

2024 IEEE/ACM International Workshop on Natural Language-Based Software Engineering (NLBSE 2024)

**Lisbon, Portugal
20 April 2024**



**IEEE Catalog Number: CFP24CD3-POD
ISBN: 979-8-3503-5255-9**

**Copyright © 2024, Association for Computing Machinery
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:	CFP24CD3-POD
ISBN (Print-On-Demand):	979-8-3503-5255-9
ISBN (Online):	979-8-4007-0576-2

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

CURRAN ASSOCIATES INC.
proceedings
.com

2024 ACM/IEEE International Workshop on NL-based Software Engineering (NLBSE) **NLBSE 2024**

Table of Contents

Message from NLBSE 2024 Program Chairs	vii
NLBSE 2024 Conference Organization	viii

SESSION 1 - Language and Code Dynamics

Aligning Programming Language and Natural Language: Exploring Design Choices in Multi-Modal Transformer-Based Embedding for Bug Localization	1
<i>Partha Chakraborty (University of Waterloo), Venkatraman Arumugam (University of Waterloo), and Meiyappan Nagappan (University of Waterloo)</i>	
Unveiling Disparities: NLP Analysis of Software Industry and Vocational Education Gaps	9
<i>Emil Bäckstrand (Mid Sweden University), Rasmus Djupedal (Mid Sweden University), Lena-Maria Öberg (Mid Sweden University), and Francisco Gomes de Oliveira Neto (Chalmers University of Gothenburg)</i>	
What's in a Display Name? An Empirical Study on the Use of Display Names in Open-Source JUnit Tests	17
<i>Yining Qiao (The University of Sheffield, United Kingdom) and José Miguel Rojas (The University of Sheffield, United Kingdom)</i>	
Software Vulnerability and Functionality Assessment using LLMs	25
<i>Rasmus Ingemann Tuffoeson Jensen (JP Morgan AI Research, United Kingdom), Vali Tawosi (JP Morgan AI Research, United Kingdom), and Salwa Alamir (JP Morgan AI Research, United Kingdom)</i>	
Towards Automatic Translation of Machine Learning Visual Insights to Analytical Assertions.....	29
<i>Arumoy Shome (Delft University of Technology, Netherlands), Luis Cruz (Delft University of Technology, Netherlands), and Arie van Deursen (Delft University of Technology, Netherlands)</i>	

Tool Competition

The NLBSE'24 Tool Competition	33
<i>Rafael Kallis (Rafael Kallis Consulting, Switzerland), Giuseppe Colavito (University of Bari, Italy), Ali Al-Kaswan (Delft University of Technology, The Netherlands), Luca Pascarella (ETH Zurich, Switzerland), Oscar Chaparro (College of William & Mary, USA), and Pooja Rani (University of Zurich, Switzerland)</i>	
Few-Shot Issue Report Classification with Adapters	41
<i>Fahad Ebrahim (University of Warwick) and Mike Joy (University of Warwick)</i>	
Lessons from the NLBSE 2024 Competition: Towards Building Efficient Models for GitHub Issue Classification	45
<i>Daniel Fernando Gómez-Barrera (Universidad de los Andes, Colombia), Luccas Rojas Becerra (Universidad de los Andes, Colombia), Juan Pinzón Roncancio (Universidad de los Andes, Colombia), David Ortiz Almanza (Universidad de los Andes, Colombia), Juan Arboleda (Universidad de los Andes, Colombia), Mario Linares-Vásquez (Universidad de los Andes, Colombia), and Rubén Francisco Manrique (Universidad de los Andes, Colombia)</i>	
ClassifAI: Automating Issue Reports Classification using Pre-Trained BERT (Bidirectional Encoder Representations from Transformers) Language Models	49
<i>Khubaib Amjad Alam (National University of Computer and Emerging Sciences (FAST-NUCES) Islamabad, Pakistan), Ashish Jumani (National University of Computer and Emerging Sciences (FAST-NUCES) Islamabad, Pakistan), Harris Aamir (National University of Computer and Emerging Sciences (FAST-NUCES) Islamabad, Pakistan), and Muhammad Uzair (National University of Computer and Emerging Sciences (FAST-NUCES) Islamabad, Pakistan)</i>	
Text-To-Text Generation for Issue Report Classification	53
<i>Gokul Rejithkumar (TCS Research), Preethu Rose Anish (TCS Research), and Smita Ghaisas (TCS Research)</i>	
Applying Large Language Models to Issue Classification	57
<i>Gabriel Aracena (Grand Canyon University, USA), Kyle Luster (Grand Canyon University, USA), Fabio Santos (Grand Canyon University, USA), Igor Steinmacher (Northern Arizona University, USA), and Marco Gerosa (Northern Arizona University, USA)</i>	
Dopamin: Transformer-based Comment Classifiers through Domain Post-Training and Multi-level Layer Aggregation	61
<i>Nam Le Hai (FPT Software AI Center) and Nghi D. Q. Bui (Fulbright University)</i>	

SESSION 2 - Frontiers of Collaborative Development

Towards LLM-Generated Code Tours for Onboarding	65
<i>Balfroid Martin (NADI, University of Namur), Benoît Vanderose (NADI, University of Namur), and Xavier Devroey (NADI, University of Namur)</i>	

Automated Extraction of Compliance Elements in Software Engineering Contracts Using Natural Language Generation	69
<i>Gokul Rejithkumar (TCS Research), Preethu Rose Anish (TCS Research), Pratik Sonar (TCS Research), and Smita Ghaisas (TCS Research)</i>	
Emotion Classification In Software Engineering Texts: A Comparative Analysis of Pre-trained Transformers Language Models	73
<i>Mia Mohammad Imran (Virginia Commonwealth University)</i>	
Understanding Emojis in Useful Code Review Comments	81
<i>Sharif Ahmed (Boise State University) and Nasir U. Eisty (Boise State University)</i>	
Author Index	85