

2025 IEEE/ACM International Workshop on Natural Language-Based Software Engineering (NLBSE 2025)

**Ottawa, Ontario, Canada
27 April 2025**



**IEEE Catalog Number: CFP25CD3-POD
ISBN: 979-8-3315-3865-1**

**Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***** *This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP25CD3-POD
ISBN (Print-On-Demand):	979-8-3315-3865-1
ISBN (Online):	979-8-3315-3864-4

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

2025 IEEE/ACM International Workshop on Natural Language-Based Software Engineering (NLBSE)

NLBSE 2025

Table of Contents

Message from NLBSE 2025 Program Chairs	vii
--	-----

Session 1 - Research papers

Analyzing Toxicity in Open Source Software Communications Using Psycholinguistics and Moral Foundations Theory	1
<i>Ramtin Ehsani (Drexel University, USA), Rezvaneh Rezapour (Drexel University, USA), and Preetha Chatterjee (Drexel University, USA)</i>	
SkillScope: A Tool to Predict Fine-Grained Skills Needed to Solve Issues on GitHub	9
<i>Benjamin C. Carter (Grand Canyon University, USA), Jonathan Rivas Contreras (Grand Canyon University, USA), Carlos A. Llanes Villegas (Grand Canyon University, USA), Pawan Acharya (Northern Arizona University, USA), Jack Utzerath (Grand Canyon University, USA), Adonijah O. Farner (Grand Canyon University, USA), Hunter Jenkins (Grand Canyon University, USA), Dylan Johnson (Grand Canyon University, USA), Jacob Penney (Northern Arizona University, USA), Igor Steinmacher (Northern Arizona University, USA), Marco A. Gerosa (Northern Arizona University, USA), and Fabio Santos (Colorado State University, USA)</i>	
Synthesizing Access Control Policies using Large Language Models	13
<i>Adarsh Vatsa (Stevens Institute of Technology), Pratyush Patel (Stevens Institute of Technology), and William Eiers (Stevens Institute of Technology)</i>	
Semantic API Alignment: Linking High-level User Goals to APIs	17
<i>Robert Feldt (Chalmers University of Technology) and Riccardo Coppola (Politecnico di Torino)</i>	
Evaluating the Performance and Efficiency of Sentence-BERT for Code Comment Classification.....	21
<i>Fabian C. Peña (University of Passau, Germany) and Steffen Herbold (University of Passau, Germany)</i>	

Automating Benchmark Generation for LLMs in Software Engineering: Challenges and Opportunities	25
<i>Nimrod Busany (Accenture Labs, Israel), Hananel Hadad (Accenture Labs, Israel), Gil Rosenblum (Accenture Labs, Israel), Ramya Ramachandran (Accenture Labs, Israel), Zofia Maszlanka (Avanade, Poland), Rohit Shashank Shelke (University of Ottawa, Canada), Okhaide Akhigbe (University of Ottawa, Canada), and Daniel Amyot (University of Ottawa, Canada)</i>	

Session 2 - Tool competition

The NLBSE'25 Tool Competition	27
<i>Ali Al-Kaswan (Delft University of Technology), Giuseppe Colavito (University of Bari), Nataliia Stulova (MacPaw), and Pooja Rani (University of Zurich)</i>	
Code Comment Classification with Data Augmentation and Transformer-Based Models	33
<i>Mushfiqur Rahman (Bangladesh University of Engineering and Technology) and Mohammed Latif Siddiq (University of Notre Dame)</i>	
GRAPHiC: Utilizing Graph Structures and Class Weights in Code Comment Classification with Pretrained BERT Models	37
<i>Pir Sami Ullah Shah (National University of Computer and Emerging Sciences), Shahela Saif (National University of Computer and Emerging Sciences), Muhammad Haris Athar (National University of Computer and Emerging Sciences), Muhammad Riyaan Tariq (National University of Computer and Emerging Sciences), and Abdur Rehman Afzal (National University of Computer and Emerging Sciences)</i>	
Evaluating Large Language Models in Exercises of UML Use Case Diagrams Modeling	41
<i>Giacomo Garaccione (Politecnico di Torino, Italy), Pablo Federico Vega Carrazan (Politecnico di Torino, Italy), Riccardo Coppola (Politecnico di Torino, Italy), and Luca Ardito (Politecnico di Torino, Italy)</i>	
Optimizing Deep Learning Models to Address Class Imbalance in Code Comment Classification	45
<i>Moritz Mock (Free University of Bozen-Bolzano), Thomas Borsani (Free University of Bozen-Bolzano), Giuseppe Di Fatta (Free University of Bozen-Bolzano), and Barbara Russo (Free University of Bozen-Bolzano)</i>	
CodeComClassify: Automating Code Comments Classification using BERT-based Language Models ..	49
<i>Khubaib Amjad Alam (College of Engineering, Al Ain University, UAE), Wajid Ali (National University of Computer & emerging Sciences (FAST-NUCES)), Nadeem Abbas (Linnæus University, Sweden), Muhammad Haroon (National University of Computer & emerging Sciences (FAST-NUCES)), Summan Aziz (National University of Computer & emerging Sciences (FAST-NUCES)), Meer Hasham Khan (National University of Computer & emerging Sciences (FAST-NUCES)), and Zahoor Ahmad (National University of Computer & emerging Sciences (FAST-NUCES))</i>	
Author Index	55