

SDP 2025

**Fifth Workshop on Scholarly Document Processing**

**Proceedings of the Workshop**

July 31, 2025

©2025 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
317 Sidney Baker St. S  
Suite 400 - 134  
Kerrville, TX 78028  
USA  
Tel: +1-855-225-1962  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 979-8-89176-265-7

## **Message from the SDP 2025 Organizing Committee**

Welcome to the Fifth Workshop on Scholarly Document Processing (SDP) at ACL 2025. As the body of scholarly literature grows, automated methods in NLP, text mining, information retrieval, document understanding etc. are needed to address issues of information overload, disinformation, reproducibility, and more. Though progress has been made, there are significant unique challenges to processing scholarly text that require dedicated attention. The goal of the Scholarly Document Processing series of workshops is to provide a venue for addressing these challenges, as well as a platform for tasks and resources supporting the processing of scientific documents. Our long-term objective is to establish scholarly and scientific texts as an essential domain for NLP research, to supplement current efforts on web text and news articles. This workshop builds on the success of prior workshops: the SDP workshop held at ACL in 2024, COLING in 2022, NAACL in 2021 and EMNLP in 2020, and the SciNLP workshop held at AKBC 2020 and AKBC 2021. As in previous years, we have sought to bring together researchers and practitioners from various backgrounds focusing on different aspects of scholarly document processing. We believe that the interdisciplinary nature of the ACL venues greatly assists in encouraging submissions from a diverse set of fields.

# Organizing Committee

## Program Chairs

Tirthankar Ghosal, Oak Ridge National Laboratory, USA

Philipp Mayr, GESIS - Leibniz Institute for the Social Sciences, Germany

Amanpreet Singh, Allen Institute for AI, USA

Aakanksha Naik, Allen Institute for AI, USA

Georg Rehm, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) and Humboldt-Universität zu Berlin, Germany

Dayne Freitag, SRI International, USA

Dan Li, Elsevier, Netherlands

Sonja Schimmler, Technische Universität Berlin and Fraunhofer FOKUS, Germany

Anita De Waard, Elsevier, Netherlands

## Program Committee

### Reviewers

Raia Abu Ahmad, Anurag Acharya, Daniel Acuna, Akiko Aizawa, Hamed Alhoori, Tamjid Azad, Ibrahim Al Azher

Fabio Barth, Ekaterina Borisova, Ioana Buhnila

Yupeng Cao, Soham Chitnis

Anita De Waard, Jay DeYoung, James Dunham, Nicolau Duran-Silva

Dayne Freitag

German Gritsai

Allan Hanbury, Tosho Hirasawa, Sameera Horawalavithana

Daisuke Ikeda

Yavuz Selim Kartal, Roman Kern, Petr Knoth, Buse Sibel Korkmaz

Dan Li, Xinyuan Lu

Biswadip Mandal, Yoshitomo Matsubara, Philipp Mayr, Shufan Ming, Miftahul Jannat Mokar-rama

Wolfgang Otto

Antonio Pieri

Tohida Rehman, Allen G Roush

Sebastian Schellhammer, Pierre Senellart, Amanpreet Singh, Neil R. Smalheiser, Wojtek Sylwe-strzak

Sotaro Takeshita, Hiroki Teranishi

Sharmila Upadhyaya, Aida Usmanova

Boris Veytsman

Taro Watanabe

Shiyuan Zhang, Wuhe Zou

# Table of Contents

## *Overview of the Fifth Workshop on Scholarly Document Processing*

Tirthankar Ghosal, Philipp Mayr, Anita De Waard, Aakanksha Naik, Amanpreet Singh, Dayne Freitag, Georg Rehm, Sonja Schimmmler and Dan Li . . . . . 1

## *TeXpert: A Multi-Level Benchmark for Evaluating L<sup>A</sup>T<sub>E</sub>X Code Generation by LLMs*

Sahil Kale and Vijaykant Nadadur . . . . . 7

## *MathD2: Towards Disambiguation of Mathematical Terms*

Shufan Jiang, Mary Ann Tan and Harald Sack . . . . . 17

## *GraphTranslate: Predicting Clinical Trial Translation using Graph Neural Networks on Biomedical Literature*

Emily Muller, Justin Boylan-Toomey, Jack Ekinsmyth, Arne Robben, María De La Paz Cardona and Antonia Langfelder . . . . . 31

## *The ClimateCheck Dataset: Mapping Social Media Claims About Climate Change to Corresponding Scholarly Articles*

Raia Abu Ahmad, Aida Usmanova and Georg Rehm . . . . . 42

## *Analyzing the Evolution of Scientific Misconduct Based on the Language of Retracted Papers*

Christof Bless, Andreas Waldis, Angelina Parfenova, Maria A. Rodriguez and Andreas Marfurt 57

## *Collage: Decomposable Rapid Prototyping for Co-Designed Information Extraction on Scientific PDFs*

Sireesh Gururaja, Yueheng Zhang, Guannan Tang, Tianhao Zhang, Kevin Murphy, Yu-Tsen Yi, Junwon Seo, Anthony Rollett and Emma Strubell . . . . . 72

## *Literature discovery with natural language queries*

Anna Kiepora, Jessica Lam, Nianlong Gu and Richard Hahnloser . . . . . 83

## *Literature-Grounded Novelty Assessment of Scientific Ideas*

Simra Shahid, Marissa Radensky, Raymond Fok, Pao Siangliulue, Daniel S Weld and Tom Hope 96

## *Data Gatherer: LLM-Powered Dataset Reference Extraction from Scientific Literature*

Pietro Marini, Aécio Santos, Nicole Contaxis and Juliana Freire . . . . . 114

## *Predicting The Scholarly Impact of Research Papers Using Retrieval-Augmented LLMs*

Tamjid Azad, Ibrahim Al Azher, Sagnik Ray Choudhury and Hamed Alhoori . . . . . 124

## *Document Attribution: Examining Citation Relationships using Large Language Models*

Vipula Rawte, Ryan A. Rossi, Franck Dernoncourt and Nedim Lipka . . . . . 132

## *SOMD2025: A Challenging Shared Tasks for Software Related Information Extraction*

Sharmila Upadhyaya, Wolfgang Otto, Frank Krüger and Stefan Dietze . . . . . 137

## *From In-Distribution to Out-of-Distribution: Joint Loss for Improving Generalization in Software Mention and Relation Extraction*

Stasa Mandic, Georg Niess and Roman Kern . . . . . 146

## *SOMD 2025: Fine-tuning ModernBERT for In- and Out-of-Distribution NER and Relation Extraction of Software Mentions in Scientific Texts*

Vaghawan Ojha, Projan Shakya, Kristina Ghimire, Kashish Bataju, Ashwini Mandal, Sadikshya Gyawali, Manish Dahal, Manish Awale, Shital Adhikari and Sanjay Rijal . . . . . 154

|                                                                                                                                                                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Inductive Learning on Heterogeneous Graphs Enhanced by LLMs for Software Mention Detection</i><br>Gabriel Silva, Mário Rodrigues, António Teixeira and Marlene Amorim .....                                                                            | 164 |
| <i>Extracting Software Mentions and Relations using Transformers and LLM-Generated Synthetic Data at SOMD 2025</i><br>Pranshu Rastogi and Rajneesh Tiwari .....                                                                                           | 173 |
| <i>SciVQA 2025: Overview of the First Scientific Visual Question Answering Shared Task</i><br>Ekaterina Borisova, Nikolas Rauscher and Georg Rehm .....                                                                                                   | 182 |
| <i>Visual Question Answering on Scientific Charts Using Fine-Tuned Vision-Language Models</i><br>Florian Schleid, Jan Strich and Chris Biemann .....                                                                                                      | 211 |
| <i>ExpertNeurons at SciVQA-2025: Retrieval Augmented VQA with Vision Language Model (RAVQA-VLM)</i><br>Nagaraj N Bhat, Joydeb Mondal and Srijon Sarkar .....                                                                                              | 221 |
| <i>Coling-UniA at SciVQA 2025: Few-Shot Example Retrieval and Confidence-Informed Ensembling for Multimodal Large Language Models</i><br>Christian Jaumann, Annemarie Friedrich and Rainer Lienhart .....                                                 | 230 |
| <i>Instruction-tuned QwenChart for Chart Question Answering</i><br>Viviana Ventura, Lukas Amadeus Kleybolte and Alessandra Zarcone .....                                                                                                                  | 240 |
| <i>Enhancing Scientific Visual Question Answering through Multimodal Reasoning and Ensemble Modeling</i><br>Prahitha Movva and Naga Harshita Marupaka .....                                                                                               | 252 |
| <i>The ClimateCheck Shared Task: Scientific Fact-Checking of Social Media Claims about Climate Change</i><br>Raia Abu Ahmad, Aida Usmanova and Georg Rehm .....                                                                                           | 263 |
| <i>Winning ClimateCheck: A Multi-Stage System with BM25, BGE-Reranker Ensembles, and LLM-based Analysis for Scientific Abstract Retrieval</i><br>Junjun Wang, Kunlong Chen, Zhaoqun Chen, Peng He and Wenlu Zheng .....                                   | 276 |
| <i>Comparing LLMs and BERT-based Classifiers for Resource-Sensitive Claim Verification in Social Media</i><br>Max Upravitelev, Nicolau Duran-Silva, Christian Woerle, Giuseppe Guarino, Salar Mohtaj, Jing Yang, Veronika Solopova and Vera Schmitt ..... | 281 |
| <i>AlexUNLP-FMT at ClimateCheck Shared Task: Hybrid Retrieval with Adaptive Similarity Graph-based Reranking for Climate-related Social Media Claims Fact Checking</i><br>Mahmoud Fathallah, Nagwa El-Makky and Marwan Torki .....                        | 288 |
| <i>ClimateCheck2025: Multi-Stage Retrieval Meets LLMs for Automated Scientific Fact-Checking</i><br>Anna Kiepura and Jessica Lam .....                                                                                                                    | 293 |
| <i>Overview of the SciHal25 Shared Task on Hallucination Detection for Scientific Content</i><br>Dan Li, Bogdan Palfi, Colin Zhang, Jaiganesh Subramanian, Adrian Raudaschl, Yoshiko Kakita, Anita De Waard, Zubair Afzal and Georgios Tsatsaronis .....  | 307 |
| <i>Detecting Hallucinations in Scientific Claims by Combining Prompting Strategies and Internal State Classification</i><br>Yupeng Cao, Chun-Nam Yu and K.p. Subbalakshmi .....                                                                           | 316 |

|                                                                                                                     |     |
|---------------------------------------------------------------------------------------------------------------------|-----|
| <i>A.M.P at SciHal2025: Automated Hallucination Detection in Scientific Content via LLMs and Prompt Engineering</i> |     |
| Le Nguyen Anh Khoa and Thèn Đặng Văn .....                                                                          | 328 |
| <i>SciBERT Meets Contrastive Learning: A Solution for Scientific Hallucination Detection</i>                        |     |
| Crivoi Carla and Ana Sabina Uban .....                                                                              | 336 |
| <i>Natural Language Inference Fine-tuning for Scientific Hallucination Detection</i>                                |     |
| Tim Schopf, Juraj Vladika, Michael Färber and Florian Matthes .....                                                 | 344 |
| <i>From RAG to Reality: Coarse-Grained Hallucination Detection via NLI Fine-Tuning</i>                              |     |
| Daria Galimzianova, Aleksandr Boriskin and Grigory Arshinov .....                                                   | 353 |



# Program

**Thursday, July 31, 2025**

17:00 - 09:00     *For the final SDP2025 Program Schedule, see the workshop website: <https://sdproc.org/2025/program.html>.*