme and dataset. Working from what we have already laid out in this paper, we will carefully consider what information can be represented by meta-relations. So far we have suggested temporal, causal, and comparative meta-relations, as they hold between relations, however more thorough exploration of the problem space is necessary to understand how to conceptualise the task of meta-relation extraction, and how best to create an annotation scheme incorporating these higher-order relations. To accomplish this, we will perform an in-depth investigation of how relations that we consider meta-relations have been treated in the literature and in existing annotation schemes. We will also consider a type system for lower-level (traditional) relations - in addition to named entities - in our conceptualisation. A type system would allow us to constrain the types of named entities a relation could apply to, and relations that meta-relations can apply to, meaning that not every entity/relation will need to be considered for every type of relation/meta-relation. This would also help

in the development of distantly supervised methods for meta-relation extraction, similar to Xu et al. (2013) where they incorporate type constraints into their automatic labelling process.

We will use the annotation scheme presented by Belz et al. (2019) as the basis for our annotation scheme, revising as necessary to incorporate meta-relations, typed relations, and any additional entity labels in order to fit our conceptualisation. As mentioned earlier, asking workers to fully label the data in one step would be too cognitively demanding, so we plan to split the annotation process into several smaller phases. The presence of meta-relations will help when breaking the annotation down into smaller phases, as entities and lower level relations can be annotated separately, and should be relatively straightforward to label as they will represent the simplest units of information. Then, more complex higher-order information can be labeled as meta-relations in subsequent phases. Additionally, we will ensure annotators fully understand the labelling task by running an instructive course before they begin the labelling process. By splitting the process up into smaller tasks - paired with the instructive course - the annotators job will be more straightforward and less challenging, which should lead to both faster annotating and fewer mistakes made by annotators.

Our newly annotated dataset will then allow us to perform a series of experiments to explore the task of meta-RE. First, we will create a baseline approach for meta-RE, this will demonstrate how successful current state-of-the-art models are at extracting meta-relations, and show the accuracy with which different types of mental relations can be extracted using current methods. As we have described in Section 3, the task of meta-relation extraction can be framed as a pipelined extension of

relation extraction, passing through the output of a traditional RE system (entities with corresponding relations) to another RE model, with some minor modification in order to perform relation extraction over the relations (as opposed to the entities). In this model we will follow the steps outlined in section 3: Use an NER model to identify the entities, before passing them into an RE model, and finally pass the entities and their corresponding relations through a modified RE model to extract meta-relations. This will be a very simple baseline model with little modification from the base RE model, meaning that we should be able to use most well performing RE models that take entities as input. We currently plan to use REDN (Li and Tian, 2020) due to its high performance on the popular shared task SemEval-2010 Task 8 (Hendrickx et al., 2009). The modifications we intend to make to the model for the final step in our baseline approach, are to treat extracted relations as entities. If we were to only input the relations (and text) in this step, the model will not be able to predict meta-relations that hold between relations and entities - instead only predicting meta-relations that hold between multiple relations; to avoid this, we will pass both the relations and entities through together, treating both as 'entities'.

We will then go on to explore alternative approaches and models developed specifically for meta-relations with the aim of achieving better results both quantitatively in the meta-relation extraction task, but also qualitatively in the quality of information the model can extract (in the form of entities, relations, and meta-relations). One alternative approach we will explore will be joint extraction of entities, relations, and meta-relations. It has been reported that jointly extracting entities and relations can lead to better overall performance as this reduces error propagation and allows more information to be leveraged from the text for both entity recognition and relation extraction (Cohen et al., 2020). In our experiments we will investigate how well meta-relations could be extracted jointly with entities and lower order relations; there are many RE models that jointly extract entities and relations, and one that we have identified for our experiments is TPLinker (Wang et al., 2020). The handshake tagging system used in the TPLinker model could be modified to also incorporate meta-relations which would enable the joint extraction of entities, relations and meta-relations.

## 5 Conclusion

In this paper we have provided the research context that our work will contribute to, identified various limitations in current work, including data collection/creation, cross domain adaptation, higher-order relation extraction, and distantly supervised relation extraction. We then proposed using meta-relations to extract more knowledge from text and discussed our plans to conceptualise the task of meta-relation extraction, and to incorporate meta-relations in an annotation scheme and dataset we will create. We also identified challenges we anticipate in addition to the data and methods we plan to use, and how we will utilise these methods in our work. Focusing on extraction of temporal, comparative, and causal relations in user-generated texts in the health domain, the expected contributions of our thesis are as follows:

- Formal conceptualisation of temporal, comparative, and causal meta-relation extraction.
- A new annotation scheme incorporating typed relations and meta-relations as well as lower-level relations.
- A new dataset for meta-relation extraction created with the above annotation scheme.
- A baseline pipelined approach and trained model for meta-relation extraction for drug effect and nonadherence information from user generated text.
- An alternative model for joint entity, relation and, meta-relation extraction for the above user generated text.
- Methods to address the dataset-creation problem capable of resulting in high quality relation extraction models trained on them, where access to data is limited.

## Acknowledgements

We thank the anonymous reviewers for their very helpful comments.

## References

- Ghada Alfattni, Niels Peek, and Goran Nenadic. 2020. Extraction of temporal relations from clinical free text: A systematic review of current approaches. *Journal of Biomedical Informatics*, page 103488.

- Maksim Belousov, William G Dixon, and Goran Nenadic. 2019. Mednorm: A corpus and embeddings for cross-terminology medical concept normalisation. In *Proceedings of the Fourth Social Media Mining for Health Applications (# SMM4H) Workshop & Shared Task*, pages 31–39.
- Anya Belz, Richard Hoile, Elizabeth Ford, and Azam Mullick. 2019. Conceptualisation and annotation of drug nonadherence information for knowledge extraction from patient-generated texts. In *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pages 202–211.
- Robert Bossy, Louise Deléger, Estelle Chaix, Mouhamadou Ba, and Claire Nédellec. 2019. Bacteria biotope at bionlp open shared tasks 2019. In *Proceedings of The 5th Workshop on BioNLP Open Shared Tasks*, pages 121–131.
- Elliot G Brown, Louise Wood, and Sue Wood. 1999. The medical dictionary for regulatory activities (meddra). *Drug safety*, 20(2):109–117.
- Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. A walk-based model on entity graphs for relation extraction. *arXiv preprint arXiv:1902.07023*.
- Amir DN Cohen, Shachar Rosenman, and Yoav Goldberg. 2020. Relation extraction as two-way span prediction. *arXiv preprint arXiv:2010.04829*.
- Lei Cui, Furu Wei, and Ming Zhou. 2018. Neural open information extraction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 407–413, Melbourne, Australia. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Tianyu Gao, Xu Han, Hao Zhu, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. 2019. Fewrel 2.0: Towards more challenging few-shot relation classification. *arXiv preprint arXiv:1910.07124*.
- ZhiQiang Geng, GuoFei Chen, YongMing Han, Gang Lu, and Fang Li. 2020. Semantic relation extraction using sequential and tree-structured lstm with attention. *Information Sciences*, 509:183–192.
- Xu Han, Tianyu Gao, Yankai Lin, Hao Peng, Yao-liang Yang, Chaojun Xiao, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. 2020. More data, more relations, more context and more openness: A review and outlook for relation extraction. *arXiv preprint arXiv:2004.03186*.
- Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Sebastian Padó, Marco Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 2009. Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions*, SEW '09, page 94–99, USA. Association for Computational Linguistics.
- Raphael Hoffmann, Congle Zhang, Xiao Ling, Luke Zettlemoyer, and Daniel S Weld. 2011. Knowledge-based weak supervision for information extraction of overlapping relations. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*, pages 541–550.
- Zhiheng Huang, Wei Xu, and Kai Yu. 2015. *Bidirectional lstm-crf models for sequence tagging*.
- V Ivanin, E Artemova, T Batura, V Ivanov, V Sarkisyan, E Tutubalina, and I Smurov. 2020. Rurebus-2020 shared task: Russian relation extraction for business. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog”*.
- Cheng Li and Ye Tian. 2020. Downstream model design of pre-trained language model for relation extraction task. *arXiv preprint arXiv:2004.03786*.
- Yunshan Ma, Lizi Liao, and Tat-Seng Chua. 2019. Automatic fashion knowledge extraction from social media. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 2223–2224.
- Sanjeev Manchanda and Ajeet Phansalkar. 2019. Automating knowledge extraction from unstructured knowledge articles. In *2019 IEEE Pune Section International Conference (PuneCon)*, pages 1–6. IEEE.
- Bonan Min, Ralph Grishman, Li Wan, Chang Wang, and David Gondek. 2013. Distant supervision for relation extraction with an incomplete knowledge base. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 777–782.
- Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 1003–1011.
- Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. Modeling relations and their mentions without labeled text. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 148–163. Springer.
- Alan Ritter, Luke Zettlemoyer, Mausam, and Oren Etzioni. 2013. Modeling missing data in distant supervision for information extraction. *Transactions of the Association for Computational Linguistics*, 1:367–378.

- Alisa Smirnova and Philippe Cudré-Mauroux. 2018. Relation extraction using distant supervision: A survey. *ACM Computing Surveys (CSUR)*, 51(5):1–35.
- Michael Q Stearns, Colin Price, Kent A Spackman, and Amy Y Wang. 2001. Snomed clinical terms: overview of the development process and project status. In *Proceedings of the AMIA Symposium*, page 662. American Medical Informatics Association.
- Mihai Surdeanu, Julie Tibshirani, Ramesh Nallapati, and Christopher D Manning. 2012. Multi-instance multi-label learning for relation extraction. In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, pages 455–465.
- Haitao Wang, Zhengqiu He, Tong Zhu, Hao Shao, Wenliang Chen, and Min Zhang. 2019a. Ccks 2019 shared task on inter-personal relationship extraction. *arXiv preprint arXiv:1908.11337*.
- Yucheng Wang, Bowen Yu, Yueyang Zhang, Tingwen Liu, Hongsong Zhu, and Limin Sun. 2020. Tplinker: Single-stage joint extraction of entities and relations through token pair linking. *arXiv preprint arXiv:2010.13415*.
- Yuxing Wang, Kaiyin Zhou, Mina Gachloo, and Jingbo Xia. 2019b. An overview of the active gene annotation corpus and the bionlp ost 2019 agac track tasks. In *Proceedings of The 5th Workshop on BioNLP Open Shared Tasks*, pages 62–71.
- Ruidong Wu, Yuan Yao, Xu Han, Ruobing Xie, Zhiyuan Liu, Fen Lin, Leyu Lin, and Maosong Sun. 2019. Open relation extraction: Relational knowledge transfer from supervised data to unsupervised data. In *Proceedings of EMNLP-IJCNLP 2019*, pages 219–228, Hong Kong, China. Association for Computational Linguistics.
- Wei Xu, Raphael Hoffmann, Le Zhao, and Ralph Grishman. 2013. Filling knowledge base gaps for distant supervision of relation extraction. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 665–670.
- Jie Yang, Soyeon Caren Han, and Josiah Poon. 2021. A survey on extraction of causal relations from natural language text. *arXiv preprint arXiv:2101.06426*.