New Frontiers of Information Extraction

Muhao Chen[1], **Lifu Huang**[2], **Manling Li**[3], **Ben Zhou**[4], **Heng Ji**[3,5], **Dan Roth**[4,6]
[1]Department of Computer Science, USC
[2]Department of Computer Science, Virginia Tech
[3]Department of Computer Science, UIUC
[4]Department of Computer and Information Science, UPenn
[5]Alexa AI & [6]AWS AI Labs

Abstract

This tutorial targets researchers and practitioners who are interested in AI and ML technologies for structural information extraction (IE) from unstructured textual sources. In particular, this tutorial will provide audience with a systematic introduction to recent advances in IE, by addressing several important research questions. These questions include (i) how to develop a robust IE system from a small amount of noisy training data, while ensuring the reliability of its prediction? (ii) how to foster the generalizability of IE through enhancing the system's cross-lingual, cross-domain, cross-task and cross-modal transferability? (iii) how to support extracting structural information with extremely fine-grained and diverse labels? (iv) how to further improve IE by leveraging indirect supervision from other NLP tasks, such as Natural Language Generation (NLG), Natural Language Inference (NLI), Question Answering (QA) or summarization, and pre-trained language models? (v) how to acquire knowledge to guide inference in IE systems? We will discuss several lines of frontier research that tackle those challenges, and will conclude the tutorial by outlining directions for further investigation.

1 Introduction

Information extraction (IE) is the process of automatically extracting structural information from unstructured or semi-structured data. It provides the essential support for natural language understanding by recognizing and resolving the concepts, entities, events described in text, and inferring the relations among them. In various application domains, IE automates the costly acquisition process of domain-specific knowledge representations that have been the backbone of any knowledge-driven AI systems. For example, automated knowledge base construction has relied on technologies for entity-centric IE (Carlson et al., 2010; Lehmann et al., 2015). Extraction of events and event chains assists machines with narrative prediction (Zhang et al., 2021b; Chaturvedi et al., 2017) and summarization tasks (Liu et al., 2018; Chen et al., 2019b). Medical IE also benefits important but expensive clinical tasks such as drug discovery and repurposing (Sosa et al., 2019; Munkhdalai et al., 2018). Despite the importance, frontier research in IE still faces several key challenges. The first challenge is that existing dominant methods using language modeling representation cannot sufficiently capture the essential knowledge and structures required for IE tasks. The second challenge is on the development of extraction models for fine-grained information with less supervision, considering that obtaining structural annotation on unlabeled data has been very costly. The third challenge is to extend the reliability and generalizability of IE systems in real-world scenarios, where data sources often contain incorrect, invalid or unrecognizable inputs, as well as inputs containing unseen labels and mixture of modalities. By tackling those critical challenges, recent literature is leading to transformative advancement in principles and methodologies of IE system development. We believe it is necessary to present a timely tutorial to comprehensively summarize the new frontiers in IE research and point out the emerging challenges that deserve further investigation.

In this tutorial, we will systematically review several lines of frontier research on developing robust, reliable and adaptive learning systems for extracting rich structured information. Beyond introducing robust learning and inference methods for unsupervised denoising, constraint capture and novelty detection, we will discuss recent approaches for leveraging indirect supervision from natural language inference and generation tasks to improve IE. We will also review recent minimally supervised methods for training IE models with distant supervision from linguistic patterns, corpus statistics or language modeling objectives. In addition, we will

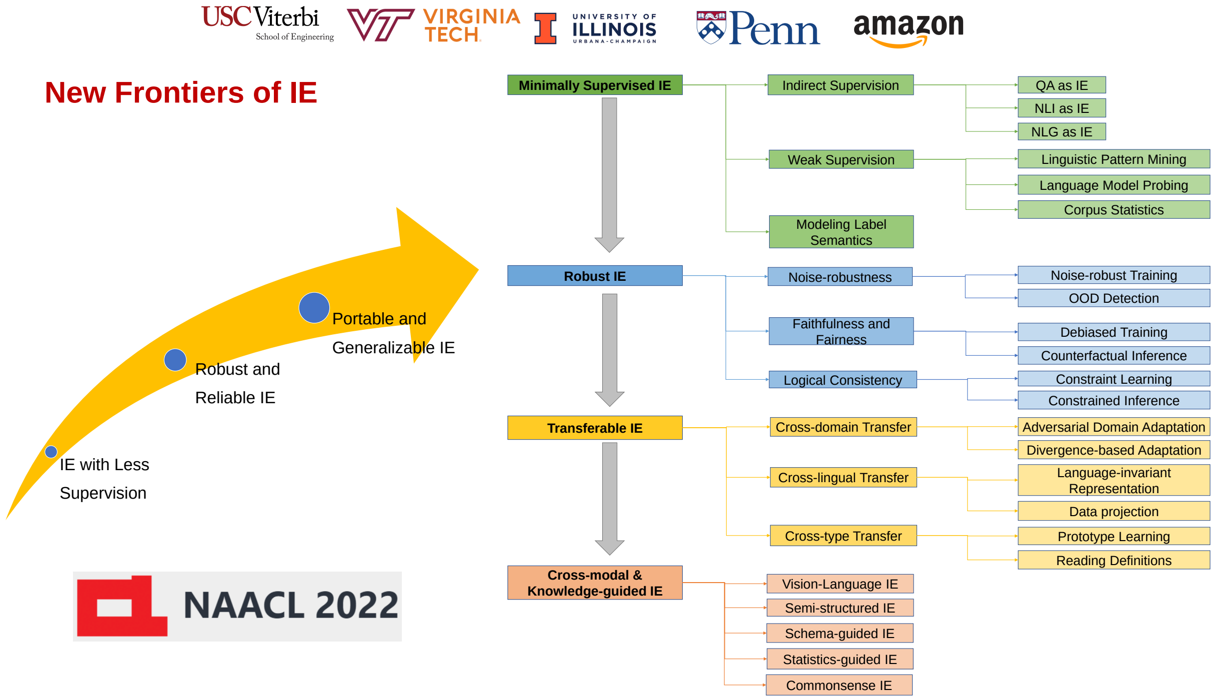

Figure 1: A roadmap for new frontiers of information extraction and a graphical abstract of this tutorial.

illustrate how a model trained on a close domain can be reliably adapted to produce extraction from data sources in different domains, languages and modalities, or acquiring global knowledge (e.g., event schemas) to guide the extraction on a highly diverse open label space. Participants will learn about recent trends and emerging challenges in this topic, representative tools and learning resources to obtain ready-to-use models, and how related technologies benefit end-user NLP applications. A graphical abstract of this tutorial is provided as Fig. 1, which serves as our roadmap for new frontiers of information extraction.

2 Outline of Tutorial Content

This **half-day** tutorial presents a systematic overview of recent advancement in IE technologies. We will begin motivating this topic with a selection of real-world applications and emerging challenges of IE. Then, we will introduce robust learning methods and inference methods to tackle noisy supervision, prediction inconsistency and out-of-distribution (OOD) inputs. We will also discuss about indirect supervision and minimal supervision methods that further improves IE model development under limited learning resources. Based on the robust IE systems developed in close-domain settings, we will explain how transfer learning technologies can adaptively extend the utility of the systems across domains, languages and tasks, and how complementary information can be extracted from data modalities other than human language. Moreover, we will exemplify the use of aforementioned technologies in various end-user NLP applications such as misinformation detection and scientific discovery, and will outline emerging research challenges that may catalyze further investigation on developing reliable and adaptive learning systems for IE. The detailed contents are outlined below.

2.1 Background and Motivation [20min]

We will define the main research problem and motivate the topic by presenting several real-world NLP and knowledge-driven AI applications of IE technologies, as well as several key challenges that are at the core of frontier research in this area.

2.2 Robust Learning and Inference for IE [35min]

We will introduce methodologies that enhance the robustness of learning systems for IE in both their learning and inference phases. Those methodologies involve self-supervised denoising techniques for training noise-robust IE models based on co-regularized knowledge distillation (Zhou and Chen, 2021; Liang et al., 2021), label re-weighting (Wang et al., 2019b) and label smoothing (Lukasik et al., 2020). Besides, we will also discuss about unsupervised techniques for out-of-distribution (OOD) detection (Zhou et al., 2021b; Hendrycks et al., 2020), prediction with abstention (Dhamija et al., 2018; Hendrycks et al., 2018) and novelty class detection (Perera and Patel, 2019) that seek to help the IE model identify invalid inputs or inputs with semantic shifts during its inference phase. Specifically, to demonstrate how models can ensure the global consistency of the extraction, we will cover constraint learning methods that automatically capture logical constraints among relations (Wang et al., 2021a, 2022c; Pan et al., 2020), and techniques to enforce the constraints in inference (Wang et al., 2020; Li et al., 2019a; Han et al., 2019; Lin et al., 2020). To assess if the systems give faithful extracts, we will also talk about the spurious correlation problems of current IE models and how to address them with counterfactual analysis (Wang et al., 2022b; Qian et al., 2021).

2.3 Minimally and Indirectly Supervised IE [35min]

We will introduce effective approaches that use alternative supervision sources for IE, that is, to use supervision signals from related tasks to make up for the lack of quantity and comprehensiveness in IE-specific training data. This includes indirect supervision sources such as question answering and reading comprehension (Wu et al., 2020; Lyu et al., 2021; Levy et al., 2017; Li et al., 2019b; Du and Cardie, 2020), natural language inference (Li et al., 2022a; Yin et al., 2020) and generation (Lu et al., 2021; Li et al., 2021b). We will also cover the use of weak supervision sources such as structural texts (e.g., Wikipedia) (Ji et al., 2017; Zhou et al., 2018) and global biases (Ning et al., 2018b). With the breakthrough of large-scale pretrained language models (Devlin et al., 2019; Li et al., 2022c), methodologies have been proposed to explore the language model objective as indirect supervision for IE. To this end, we will cover methods includes direct probing (Feldman et al., 2019; Zhang et al., 2020c), and more recently, pretraining with distant signals acquired from linguistic patterns (Zhou et al., 2020, 2021a).

2.4 Transferablity of IE Systems [35min]

One important challenge of developing IE systems lies in the limited coverage of predefined schemas (e.g., predefined types of entities, relations or events) and the heavy reliance on human annotations. When moving to new types, domains or languages, we have to start from scratch by creating annotations and re-training the extraction models. In this part of tutorial, we will cover the recent advances in improving the transferability of IE, including (1) cross-lingual transfer by leveraging adversarial training (Chen et al., 2019a; Huang et al., 2019; Zhou et al., 2019), language-invariant representations (Huang et al., 2018a; Subburathinam et al., 2019) and resources (Tsai et al., 2016; Pan et al., 2017), pre-trained multilingual language models (Wu and Dredze, 2019; Conneau et al., 2020) as well as data projection (Ni et al., 2017; Yarmohammadi et al., 2021), (2) cross-type transfer including zero-shot and few-shot IE by learning prototypes (Huang et al., 2018b; Chan et al., 2019; Huang and Ji, 2020), reading the definitions (Chen et al., 2020b; Logeswaran et al., 2019; Obeidat et al., 2019; Yu et al., 2022; Wang et al., 2022a), answering questions (Levy et al., 2017; Liu et al., 2020; Lyu et al., 2021), and (3) transfer across different benchmark datasets (Xia and Van Durme, 2021; Wang et al., 2021b). Finally, we will also discuss the progress on life-long learning for IE (Wang et al., 2019a; Cao et al., 2020; Yu et al., 2021; Liu et al., 2022) to enable knowledge transfer across incrementally updated models.

2.5 Cross-modal IE [20min]

Cross-modal IE aims to extract structured knowledge from multiple modalities, including unstructured and semi-structured text, images, videos, tables, etc. We will start from visual event and argument extraction from images (Yatskar et al., 2016; Gkioxari et al., 2018; Pratt et al., 2020; Zareian et al., 2020; Li et al., 2022b) and videos (Gu et al., 2018; Sadhu et al., 2021; Chen et al., 2021a). To extract multimedia events, the key challenge is to identify the cross-modal coreference and linking (Deng et al., 2018; Akbari et al., 2019; Zeng et al., 2019) and represent both text and visual knowledge in a common semantic space (Li et al., 2020a; Chen et al., 2021b; Zhang et al., 2021a; Li et al., 2022b). We will also introduce the information extraction from semi-structured data (Katti et al., 2018; Qian et al., 2019) and tabular data (Herzig et al., 2020).

2.6 Knowledge-guided IE [15min]

Global knowledge representation induced from large-scale corpora can guide the inference about the complicated connections between knowledge elements and help fix the extraction errors. We will introduce cross-task and cross-instance statistical constraint knowledge (Lin et al., 2020; Van Nguyen et al., 2021), commonsense knowledge (Ning et al., 2018a), and global event schema knowledge (Li et al., 2020b; Wen et al., 2021; Li et al., 2021a; Jin et al., 2022) that help jointly extract entities, relations, and events.

2.7 Future Research Directions [30min]

IE is a key component in supporting knowledge acquisition and it impacts a wide spectrum of knowledge-driven AI applications. We will conclude the tutorial by presenting further challenges and potential research topics in identifying trustworthiness of extracted content (Zhang et al., 2019, 2020b), IE with quantitative reasoning (Elazar et al., 2019; Zhang et al., 2020a), cross-document IE (Caciularu et al., 2021), incorporating domain-specific knowledge (Lai et al., 2021; Zhang et al., 2021c), extension to knowledge reasoning and prediction, modeling of label semantics (Huang et al., 2022; Mueller et al., 2022; Ma et al., 2022; Chen et al., 2020a), and challenges for acquiring implicit but essential information from corpora that potentially involve reporting bias (Sap et al., 2020).

3 Specification of the Tutorial

The proposed tutorial is considered a **cutting-edge** tutorial that introduces new frontiers in IE research. The presented topic has not been covered by ACL/EMNLP/NAACL/EACL/COLING tutorials in the past 4 years. One exception is the ACL 2020 tutorial "Multi-modal Information Extraction from Text, Semi-structured, and Tabular Data on the Web" that is partly relevant to one of our technical sections (§2.5). That particular section of our talk will focus on IE from visual and multi-media data in addition to semi-structured data, being different from the aforementioned ACL 2020 tutorial that has mainly covered topics on semi-structured data.

Audience and Prerequisites Based on the level of interest in this topic, we expect around 150 participants. While no specific background knowledge is assumed of the audience, it would be the best for the attendees to know about basic deep learning technologies, pre-trained word embeddings (e.g. Word2Vec) and language models (e.g. BERT). A reading list that could help provide background knowledge to the audience before attending this tutorial is given in Appx. §A.2.

Open Access All the materials are openly available at https://cogcomp.seas.upenn.edu/page/tutorial.202207.

4 Tutorial Instructors

The following are biographies of the speaker. Past tutorials given by us are listed in Appx. §A.1.

Muhao Chen is an Assistant Research Professor of Computer Science at USC, where he directs the Language Understanding and Knowledge Acquisition (LUKA) Group. His research focuses on data-driven machine learning approaches for natural language understanding and knowledge acquisition. His work has been recognized with an NSF CRII Award, a Cisco Faculty Research Award, an ACM SIGBio Best Student Paper Award, and a Best Paper Nomination at CoNLL. Muhao obtained his B.S. in Computer Science degree from Fudan University in 2014, his PhD degree from UCLA Department of Computer Science in 2019, and was a postdoctoral researcher at UPenn prior to joining USC. Additional information is available at http://muhaochen.github.io.

Lifu Huang is an Assistant Professor at the Computer Science department of Virginia Tech. He obtained a PhD in Computer Science from UIUC. He has a wide range of research interests in NLP, including extracting structured knowledge with limited supervision, natural language understanding and reasoning with external knowledge and commonsense, natural language generation, representation learning for cross-lingual and cross-domain transfer, and multi-modality learning. He is a recipient of the 2019 AI2 Fellowship and 2021 Amazon Research Award. Additional information is available at https://wilburone.github.io/.

Manling Li is a fourth-year Ph.D. student at the Computer Science Department of UIUC. Manling has won the Best Demo Paper Award at ACL'20, the Best Demo Paper Award at NAACL'21, C.L. Dave and Jane W.S. Liu Award, and has been selected as Mavis Future Faculty Fellow. She is a recipient of Microsoft Research PhD Fellowship. She has more than 30 publications on knowledge extraction and reasoning from multimedia data. Additional information is available at https://limanling.github.io.

Ben Zhou is a third-year Ph.D. student at the Department of Computer and Information Science, University of Pennsylvania. He obtained his B.S. from UIUC in 2019. Ben's research interests are distant supervision extraction and experiential knowledge reasoning, and he has more than 5 recent papers on related topics. He is a recipient of the ENIAC fellowship from the University of Pennsylvania, and a finalist of the CRA outstanding undergraduate researcher award. Additional information is available at http://xuanyu.me/.

Heng Ji is a Professor at Computer Science Department of University of Illinois Urbana-Champaign, and an Amazon Scholar. She received her B.A. and M. A. in Computational Linguistics from Tsinghua University, and her M.S. and Ph.D. in Computer Science from New York University. Her research interests focus on NLP, especially on Multimedia Multilingual Information Extraction, Knowledge Base Population and Knowledge-driven Generation. She was selected as "Young Scientist" and a member of the Global Future Council on the Future of Computing by the World Economic Forum. The awards she received include "AI's 10 to Watch" Award, NSF CAREER award, Google Research Award, IBM Watson Faculty Award, Bosch Research Award, and Amazon AWS Award, ACL2020 Best Demo Paper Award, and NAACL2021 Best Demo Paper Award. Additional information is available at https://blender.cs.illinois.edu/hengji.html.

Dan Roth is the Eduardo D. Glandt Distinguished Professor at the Department of Computer and Information Science, UPenn, the NLP Lead at AWS AI Labs, and a Fellow of the AAAS, ACM, AAAI, and ACL. In 2017 Roth was awarded the John McCarthy Award, the highest award the AI community gives to mid-career AI researchers. Roth was recognized "for major conceptual and theoretical advances in the modeling of natural language understanding, machine learning, and reasoning." Roth has published broadly in machine learning, NLP, KRR, and learning theory, and has given keynote talks and tutorials in all ACL and AAAI major conferences. Roth was the Editor-in-Chief of JAIR until 2017, and was the program chair of AAAI'11, ACL'03 and CoNLL'02; he serves regularly as an area chair and senior program committee member in the major conferences in his research areas. Prof. Roth received his B.A Summa cum laude in Mathematics from the Technion, and his Ph.D. in Computer Science from Harvard University in 1995. Additional information is available at http://www.cis.upenn.edu/~danroth/.

Acknowledgement

This presenters' research is supported in part by U.S. DARPA KAIROS Program No. FA8750-19-2-1004, DARPA AIDA Program No. FA8750-18-2-0014, DARPA MCS program No. N660011924033, and by the NSF of United States Grant IIS 2105329. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

Ethical Considerations

Innovations in technology often face the ethical dilemma of dual use: the same advance may offer potential benefits and harms. For the IE technologies introduced in this tutorial, the distinction between beneficial use and harmful use depends mainly on the data. Proper use of the technology requires that input text corpora, as well as other modalities of inputs, are legally and ethically obtained. Regulation and standards provide a legal framework for ensuring that such data is properly used and that any individual whose data is used has the right to request its removal. In the absence of such regulation, society relies on those who apply technology to ensure that data is used in an ethical way. Besides, training and assessment data may be biased in ways that limit system accuracy on less well represented populations and in new domains, for example causing disparity of performance for different sub-populations based on ethnic, racial, gender, and other attributes. Furthermore, trained systems degrade when used on new data that is distant from their training data. Thus questions concerning generalizability and fairness need to be carefully considered when applying the IE technologies to specific datasets.

A general approach to ensure proper, rather than malicious, application of dual-use technology should: incorporate ethics considerations as the first-order principles in every step of the system design, maintain a high degree of transparency and interpretability of data, algorithms, models, and functionality throughout the system, make software available as open source for public verification and auditing, and explore countermeasures to protect vulnerable groups.

References

Hassan Akbari, Svebor Karaman, Surabhi Bhargava, Brian Chen, Carl Vondrick, and Shih-Fu Chang. 2019. Multi-level multimodal common semantic space for image-phrase grounding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12476–12486.

Avi Caciularu, Arman Cohan, Iz Beltagy, Matthew E Peters, Arie Cattan, and Ido Dagan. 2021. Cdlm: Cross-document language modeling. In *Findings of ACL: EMNLP*.

Pengfei Cao, Yubo Chen, Jun Zhao, and Taifeng Wang. 2020. Incremental event detection via knowledge consolidation networks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 707–717.

Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R Hruschka, and Tom M Mitchell. 2010. Toward an architecture for never-ending language learning. In *AAAI*.

Yee Seng Chan, Joshua Fasching, Haoling Qiu, and Bonan Min. 2019. Rapid customization for event extraction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 31–36.

Snigdha Chaturvedi, Haoruo Peng, and Dan Roth. 2017. Story comprehension for predicting what happens next. In *EMNLP*, pages 1603–1614.

Brian Chen, Xudong Lin, Christopher Thomas, Manling Li, Shoya Yoshida, Lovish Chum, Heng Ji, and Shih-Fu Chang. 2021a. Joint multimedia event extraction from video and article. In *Proc. The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP2021)*.

Brian Chen, Xudong Lin, Christopher Thomas, Manling Li, Shoya Yoshida, Lovish Chum, Heng Ji, and Shih-Fu Chang. 2021b. Joint multimedia event extraction from video and article. *EMNLP Findings*.

Muhao Chen, Hongming Zhang, Haoyu Wang, and Dan Roth. 2020a. "what are you trying to do?" semantic typing of event processes. In *Proceedings of the 24th Conference on Computational Natural Language Learning (CoNLL 2020)*. Association for Computational Linguistics.

Xilun Chen, Ahmed Hassan, Hany Hassan, Wei Wang, and Claire Cardie. 2019a. Multi-source cross-lingual model transfer: Learning what to share. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3098–3112.

Xiuying Chen, Zhangming Chan, Shen Gao, Meng-Hsuan Yu, Dongyan Zhao, and Rui Yan. 2019b. Learning towards abstractive timeline summarization. In *IJCAI*, pages 4939–4945.

Yunmo Chen, Tongfei Chen, Seth Ebner, Aaron Steven White, and Benjamin Van Durme. 2020b. Reading the manual: Event extraction as definition comprehension. In *Proceedings of the Fourth Workshop on Structured Prediction for NLP*, pages 74–83.

Alexis Conneau, Shijie Wu, Haoran Li, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Emerging cross-lingual structure in pretrained language models. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6022–6034.

Chaorui Deng, Qi Wu, Qingyao Wu, Fuyuan Hu, Fan Lyu, and Mingkui Tan. 2018. Visual grounding via accumulated attention. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7746–7755.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL*.

Akshay Raj Dhamija, Manuel Günther, and Terrance E Boult. 2018. Reducing network agnostophobia. In *NeurIPS*, pages 9175–9186.

Xinya Du and Claire Cardie. 2020. Event extraction by answering (almost) natural questions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 671–683.

Yanai Elazar, Abhijit Mahabaly, Deepak Ramachandran, Tania Bedrax-Weiss, and Dan Roth. 2019. How Large Are Lions? Inducing Distributions over Quantitative Attributes. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*.

Joshua Feldman, Joe Davison, and Alexander M. Rush. 2019. Commonsense knowledge mining from pretrained models. *EMNLP*.

Georgia Gkioxari, Ross Girshick, Piotr Dollár, and Kaiming He. 2018. Detecting and recognizing human-object interactions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8359–8367.

Chunhui Gu, Chen Sun, David A Ross, Carl Vondrick, Caroline Pantofaru, Yeqing Li, Sudheendra Vijayanarasimhan, George Toderici, Susanna Ricco, Rahul Sukthankar, et al. 2018. Ava: A video dataset of spatio-temporally localized atomic visual actions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6047–6056.

Rujun Han, Qiang Ning, and Nanyun Peng. 2019. Joint event and temporal relation extraction with shared representations and structured prediction. 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 434–444.

Dan Hendrycks, Xiaoyuan Liu, Eric Wallace, Adam Dziedzic, Rishabh Krishnan, and Dawn Song. 2020. Pretrained transformers improve out-of-distribution robustness. In *ACL*.

Dan Hendrycks, Mantas Mazeika, and Thomas Dietterich. 2018. Deep anomaly detection with outlier exposure. In *International Conference on Learning Representations*.

Jonathan Herzig, Pawel Krzysztof Nowak, Thomas Mueller, Francesco Piccinno, and Julian Eisenschlos. 2020. Tapas: Weakly supervised table parsing via pre-training. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4320–4333.

James Y. Huang, Bangzheng Li Li, Jiashu Xu, and Muhao Chen. 2022. Unified semantic typing with meaningful label inference. In *Proceedings of the 20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*.

Lifu Huang, Kyunghyun Cho, Boliang Zhang, Heng Ji, and Kevin Knight. 2018a. Multi-lingual common semantic space construction via cluster-consistent word embedding. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 250–260.

Lifu Huang and Heng Ji. 2020. Semi-supervised new event type induction and event detection. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 718–724.

Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, and Clare Voss. 2018b. Zero-shot transfer learning for event extraction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2160–2170.

Lifu Huang, Heng Ji, and Jonathan May. 2019. Cross-lingual multi-level adversarial transfer to enhance low-resource name tagging. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3823–3833.

Guoliang Ji, Kang Liu, Shizhu He, and Jun Zhao. 2017. Distant supervision for relation extraction with sentence-level attention and entity descriptions. In *AAAI*.

Xiaomeng Jin, Manling Li, and Heng Ji. 2022. Event schema induction with double graph autoencoders. In *Proc. The 2022 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT2022)*.

Anoop R Katti, Christian Reisswig, Cordula Guder, Sebastian Brarda, Steffen Bickel, Johannes Höhne, and Jean Baptiste Faddoul. 2018. Chargrid: Towards understanding 2d documents. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4459–4469.

Tuan Lai, Heng Ji, ChengXiang Zhai, and Quan Hung Tran. 2021. Joint biomedical entity and relation extraction with knowledge-enhanced collective inference. In *Proc. The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*.

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web*, 6(2):167–195.

Omer Levy, Minjoon Seo, Eunsol Choi, and Luke Zettlemoyer. 2017. Zero-shot relation extraction via reading comprehension. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages 333–342.

Bangzheng Li, Wenpeng Yin, and Muhao Chen. 2022a. Ultra-fine entity typing with indirect supervision from natural language inference. *Transactions of the Association for Computational Linguistics*.

Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, and Jiawei Han. 2021a. The future is not one-dimensional: Complex event schema induction via graph modeling for event prediction. In *Proc. The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP2021)*.

Manling Li, Ruochen Xu, Shuohang Wang, Luowei Zhou, Xudong Lin, Chenguang Zhu, Michael Zeng, Heng Ji, and Shih-Fu Chang. 2022b. Clip-event: Connecting text and images with event structures. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR2022)*.

Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, and Shih-Fu Chang. 2020a. Cross-media structured common space for multimedia event extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2557–2568.

Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers, and Clare Voss. 2020b. 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, Heng Ji, and Jiawei Han. 2021b. Document-level event argument extraction by conditional generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 894–908.

Sha Li, Liyuan Liu, Yiqing Xie, Heng Ji, and Jiawei Han. 2022c. Piled: An identify-and-localize framework for few-shot event detection. In *arXiv:2202.07615*.

Tao Li, Vivek Gupta, Maitrey Mehta, and Vivek Srikumar. 2019a. A logic-driven framework for consistency of neural models. In *EMNLP*.

Xiaoya Li, Fan Yin, Zijun Sun, Xiayu Li, Arianna Yuan, Duo Chai, Mingxin Zhou, and Jiwei Li. 2019b. Entity-relation extraction as multi-turn question answering. In *ACL*.

Xiaobo Liang, Lijun Wu, Juntao Li, Yue Wang, Qi Meng, Tao Qin, Wei Chen, Min Zhang, and Tie-Yan Liu. 2021. R-drop: Regularized dropout for neural networks. *arXiv preprint arXiv:2106.14448*.

Ying Lin, Heng Ji, Fei Huang, and Lingfei Wu. 2020. A joint neural model for information extraction with global features. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7999–8009.

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.

Minqian Liu, Shiyu Chang, and Lifu Huang. 2022. Incremental prompting: Episodic memory prompt for lifelong event detection. *arXiv preprint arXiv:2204.07275*.

Zhengzhong Liu, Chenyan Xiong, Teruko Mitamura, and Eduard Hovy. 2018. Automatic event salience identification. In *EMNLP*, pages 1226–1236.

Lajanugen Logeswaran, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin, and Honglak Lee. 2019. Zero-shot entity linking by reading entity descriptions. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3449–3460.

Yaojie Lu, Hongyu Lin, Jin Xu, Xianpei Han, Jialong Tang, Annan Li, Le Sun, Meng Liao, and Shaoyi Chen. 2021. Text2event: Controllable sequence-to-structure generation for end-to-end event extraction. In *ACL*.

Michal Lukasik, Srinadh Bhojanapalli, Aditya Menon, and Sanjiv Kumar. 2020. Does label smoothing mitigate label noise? In *ICML*, pages 6448–6458.

Qing Lyu, Hongming Zhang, Elior Sulem, and Dan Roth. 2021. Zero-shot event extraction via transfer learning: Challenges and insights. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 322–332.

Jie Ma, Miguel Ballesteros, Srikanth Doss, Rishita Anubhai, Sunil Mallya, Yaser Al-Onaizan, and Dan Roth. 2022. Label semantics for few shot named entity recognition. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 1956–1971.

Aaron Mueller, Jason Krone, Salvatore Romeo, Saab Mansour, Elman Mansimov, Yi Zhang, and Dan Roth. 2022. Label semantic aware pre-training for few-shot text classification. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8318–8334, Dublin, Ireland. Association for Computational Linguistics.

Tsendsuren Munkhdalai, Feifan Liu, Hong Yu, et al. 2018. Clinical relation extraction toward drug safety surveillance using electronic health record narratives: classical learning versus deep learning. *JMIR public health and surveillance*, 4(2):e9361.

Jian Ni, Georgiana Dinu, and Radu Florian. 2017. Weakly supervised cross-lingual named entity recognition via effective annotation and representation projection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1470–1480.

Qiang Ning, Zhili Feng, Hao Wu, and Dan Roth. 2018a. Joint reasoning for temporal and causal relations. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2278–2288.

Qiang Ning, Hao Wu, Haoruo Peng, and Dan Roth. 2018b. Improving temporal relation extraction with a globally acquired statistical resource. In *NAACL*.

Rasha Obeidat, Xiaoli Fern, Hamed Shahbazi, and Prasad Tadepalli. 2019. Description-based zero-shot fine-grained entity typing. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 807–814.

Xiaoman Pan, Boliang Zhang, Jonathan May, Joel Nothman, Kevin Knight, and Heng Ji. 2017. Cross-lingual name tagging and linking for 282 languages. In *Proc. the 55th Annual Meeting of the Association for Computational Linguistics (ACL2017)*.

Xingyuan Pan, Maitrey Mehta, and Vivek Srikumar. 2020. Learning constraints for structured prediction using rectifier networks. In *ACL*, pages 4843–4858.

Pramuditha Perera and Vishal M Patel. 2019. Deep transfer learning for multiple class novelty detection. In *CVPR*, pages 11544–11552.

Sarah Pratt, Mark Yatskar, Luca Weihs, Ali Farhadi, and Aniruddha Kembhavi. 2020. Grounded situation recognition. In *European Conference on Computer Vision*, pages 314–332. Springer.

Chen Qian, Fuli Feng, Lijie Wen, Chunping Ma, and Pengjun Xie. 2021. Counterfactual inference for text classification debiasing. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5434–5445.

Yujie Qian, Enrico Santus, Zhijing Jin, Jiang Guo, and Regina Barzilay. 2019. Graphie: A graph-based framework for information extraction. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 751–761.

Arka Sadhu, Tanmay Gupta, Mark Yatskar, Ram Nevatia, and Aniruddha Kembhavi. 2021. Visual semantic role labeling for video understanding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5589–5600.

Maarten Sap, Vered Shwartz, Antoine Bosselut, Yejin Choi, and Dan Roth. 2020. Commonsense reasoning for natural language processing. In *ACL: Tutorial Abstracts*, pages 27–33.

Daniel N Sosa, Alexander Derry, Margaret Guo, Eric Wei, Connor Brinton, and Russ B Altman. 2019. A literature-based knowledge graph embedding method for identifying drug repurposing opportunities in rare diseases. In *PSB*, pages 463–474.

Ananya Subburathinam, Di Lu, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, and Clare Voss. 2019. Cross-lingual structure transfer for relation and event 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 313–325.

Chen-Tse Tsai, Stephen Mayhew, and Dan Roth. 2016. Cross-lingual named entity recognition via wikification. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*, pages 219–228.

Minh Van Nguyen, Viet Lai, and Thien Huu Nguyen. 2021. Cross-task instance representation interactions and label dependencies for joint information extraction with graph convolutional networks. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 27–38.

Haoyu Wang, Muhao Chen, Hongming Zhang, and Dan Roth. 2020. Joint constrained learning for event-event relation extraction. In *EMNLP*.

Haoyu Wang, Hongming Zhang, Muhao Chen, and Dan Roth. 2021a. Learning constraints and descriptive segmentation for subevent detection. In *EMNLP*.

Hong Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang, and William Yang Wang. 2019a. Sentence embedding alignment for lifelong relation extraction. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 796–806.

Sijia Wang, Mo Yu, Shiyu Chang, Lichao Sun, and Lifu Huang. 2021b. Query and extract: Refining event extraction as type-oriented binary decoding. *arXiv preprint arXiv:2110.07476*.

Sijia Wang, Mo Yu, and Lifu Huang. 2022a. The art of prompting: Event detection based on type specific prompts. *arXiv preprint arXiv:2204.07241*.

Yiwei Wang, Muhao Chen, Wenxuan Zhou, Yujun Cai, Yuxuan Liang, Dayiheng Liu Liu, Baosong Yang, Juncheng Liu, and Bryan Hooi. 2022b. Should we rely on entity mentions for relation extraction? debiasing relation extraction with counterfactual analysis. In *Proceedings of the 20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*.

Yiwei Wang, Muhao Chen, Wenxuan Zhou Zhou, Yujun Cai Cai, Yuxuan Liang, and Bryan Hooi. 2022c. Graphcache: Message passing as caching for sentence-level relation extraction. In *Findings of NAACL*.

Zihan Wang, Jingbo Shang, Liyuan Liu, Lihao Lu, Jiacheng Liu, and Jiawei Han. 2019b. CrossWeigh: Training named entity tagger from imperfect annotations. In *EMNLP*, pages 5154–5163, Hong Kong, China.

Haoyang Wen, Ying Lin, Tuan Lai, Xiaoman Pan, Sha Li, Xudong Lin, Ben Zhou, Manling Li, Haoyu Wang, Hongming Zhang, et al. 2021. Resin: A dockerized schema-guided cross-document cross-lingual cross-media information extraction and event tracking system. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations*, pages 133–143.

Shijie Wu and Mark Dredze. 2019. Beto, bentz, becas: The surprising cross-lingual effectiveness of bert. 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 833–844.

Wei Wu, Fei Wang, Arianna Yuan, Fei Wu, and Jiwei Li. 2020. CorefQA: Coreference resolution as query-based span prediction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6953–6963, Online. Association for Computational Linguistics.

Patrick Xia and Benjamin Van Durme. 2021. Moving on from ontonotes: Coreference resolution model transfer. *arXiv preprint arXiv:2104.08457*.

Mahsa Yarmohammadi, Shijie Wu, Marc Marone, Haoran Xu, Seth Ebner, Guanghui Qin, Yunmo Chen, Jialiang Guo, Craig Harman, Kenton Murray, et al. 2021. Everything is all it takes: A multipronged strategy for zero-shot cross-lingual information extraction. *arXiv preprint arXiv:2109.06798*.

Mark Yatskar, Luke Zettlemoyer, and Ali Farhadi. 2016. Situation recognition: Visual semantic role labeling for image understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5534–5542.

Wenpeng Yin, Nazneen Fatema Rajani, Dragomir Radev, Richard Socher, and Caiming Xiong. 2020. Universal natural language processing with limited annotations: Try few-shot textual entailment as a start. In *EMNLP*, pages 8229–8239.

Pengfei Yu, Heng Ji, and Premkumar Natarajan. 2021. Lifelong event detection with knowledge transfer. In *Proc. The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP2021)*.

Pengfei Yu, Zixuan Zhang, Clare Voss, Jonathan May, and Heng Ji. 2022. Event extractor with only a few examples. In *Proc. NAACL2022 workshop on Deep Learning for Low Resource NLP*.

Alireza Zareian, Svebor Karaman, and Shih-Fu Chang. 2020. Weakly supervised visual semantic parsing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3736–3745.

Runhao Zeng, Wenbing Huang, Mingkui Tan, Yu Rong, Peilin Zhao, Junzhou Huang, and Chuang Gan. 2019. Graph convolutional networks for temporal action localization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7094–7103.

Linhai Zhang, Deyu Zhou, Yulan He, and Zeng Yang. 2021a. Merl: Multimodal event representation learning in heterogeneous embedding spaces. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 14420–14427.

Xikun Zhang, Deepak Ramachandran, Ian Tenney, Yanai Elazar, and Dan Roth. 2020a. Do language embeddings capture scales? In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*, pages 4889–4896.

Xiyang Zhang, Muhao Chen, and Jonathan May. 2021b. Salience-aware event chain modeling for narrative understanding. In *EMNLP*.

Yi Zhang, Zachary Ives, and Dan Roth. 2019. Evidence-based trustworthiness. In *ACL*, pages 413–423.

Yi Zhang, Zachary G. Ives, and Dan Roth. 2020b. "Who said it, and Why?" Provenance for Natural Language Claims. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*.

Yunyi Zhang, Jiaming Shen, Jingbo Shang, and Jiawei Han. 2020c. Empower entity set expansion via language model probing. *ACL*.

Zixuan Zhang, Nikolaus Nova Parulian, Heng Ji, Ahmed S. Elsayed, Skatje Myers, and Martha Palmer. 2021c. Biomedical information extraction based on knowledge-enriched abstract meaning representation. In *Proc. The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*.

Ben Zhou, Daniel Khashabi, Chen-Tse Tsai, and Dan Roth. 2018. Zero-shot open entity typing as type-compatible grounding. In *EMNLP*.

Ben Zhou, Qiang Ning, Daniel Khashabi, and Dan Roth. 2020. Temporal common sense acquisition with minimal supervision. *ACL*.

Ben Zhou, Kyle Richardson, Qiang Ning, Tushar Khot, Ashish Sabharwal, and Dan Roth. 2021a. Temporal reasoning on implicit events from distant supervision. *NAACL*.

Joey Tianyi Zhou, Hao Zhang, Di Jin, Hongyuan Zhu, Meng Fang, Rick Siow Mong Goh, and Kenneth Kwok. 2019. Dual adversarial neural transfer for low-resource named entity recognition. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3461–3471.

Wenxuan Zhou and Muhao Chen. 2021. Learning from noisy labels for entity-centric information extraction. In *EMNLP*.

Wenxuan Zhou, Fangyu Liu, and Muhao Chen. 2021b. Contrastive out-of-distribution detection for pretrained transformers. In *EMNLP*.

A Appendix

A.1 Past Tutorials by the Instructors

The presenters of this tutorial have given the following tutorials at leading international conferences and venues in the past:

- Muhao Chen:

- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.
- KDD'21: From Tables to Knowledge: Recent Advances in Table Understanding.
- AAAI'20: Recent Advances of Transferable Representation Learning.

• Manling Li:

- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.

• Heng Ji:

- KDD'22: The Battlefront of Combating Misinformation and Coping with Media Bias.
- AACL'22: The Battlefront of Combating Misinformation and Coping with Media Bias.
- KDD'22: New Frontiers of Scientific Text Mining: Tasks, Data, and Tools.
- WWW'22: Modern Natural Language Processing Techniques for Scientific Web Mining: Tasks, Data, and Tools.
- AAAI'22: Deep Learning on Graphs for Natural Language Processing.
- KDD'21: Deep Learning on Graphs for Natural Language Processing.
- IJCAI'21: Deep Learning on Graphs for Natural Language Processing.
- SIGIR'21: Deep Learning on Graphs for Natural Language Processing.
- EMNLP'21: Knowledge-Enriched Natural Language Generation.
- ACL'21: Event-Centric Natural Language Processing.
- NAACL'21: Deep Learning on Graphs for Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.
- CCL'18 and NLP-NADB'18: Multi-lingual Entity Discovery and Linking.
- ACL'18: Multi-lingual Entity Discovery and Linking.
- SIGMOD'16: Automatic Entity Recognition and Typing in Massive Text Data.
- ACL'15: Successful Data Mining Methods for NLP.
- ACL'14: Wikification and Beyond: The Challenges of Entity and Concept Grounding.
- NLPCC'14: Wikification and Beyond: The Challenges of Entity and Concept Grounding.
- COLING'12: Temporal Information Extraction and Shallow Temporal Reasoning.

• Dan Roth:

- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.
- ACL'20: Commonsense Reasoning for Natural Language Processing.
- AAAI'20: Recent Advances of Transferable Representation Learning.
- ACL'18: A tutorial on Multi-lingual Entity Discovery and Linking.
- EACL'17: A tutorial on Integer Linear Programming Formulations in Natural Language Processing.
- AAAI'16: A tutorial on Structured Prediction.
- ACL'14: A tutorial on Wikification and Entity Linking.
- AAAI'13: Information Trustworthiness.
- COLING'12: A Tutorial on Temporal Information Extraction and Shallow Temporal Reasoning.
- NAACL'12: A Tutorial on Constrained Conditional Models: Structured Predictions in NLP.
- NAACL'10: A Tutorial on Integer Linear Programming Methods in NLP.
- EACL'09: A Tutorial on Constrained Conditional Models.
- ACL'07: A Tutorial on Textual Entailment.

A.2 Recommended Paper List

The following is a reading list that could help provide background knowledge to the audience before attending this tutorial:

• Wenxuan Zhou, Muhao Chen. Learning from Noisy Labels for Entity-Centric Information Extraction. EMNLP, 2021.

• Wenxuan Zhou, Fanyu Liu, Muhao Chen. Contrastive Out-of-Distribution Detection for Pretrained Transformers. EMNLP, 2021.

• Xingyuan Pan, Maitrey Mehta, Vivek Srikumar. Learning Constraints for Structured Prediction Using Rectifier Networks. ACL, 2020.

- Hangfeng He, Mingyuan Zhang, Qiang Ning, Dan Roth. Foreseeing the Benefits of Incidental Supervision. EMNLP, 2021.
- Ben Zhou, Qiang Ning, Daniel Khashabi, Dan Roth. Temporal Common Sense Acquisition with Minimal Supervision. ACL, 2020.
- Wenpeng Yin, Nazneen Fatema Rajani, Dragomir Radev, Richard Socher, Caiming Xiong. Universal natural language processing with limited annotations: Try few-shot textual entailment as astart. EMNLP, 2020.
- Bangzheng Li, Wenpeng Yin, Muhao Chen. Ultra-fine Entity Typing with Indirect Supervision from Natural Language Inference. TACL, 2022.
- Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, Clare Voss. Zero-shot transfer learning for event extraction. ACL, 2018.
- Ananya Subburathinam, Di Lu, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, Clare Voss. Cross-lingual structure transfer for relation and event extraction. EMNLP, 2019.
- Hong Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang, William Yang Wang. Sentence Embedding Alignment for Lifelong Relation Extraction. NAACL, 2019.
- Hassan Akbari, Svebor Karaman, Surabhi Bhargava, Brian Chen, Carl Vondrick, and Shih-Fu Chang. Multi-level multimodal common semantic space for image-phrase grounding. CVPR, 2019.
- Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, Shih-Fu Chang. Cross-media structured common space for multimedia event extraction. ACL, 2020.