The NLLP organizers gratefully acknowledge the support from the following sponsors.

In Collaboration With

Bloomberg

Human ADS

[Images of logos]
Introduction

Welcome to the fifth edition of the NLLP (Natural Legal Language Processing) Workshop, co-located with the 2023 Conference on Empirical Methods in Natural Language Processing. Different industrial sectors have embraced natural language processing (NLP) technologies, which have altered services and products in healthcare, finance, education among others. The legal domain, in particular, stands as a promising frontier for the exploration of interesting research problems. Electronic tools have become increasingly integral in the practice of law, with their usage projected to undergo exponential growth. By its very nature, the practice of law necessarily involves the analysis and interpretation of language. The potential for NLP applications to provide benefit to practitioners of law and consumers of legal services around the world is enormous.

We organized this workshop to bring together researchers, practitioners, policy makers from around the world who develop NLP techniques within the legal domain. This is an exciting opportunity to expand the boundaries of our field by identifying new problems and exploring new data as it interacts with the full inventory of NLP and machine learning approaches. In this spirit, the Organizing and Program Committee was assembled to include researchers from both academia and industry, from NLP and legal backgrounds.

We were interested in the following types of papers: (1) applications of NLP methods to legal tasks; (2) experimental results using and adapting NLP methods in legal documents; (3) descriptions of new legal tasks for NLP; (4) creation of curated and/or annotated resources; (5) descriptions of systems which use NLP technologies for legal text; (6) industrial research in this area and (7) interdisciplinary position papers.

We again received a record number of submissions. In total, we received of 64 submissions on our direct submission website, out of which we accepted 32 for an overall acceptance rate of 50%. Out of the 32 accepted papers, 22 are long and 10 are short. The accepted papers include 6 original papers submitted as non-archival, in order to accommodate publication of the work at a later date in a conference or journal. All papers were reviewed by at least 3 members of the Program Committee, one of which usually including legal scholar. In addition, we also offered a venue for presentation for 10 papers accepted to the Findings of EMNLP 2023 on the above topics. All 42 papers were invited to have an oral presentation, either in-person or remote, as well as the option to present a poster in the lunch poster session.

The papers cover a wide range of topics including new data sets for legal NLP, position papers on open legal NLP issues, legal perspectives on NLP topics, as well as applications of NLP tasks to legal documents including retrieval, information extraction, legal judgement prediction and statutory reasoning. A number of papers also explore the use of Large Language Models (LLMs) in the legal domain spanning from fundamental questions to usage methods and applications such as prompting strategies, chain-of-thought applications in retrieval tasks or e-discovery. Finally, we note the diversity of legal domains/languages with papers working in French, Hebrew, Italian, Korean or resources covering up to 24 languages.

We thank our invited speaker Jerrold Soh, Assistant Professor at the Yong Pung How School of Law and Deputy Director of the Centre of Computational Law at Singapore Management University for accepting our invitation. In the tradition of past NLLP workshops, the invited speaker is a legal scholar with an interest in empirical methods for legal analysis including NLP methods. He will be giving a talk entitled NLP in the Legal World which aims to identify the unique role that NLP can play in furthering our understanding of law and facilitating practical access to justice.

We thank everyone who expressed interest in the workshop, all authors of submitted papers, members of the Program Committee who did an excellent job at reviewing papers given a short turnaround time, everyone attending the workshop and the EMNLP 2023 conference for hosting us and the workshop. We thank our sponsors – Bloomberg and the European Research Council Starting Grant project HUMANads – for their contributions.

The NLLP Workshop organizers.

http://nllpw.org
Program Committee

Organizing Committee

Nikolaos Aletras, University of Sheffield
Leslie Barrett, Bloomberg
Ilias Chalkidis, University of Copenhagen
Catalina Goanta, Utrecht University
Daniel Preotiuc-Pietro, Bloomberg
Gerasimos Spanakis, Maastricht University

Program Committee

Tommaso Agnoloni, IGSG CNR
Ios Androutsopoulos, Athens University of Economics and Business
Elliott Ash, ETH Zurich
Breck Baldwin, Safety Third
Claire Barale, The University of Edinburgh
Thales Bertaglia, Maastricht University
Andrew Blair-Stanek, University of Maryland Law School
Stephanie Brandl, University of Copenhagen
Rajarathnam Chandramouli, Spectronn Inc
Daniel Chen, Toulouse School of Economics
Ashish Chouhan, Heidelberg University
Tony Davis, Southern Oregon University
Stefania Degaetano-Ortlieb, Saarland University
Luigi Di Caro, University of Turin
Manos Fergadiotis, Athens University of Economics and Business
Nicolas Garneau, University of Copenhagen
Ivan Habernal, Paderborn University
Ben Hagag, Bar Ilan University
Nils Holzenberger, Télécom Paris
Abderrahmane Issam, Maastricht University
Kartik Kannapur, Amazon Web Services
Constantinos Karouzos, University of Sheffield
Panagiota Katsikouli, University of Copenhagen
Daniel Katz, Illinois Tech - Chicago Kent College of Law
Ilan Kerner, K Dictionaries - Lexicala
Moniba Keymanesh, Bloomberg
Aykut Koc, Bilkent University
Manolis Koubarakis, National and Kapodistrian University of Athens
Junyi Jessy Li, University of Texas at Austin
Jiaang Li, University of Copenhagen
Ruta Liepina, University of Bologna
Antoine Louis, Maastricht University
Adam Meyers, New York University
Jelena Mitrović, University of Passau
Joel Niklaus, University of Bern
Milda Norkute, Thomson Reuters
Daphne Odekerken, Utrecht University
Keynote Talk: NLP in the Legal World

Jerrold Soh
Yong Pung How School of Law, Singapore Management University, Singapore

Abstract: This talk situates the rising field of NLLP in the context of legal scholarship and practice. It will examine how the field relates to existing inquiries in computational law, AI and Law, and computational/empirical legal studies. Similarities, differences, and opportunities for cross-fertilization and inter-field discourse will be identified along four dimensions: questions generally asked, techniques generally used, publications generally produced, and the people generally involved. By elucidating links and bridges between NLLP and adjacent legal fields, the talk aims to identify the unique role that NLLP can play, and indeed has already been playing, in furthering our understanding of law and facilitating practical access to justice.

Bio: Jerrold Soh is an Assistant Professor at the Yong Pung How School of Law, Singapore Management University and the Deputy Director of its Centre for Computational Law. His research revolves around how emerging technologies, especially artificial intelligence, should be regulated by and also used in the legal system. He has published both doctrinal work on AI and automated vehicles liability as well as scientific work developing machine learning and network models for legal use cases. He holds degrees in Law and Economics from the National University of Singapore and an LLM from Harvard Law School.
# Table of Contents

*Anthropomorphization of AI: Opportunities and Risks*
Ameet Deshpande, Tanmay Rajpurohit, Karthik Narasimhan and Ashwin Kalyan .......................... 1

*NOMOS: Navigating Obligation Mining in Official Statutes*
Andrea Pennisi, Elvira González Hernández and Nina Koivula ........................................ 8

*Long Text Classification using Transformers with Paragraph Selection Strategies*
Mohit Tuteja and Daniel González Juclà ................................................................. 17

*Do Language Models Learn about Legal Entity Types during Pretraining?*
Claire Barale, Michael Rovatsos and Nehal Bhuta ......................................................... 25

*Pretrained Language Models v. Court Ruling Predictions: A Case Study on a Small Dataset of French Court of Appeal Rulings*
Olivia Vaudaux, Caroline Bazzoli, Maximin Coavoux, Géraldine Vial and Étienne Vergès ... 38

*Italian Legislative Text Classification for Gazzetta Ufficiale*
Marco Rovera, Alessio Palmero Aprosio, Francesco Greco, Mariano Lucchese, Sara Tonelli and Antonio Antetomaso ............................................................ 44

*Mixed-domain Language Modeling for Processing Long Legal Documents*
Wenyue Hua, Yuchen Zhang, Zhe Chen, Josie Li and Melanie Weber ............................... 51

*Questions about Contracts: Prompt Templates for Structured Answer Generation*
Adam Roegiest, Radha Chitta, Jonathan Donnelly, Maya Lash, Alexandra Vtyurina and Francois Longtin ................................................................. 62

*Legal Judgment Prediction: If You Are Going to Do It, Do It Right*
Masha Medvedeva and Pauline Mcbride ................................................................. 73

*Beyond The Text: Analysis of Privacy Statements through Syntactic and Semantic Role Labeling*
Yan Shvartzshanider, Ananth Balashankar, Thomas Wies and Lakshminarayanan Subramanian85

*Towards Mitigating Perceived Unfairness in Contracts from a Non-Legal Stakeholder’s Perspective*
Anmol Singhal, Preethu Rose Anish, Shirish Karande and Smita Ghaisas .......................... 99

*Connecting Symbolic Statutory Reasoning with Legal Information Extraction*
Nils Holzenberger and Benjamin Van Durme ......................................................... 113

*Retrieval-based Evaluation for LLMs: A Case Study in Korean Legal QA*
Cheol Ryu, Seolhwa Lee, Subeen Pang, Chanyeol Choi, Hojun Choi, Myeonggee Min and Jy-Yong Sohn ................................................................. 132

*Legal NLP Meets MiCAR: Advancing the Analysis of Crypto White Papers*
Carolina Camassa ................................................................................................. 138

*Low-Resource Deontic Modality Classification in EU Legislation*
Kristina Minkova, Shashank Chakravarthry and Gijs Dijck .............................................. 149

*Automatic Anonymization of Swiss Federal Supreme Court Rulings*
Joel Niklaus, Robin Mamié, Matthias Stürmer, Daniel Brunner and Marcel Gygli .............. 159

*Exploration of Open Large Language Models for eDiscovery*
Sumit Pai, Sounak Lahiri, Ujjwal Kumar, Krishanu Baksi, Elijah Soba, Michael Suesserman, Nirmala Pudota, Jon Foster, Edward Bowen and Sanmitra Bhattacharya ......................... 166
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval-Augmented Chain-of-Thought in Semi-structured Domains</td>
<td>178</td>
</tr>
<tr>
<td>Vaibhav Mavi, Abulhair Saparov and Chen Zhao</td>
<td></td>
</tr>
<tr>
<td>Joint Learning for Legal Text Retrieval and Textual Entailment: Leveraging the Relationship between Relevancy and Affirmation</td>
<td>192</td>
</tr>
<tr>
<td>Nguyen Hai Long, Thi Hai Yen Vuong, Ha Thanh Nguyen and Xuan-Hieu Phan</td>
<td></td>
</tr>
<tr>
<td>Super-SCOTUS: A multi-sourced dataset for the Supreme Court of the US</td>
<td>202</td>
</tr>
<tr>
<td>Biaoyan Fang, Trevor Cohn, Timothy Baldwin and Lea Frermann</td>
<td></td>
</tr>
<tr>
<td>Transferring Legal Natural Language Inference Model from a US State to Another: What Makes It So Hard?</td>
<td>215</td>
</tr>
<tr>
<td>Alice Kwak, Gaetano Forte, Derek Bambauer and Mihai Surdeanu</td>
<td></td>
</tr>
<tr>
<td>Large Language Models are legal but they are not: Making the case for a powerful LegalLLM</td>
<td>223</td>
</tr>
<tr>
<td>Thanmay Jayakumar, Fauzan Farooqui and Luqman Farooqui</td>
<td></td>
</tr>
<tr>
<td>On the Potential and Limitations of Few-Shot In-Context Learning to Generate Metamorphic Specifications for Tax Preparation Software</td>
<td>230</td>
</tr>
<tr>
<td>Dananjay Srinivas, Rohan Das, Saeid Tizpaz-Niari, Ashutosh Trivedi and Maria Leonor Pacheco</td>
<td></td>
</tr>
<tr>
<td>AsyLex: A Dataset for Legal Language Processing of Refugee Claims</td>
<td>244</td>
</tr>
<tr>
<td>Claire Barale, Mark Klaisoongnoen, Pasquale Minervini, Michael Rovatsos and Nehal Bhuta</td>
<td></td>
</tr>
<tr>
<td>A Comparative Study of Prompting Strategies for Legal Text Classification</td>
<td>258</td>
</tr>
<tr>
<td>Ali Hakimi Parizi, Yuyang Liu, Prudhvi Nokku, Sina Gholamian and David Emerson</td>
<td></td>
</tr>
<tr>
<td>Tracing Influence at Scale: A Contrastive Learning Approach to Linking Public Comments and Regulator Responses</td>
<td>266</td>
</tr>
<tr>
<td>Linzi Xing, Brad Hackinen and Giuseppe Caremini</td>
<td></td>
</tr>
</tbody>
</table>
Program

Wednesday, December 7, 2022

09:00 - 09:10 Workshop Opening

09:10 - 10:30 Session 1

Anthropomorphization of AI: Opportunities and Risks
Ameet Deshpande, Tanmay Rajpurohit, Karthik Narasimhan and Ashwin Kalyan

Legal NLP Meets MiCAR: Advancing the Analysis of Crypto White Papers
Carolina Camassa

On the Potential and Limitations of Few-Shot In-Context Learning to Generate Metamorphic Specifications for Tax Preparation Software
Dananjay Srinivas, Rohan Das, Saeid Tizpaz-Niari, Ashutosh Trivedi and Maria Leonor Pacheco

NOMOS: Navigating Obligation Mining in Official Statutes
Andrea Pennisi, Elvira González Hernández and Nina Koivula

Long Text Classification using Transformers with Paragraph Selection Strategies
Mohit Tuteja and Daniel González Juclà

Italian Legislative Text Classification for Gazzetta Ufficiale
Marco Rovera, Alessio Palmero Aprosio, Francesco Greco, Mariano Lucchese, Sara Tonelli and Antonio Antetomaso

Towards Mitigating Perceived Unfairness in Contracts from a Non-Legal Stakeholder's Perspective
Anmol Singhal, Preethu Rose Anish, Shirish Karande and Smita Ghaisas

Low-Resource Deontic Modality Classification in EU Legislation
Kristina Minkova, Shashank Chakravarthy and Gijs Dijck

Transferring Legal Natural Language Inference Model from a US State to Another: What Makes It So Hard?
Alice Kwak, Gaetano Forte, Derek Bambauer and Mihai Surdeanu

Large Language Models for Legally Enforceable Hate Speech Detection
Chu Fei Luo, Rohan Bhambhoria, Samuel Dahan and Xiaodan Zhu
Wednesday, December 7, 2022 (continued)

More than Votes? Voting and Language based Partisanship in the US Supreme Court
Biaoyan Fang, Trevor Cohn, Timothy Baldwin and Lea Frermann

10:30 - 11:00  
Break

11:00 - 12:30  
Session 2

Retrieval-Augmented Chain-of-Thought in Semi-structured Domains
Vaibhav Mavi, Abulhair Saparov and Chen Zhao

Mixed-domain Language Modeling for Processing Long Legal Documents
Wenyue Hua, Yuchen Zhang, Zhe Chen, Josie Li and Melanie Weber

Large Language Models are legal but they are not: Making the case for a powerful LegalLLM
Thanmay Jayakumar, Fauzan Farooqui and Luqman Farooqui

Exploration of Open Large Language Models for eDiscovery
Sumit Pai, Sounak Lahiri, Ujjwal Kumar, Krishanu Baksi, Elijah Soba, Michael Suesserman, Nirmala Pudota, Jon Foster, Edward Bowen and Sanmitra Bhattacharya

A Comparative Study of Prompting Strategies for Legal Text Classification
Ali Hakimi Parizi, Yuyang Liu, Prudhvi Nokku, Sina Gholamian and David Emerson

A Comprehensive Evaluation of Large Language Models on Legal Judgment Prediction
Ruihao Shui, Yixin Cao, Xiang Wang and Tat-Seng Chua

SCALE: Scaling up the Complexity for Advanced Language Model Evaluation
Vishvakosenan Rasiah, Ronja Stern, Veton Matoshi, Matthias Stürmer, Ilias Chalkidis, Daniel Ho and Joel Niklaus

Super-SCOTUS: A multi-sourced dataset for the Supreme Court of the US
Biaoyan Fang, Trevor Cohn, Timothy Baldwin and Lea Frermann

MultiLegalPile: A 689GB Multilingual Legal Corpus
Joel Niklaus, Veton Matoshi, Matthias Stürmer, Ilias Chalkidis and Daniel Ho
Wednesday, December 7, 2022 (continued)

**ECtHR-PCR: A Dataset for Precedent Understanding and Prior Case Retrieval in the European Court of Human Rights**
Santosh T.y.s.s, Rashid Haddad and Matthias Grabmair

**AsyLex: A Dataset for Legal Language Processing of Refugee Claims**
Claire Barale, Mark Klaisoongnoen, Pasquale Minervini, Michael Rovatsos and Nehal Bhuta

**LEXTREME: A Multi-Lingual and Multi-Task Benchmark for the Legal Domain**
Joel Niklaus, Veton Matoshi, Pooja Rani, Andrea Galassi, Matthias Stürmer and Ilias Chalkidis

12:30 - 14:00  *Lunch and In-Person Poster Session*

14:00 - 15:00  *Keynote - NLP in the Legal World - Jerrold Soh (Yong Pung How School of Law)*

14:00 - 15:30  *Session 3*

**Retrieval-based Evaluation for LLMs: A Case Study in Korean Legal QA**
Cheol Ryu, Seolhwa Lee, Subeen Pang, Chanyeol Choi, Hojun Choi, Myeonggee Min and Jy-Yong Sohn

**Joint Learning for Legal Text Retrieval and Textual Entailment: Leveraging the Relationship between Relevancy and Affirmation**
Nguyen Hai Long, Thi Hai Yen Vuong, Ha Thanh Nguyen and Xuan-Hieu Phan

**Tracing Influence at Scale: A Contrastive Learning Approach to Linking Public Comments and Regulator Responses**
Linzi Xing, Brad Hackinen and Giuseppe Carenini

**Legal Passage Retrieval: A pragmatic approach to legal AI**
Robert Mahari, Dominik Stammbach, Elliott Ash and Alex Pentland

15:30 - 16:00  *Break*

16:00 - 17:30  *Session 4*

**Towards Explainability and Fairness in Swiss Judgement Prediction: Benchmarking on a Multilingual Dataset**
Santosh T.y.s.s, Nina Baumgartner, Matthias Stürmer, Matthias Grabmair and Joel Niklaus
LLMs – the Good, the Bad or the Indispensable?: A Use Case on Legal Statute Prediction and Legal Judgment Prediction on Indian Court Cases
Shaurya Vats, Atharva Zope, Somsubhra De, Anurag Sharma, Upal Bhattacharya, Shubham Kumar Nigam, Shouvik Kumar Guha, Koustav Rudra and Kripabandhu Ghosh

Exploiting Contrastive Learning and Numerical Evidence for Confusing Legal Judgment Prediction
Leilei Gan, Baokui Li, Kun Kuang, Yating Zhang, Lei Wang, Anh Tuan Luu, Yi Yang and Fei Wu

Multi-Defendant Legal Judgment Prediction via Hierarchical Reasoning
Yougang Lyu, Jitai Hao, Zihan Wang, Kai Zhao, Shen Gao, Pengjie Ren, Zhumin Chen, Fang Wang and Zhaochun Ren

Pretrained Language Models v. Court Ruling Predictions: A Case Study on a Small Dataset of French Court of Appeal Rulings
Olivia Vaudaux, Caroline Bazzoli, Maximin Couvoux, Géraldine Vial and Étienne Vergès

Legal Judgment Prediction: If You Are Going to Do It, Do It Right
Masha Medvedeva and Pauline Mcbride

Extracting Sentencing-Related Factors from Criminal Cases in Hebrew
Din Ezra, Maxim Bragilovski, Itay Razumenko, Keren Gorelik, Lior Kobi, Lior Rokach, Arnon Strum and Nir Grinberg

Do Language Models Learn about Legal Entity Types during Pretraining?
Claire Barale, Michael Rovatsos and Nehal Bhuta

Questions about Contracts: Prompt Templates for Structured Answer Generation
Adam Roegiest, Radha Chitta, Jonathan Donnelly, Maya Lash, Alexandra Vtyurina and Francois Longtin

Resolving Legalese: A Multilingual Exploration of Negation Scope Resolution in Legal Documents
Ramona Christen, Anastassia Shaitarova, Matthias Stürmer and Joel Niklaus

Beyond The Text: Analysis of Privacy Statements through Syntactic and Semantic Role Labeling
Yan Shvartzshnanider, Ananth Balashankar, Thomas Wies and Lakshminarayanan Subramanian

Connecting Symbolic Statutory Reasoning with Legal Information Extraction
Nils Holzenberger and Benjamin Van Durme
Wednesday, December 7, 2022 (continued)

Automatic Anonymization of Swiss Federal Supreme Court Rulings
Joel Niklaus, Robin Mamié, Matthias Stürmer, Daniel Brunner and Marcel Gygli

Automatic Anonymization of Swiss Federal Supreme Court Rulings
Joel Niklaus, Robin Mamié, Matthias Stürmer, Daniel Brunner and Marcel Gygli

Can ChatGPT Perform Reasoning Using the IRAC Method in Analyzing Legal Scenarios Like a Lawyer?
Xiaoxi Kang, Lizhen Qu, Lay-Ki Soon, Adnan Trakic, Terry Yue Zhuo, Patrick Charles Emerton and Genevieve Grant

Information Extraction from Legal Wills: How Well Does GPT-4 Do?
Alice Saebom Kwak, Cheonkam Jeong, Gaetano Vincent Forte, Derek Bambauer, Clayton T Morrison and Mihai Surdeanu

17:30 - 17:35 Closing Remarks & Best Presentation Award