

2022 IEEE/ACM 1st International Workshop on Natural Language-Based Software Engineering (NLBSE 2022)

**Pittsburgh, Pennsylvania, USA
8 May 2022**



**IEEE Catalog Number: CFP22CD3-POD
ISBN: 978-1-6654-6231-0**

**Copyright © 2022, 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:	CFP22CD3-POD
ISBN (Print-On-Demand):	978-1-6654-6231-0
ISBN (Online):	978-1-4503-9343-0

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

The 1st Intl. Workshop on Natural Language-based Software Engineering **NLBSE 2022**

Table of Contents

Message from the NLBSE 2022 Program Chairs	vii
NLBSE 2022 Program Committee	ix

The First Intl. Workshop on Natural Language-based Software Engineering

Unsupervised Extreme Multi Label Classification of Stack Overflow Posts	1
<i>Peter Devine (The University of Auckland, New Zealand) and Kelly Blincoe (The University of Auckland, New Zealand)</i>	
Understanding Digits in Identifier Names: An Exploratory Study	9
<i>Anthony Peruma (Rochester Institute of Technology, USA) and Christian D. Newman (Rochester Institute of Technology, USA)</i>	
From Zero to Hero: Generating Training Data for Question-To-Cypher Models	17
<i>Opitz Dominik (Hochschule Bonn-Rhein-Sieg, Germany) and Hochgeschwender Nico (German Aerospace Center (DLR), Germany)</i>	
Automatic Identification of Informative Code in Stack Overflow Posts	21
<i>Preetha Chatterjee (Drexel University, USA)</i>	
NLBSE'22 Tool Competition	25
<i>Rafael Kallis (Rafael Kallis Consulting, Switzerland), Oscar Chaparro (College of William & Mary, USA), Andrea Di Sorbo (University of Sannio, Italy), and Sebastiano Panichella (Zurich University of Applied Sciences, Switzerland)</i>	
Issue Report Classification using Pre-Trained Language Models	29
<i>Giuseppe Colavito (University of Bari, Italy), Filippo Lanubile (University of Bari, Italy), and Nicole Novielli (University of Bari, Italy)</i>	
BERT-Based GitHub Issue Report Classification	33
<i>Mohammed Latif Siddiq (University of Notre Dame, USA) and Joanna C. S. Santos (University of Notre Dame, USA)</i>	
Predicting Issue Types with seBERT	37
<i>Alexander Trautsch (University of Goettingen, Germany) and Steffen Herbold (TU Clausthal, Germany)</i>	

GitHub Issue Classification using BERT-Style Models	40
<i>Shikhar Bharadwaj (Indian Institute of Science (IISc), India) and Tushar Kadam (Indian Institute of Science (IISc), India)</i>	
CatIss: An Intelligent Tool for Categorizing Issues Reports using Transformers	44
<i>Maliheh Izadi (Delft University of Technology, Netherlands)</i>	
On the Evaluation of NLP-Based Models for Software Engineering	48
<i>Maliheh Izadi (Delft University of Technology, Netherlands) and Matin Nili Ahmadabadi (University of Tehran, Iran)</i>	
Identification of Intra-Domain Ambiguity using Transformer-Based Machine Learning	51
<i>Ambarish Moharil (Eindhoven University of Technology, The Netherlands) and Arpit Sharma (Indian Institute of Science Education and Research Bhopal, India)</i>	
Can NMT Understand Me? Towards Perturbation-Based Evaluation of NMT Models for Code Generation	59
<i>Pietro Liguori (University of Naples Federico II, Italy), Cristina Improta (University of Naples Federico II, Italy), Simona De Vivo (University of Naples Federico II, Italy), Roberto Natella (University of Naples Federico II, Italy), Bojan Cukic (University of North Carolina at Charlotte, Italy), and Domenico Cotroneo (University of Naples Federico II, Italy)</i>	
Supporting Systematic Literature Reviews using Deep-Learning-Based Language Models	67
<i>Rand Alchokr (Otto-von-Guericke University, Germany), Manoj Borkar (Otto-von-Guericke University, Germany), Sharanya Thotadarya (Otto-von-Guericke University, Germany), Gunter Saake (Otto-von-Guericke University, Germany), and Thomas Leich (Harz University & METOP GmbH, Germany)</i>	
Story Point Effort Estimation by Text Level Graph Neural Network	75
<i>Hung Phan (Iowa State University, USA) and Ali Jannesari (Iowa State University, USA)</i>	
Author Index	79