

2025 IEEE/ACM 22nd International Conference on Mining Software Repositories (MSR 2025)

**Ottawa, Ontario, Canada
28-29 April 2025**



**IEEE Catalog Number: CFP2578C-POD
ISBN: 979-8-3315-0184-6**

**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:	CFP2578C-POD
ISBN (Print-On-Demand):	979-8-3315-0184-6
ISBN (Online):	979-8-3315-0183-9
ISSN:	2574-3848

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

2025 IEEE/ACM 22nd International Conference on Mining Software Repositories (MSR) **MSR 2025**

Table of Contents

Message from the General, Program, and Junior PC Chairs	xviii
Message from the Data and Tool Showcase Track Co-Chairs	xxii
Message from the Industry Track Co-Chairs	xxiii
Message from the Mining Challenge Co-Chairs	xxiv
Message from the Registered Reports Track Co-Chairs	xxvi
Message from the Tutorials Track Co-Chairs	xxviii
Message from the Vision and Reflection Track Co-Chairs	xxix
Organizing Committee	xxx
Program Committees	xxxii
Keynote	xli
Vision and Reflection Track Abstracts	xlii
Tutorials	xliii

Defects, bugs, and issues

Learning from Mistakes: Understanding Ad-Hoc Logs through Analyzing Accidental Commits	1
<i>Yi-Hung Chou (University of California, Irvine, USA), Yiyang Min (Amazon, Canada), April Yi Wang (ETH Zürich, Switzerland), and James A. Jones (University of California, Irvine, USA)</i>	
On the Calibration of Just-in-Time Defect Prediction	14
<i>Khulja Shahini (University of Duisburg Essen, Germany), Jone Bartel (University of Duisburg Essen, Germany), and Klaus Pohl (University of Duisburg Essen, Germany)</i>	
An Empirical Study on Leveraging Images in Automated Bug Report Reproduction	27
<i>Dingbang Wang (University of Connecticut, USA), Zhaoxu Zhang (University of Southern California, USA), Sidong Feng (Monash University, Australia), William G. J. Halfond (University of Southern California, USA), and Tingting Yu (University of Connecticut, USA)</i>	
It's About Time: An Empirical Study of Date and Time Bugs in Open-Source Python Software	39
<i>Shrey Tiwari (Carnegie Mellon University, USA), Serena Chen (University of California San Diego, USA), Alexander Joukov (Stony Brook University, USA), Peter Vandervelde (University of California, Santa Barbara, USA), Ao Li (Carnegie Mellon University, USA), and Rohan Padhye (Carnegie Mellon University, USA)</i>	

Enhancing Just-In-Time Defect Prediction Models with Developer-Centric Features	52
<i>Emanuela Guglielmi (University of Molise, Italy), Andrea D'Aguzzo (University of Molise, Italy), Rocco Oliveto (University of Molise, Italy), and Simone Scalabrino (University of Molise, Italy)</i>	
Revisiting Defects4J for Fault Localization in Diverse Development Scenarios	63
<i>Md Nakhla Rafi (Software Performance, Analysis, and Reliability (SPEAR) Lab, Concordia University, Montreal, Quebec, Canada), An Ran Chen (University of Alberta, Edmonton, Canada), Tse-Hsun Chen (Software Performance, Analysis, and Reliability (SPEAR) Lab, Concordia University, Montreal, Quebec, Canada), and Shaohua Wang (Central University of Finance and Economics, Beijing, China)</i>	
Mining Bug Repositories for Multi-Fault Programs	76
<i>Dylan Callaghan (Stellenbosch University, South Africa) and Bernd Fischer (Stellenbosch University, South Africa)</i>	
HaPy-Bug - Human Annotated Python Bug Resolution Dataset	81
<i>Piotr Przymus (Nicolaus Copernicus University, Poland), Mikołaj Fejzer (Nicolaus Copernicus University, Poland), Jakub Narebski (Nicolaus Copernicus University, Poland), Radosław Woźniak (Nicolaus Copernicus University, Poland), Łukasz Halada (University of Wrocław, Poland), Aleksander Kazecki (Nicolaus Copernicus University, Poland), Mykhailo Molchanov (Kyiv Polytechnic Institute, Ukraine), and Krzysztof Stencel (University of Warsaw, Poland)</i>	
SPRINT: An Assistant for Issue Report Management	86
<i>Ahmed Adnan (University of Dhaka, Bangladesh), Antu Saha (William & Mary, USA), and Oscar Chaparro (William & Mary, USA)</i>	

Security and legal aspects

Wolves in the Repository: A Software Engineering Analysis of the XZ Utils Supply Chain Attack	91
<i>Piotr Przymus (Nicolaus Copernicus University in Toruń, Poland) and Thomas Durieux (TU Delft & Endor Labs, The Netherlands)</i>	
Software Composition Analysis and Supply Chain Security in Apache Projects: An Empirical Study	103
<i>Sabato Nocera (University of Salerno, Italy), Sira Vegas (Universidad Politécnica de Madrid, Spain), Giuseppe Scanniello (University of Salerno, Italy), and Natalia Juristo (Universidad Politécnica de Madrid, Spain)</i>	
Good Practice Versus Reality: A Landscape Analysis of Research Software Metadata Adoption in European Open Science Clusters	116
<i>Anas El Hounsri (University Polytechnic of Madrid, Spain) and Daniel Garijo Verdejo (University Polytechnic of Madrid, Spain)</i>	
From Industrial Practices to Academia: Uncovering the Gap in Vulnerability Research and Practice	129
<i>Zhuang Liu (Zhejiang University, China), Xing Hu (Zhejiang University, China), Jiayuan Zhou (Queen's University, Canada), and Xin Xia (Zhejiang University, China)</i>	

Patch Me If You Can — Securing the Linux Kernel	142
<i>Gunnar Kudrjavets (Amazon Web Services, USA)</i>	
OSS License Identification at Scale: A Comprehensive Dataset using World of Code	144
<i>Mahmoud Jahanshahi (University of Tennessee USA), David Reid (University of Tennessee, USA), Adam McDaniel (University of Tennessee, USA), and Audris Mockus (University of Tennessee, USA)</i>	
SCRUBD: Smart Contracts Reentrancy and Unhandled Exceptions Vulnerability Dataset	149
<i>Sujeet Chavhan Yashavant (Indian Institute of Technology Kanpur, India), MitrajSinh Chavda (Indian Institute of Technology Kanpur, India), Saurabh Kumar (Indian Institute of Technology Hyderabad, India), Amey Karkare (Indian Institute of Technology Kanpur, India), and Angshuman Karmakar (Indian Institute of Technology Kanpur, India)</i>	
ICVul: A Well-Labeled C/C++ Vulnerability Dataset with Comprehensive Metadata and VCCs	154
<i>Chaomeng Lu (KU Leuven, Belgium), Tianyu Li (KU Leuven, Belgium), Toon Dehaene (KU Leuven, Belgium), and Bert Lagaisse (KU Leuven, Belgium)</i>	
A Dataset of Contributor Activities in the NumFocus Open-Source Community	159
<i>Youness Hourri (University of Mons, Belgium), Alexandre Decan (F.R.S.-FNRS Research Associate), and Tom Mens (University of Mons, Belgium)</i>	
Wild SBOMs: A Large-Scale Dataset of Software Bills of Materials from Public Code	164
<i>Luís Soeiro (LTCl, Télécom Paris, Institut Polytechnique de Paris, France), Thomas Robert (LTCl, Télécom Paris, Institut Polytechnique de Paris, France), and Stefano Zacchiroli (LTCl, Télécom Paris, Institut Polytechnique de Paris, France)</i>	
MaLAWare: Automating the Comprehension of Malicious Software Behaviours using Large Language Models (LLMs)	169
<i>Bikash Saha (Indian Institute of Technology Kanpur, India), Nanda Rani (Indian Institute of Technology Kanpur, India), and Sandeep Kumar Shukla (Indian Institute of Technology Kanpur, India)</i>	

AI for SE (1)

Combining Large Language Models with Static Analyzers for Code Review Generation	174
<i>Imen Jaoua (Université de Montréal, Canada), Oussama Ben Sghaier (Université de Montréal, Canada), and Houari Sahraoui (Université de Montréal, Canada)</i>	
Harnessing Large Language Models for Curated Code Reviews	187
<i>Oussama Ben Sghaier (Université de Montréal, Canada), Martin Weyssow (Singapore Management University, Singapore), and Houari Sahraoui (Université de Montréal, Canada)</i>	
SMATCH-M-LLM: Semantic Similarity in Metamodel Matching with Large Language Models	199
<i>Nafisa Ahmed (Polytechnique Montréal, Canada), Hin Chi Kwok (Polytechnic University, China), Mohammad Hamdaqa (Polytechnique Montréal, USA), and Wesley K. G. Assunção (North Carolina State University, USA)</i>	

How Effective are LLMs for Data Science Coding? A Controlled Experiment	211
<i>Nathalia Nascimento (Pennsylvania State University, USA), Everton Guimaraes (Pennsylvania State University, USA), Sai Sanjna Chintakunta (Pennsylvania State University, USA), and Santhosh Anitha Boominathan (Pennsylvania State University, USA)</i>	
Do LLMs Provide Links to Code Similar to What They Generate? A Study with Gemini and Bing CoPilot	223
<i>Daniele Bifulco (University of Sannio, Italy), Pietro Cassieri (University of Salerno, Italy), Giuseppe Scanniello (University of Salerno, Italy), Massimiliano Di Penta (University of Sannio, Italy), and Fiorella Zampetti (University of Sannio, Italy)</i>	
Too Noisy to Learn: Enhancing Data Quality for Code Review Comment Generation	236
<i>Chunhua Liu (The University of Melbourne, Australia), Hong Yi Lin (The University of Melbourne, Australia), and Patanamon Thongtanunam (The University of Melbourne, Australia)</i>	
Should Code Models Learn Pedagogically? A Preliminary Evaluation of Curriculum Learning for Real-World Software Engineering Tasks	249
<i>Kyi Shin Khant (The University of Melbourne), Hong Yi Lin (The University of Melbourne), and Patanamon Thongtanunam (The University of Melbourne)</i>	
RepoChat: An LLM-Powered Chatbot for GitHub Repository Question-Answering	255
<i>Samuel Abedu (Concordia University, Canada), Laurine Menneron (CESI Graduate School of Engineering, France), SayedHassan Khatoonabadi (Concordia University, Canada), and Emad Shihab (Concordia University, Canada)</i>	

MSR 2025 Mining Challenge

Analyzing Dependency Clusters and Security Risks in the Maven Central Repository	260
<i>George Lake (Idaho State University, USA) and Minhaz F. Zibran (Idaho State University, USA)</i>	
Chasing the Clock: How Fast are Vulnerabilities Fixed in the Maven Ecosystem?	265
<i>Fazle Rabbi (Idaho State University, USA), Arifa Islam Champa (Idaho State University, USA), Rajshakhar Paul (Idaho State University, USA), and Minhaz F. Zibran (Idaho State University, USA)</i>	
Decoding Dependency Risks: A Quantitative Study of Vulnerabilities in the Maven Ecosystem	270
<i>Costain Nachuma (Idaho State University, USA), Mosharaf Hossan (Idaho State University, USA), Asif K. Turzo (Idaho State University, USA), and Minhaz F. Zibran (Idaho State University, USA)</i>	
Faster Releases, Fewer Risks: A Study on Maven Artifact Vulnerabilities and Lifecycle Management	275
<i>Shafiullah Shafin (Rajshahi University of Engineering & Technology, Bangladesh), Fazle Rabbi (Idaho State University, USA), S. M. Mahedy Hasan (Rajshahi University of Engineering & Technology, Bangladesh), and Minhaz F. Zibran (Idaho State University, USA)</i>	

Insights into Dependency Maintenance Trends in the Maven Ecosystem	280
<i>Barisha Chowdhury (Rajshahi University of Engineering & Technology, Bangladesh), Fazle Rabbi (Idaho State University, USA), S. M. Mahedy Hasan (Rajshahi University of Engineering & Technology, Bangladesh), and Minhaz F. Zibran (Idaho State University, USA)</i>	
Insights into Vulnerability Trends in Maven Artifacts: Recurrence, Popularity, and User Behavior	285
<i>Courtney Bodily (Idaho State University, USA), Eric Hill (Idaho State University, USA), Andreas Kramer (Idaho State University, USA), Leslie Kerby (Idaho State University, USA), and Minhaz Zibran (Idaho State University, USA)</i>	
Understanding Software Vulnerabilities in the Maven Ecosystem: Patterns, Timelines, and Risks	290
<i>Fazle Rabbi (Idaho State University, USA), Rajshakhar Paul (Idaho State University, USA), Arifa Islam Champa (Idaho State University, USA), and Minhaz F. Zibran (Idaho State University, USA)</i>	
Dependency Update Adoption Patterns in the Maven Software Ecosystem	295
<i>Baltasar Berretta (College of Wooster, USA), Augustus Thomas (College of Wooster, USA), and Heather Guarnera (College of Wooster, USA)</i>	
Analyzing Vulnerability Overestimation in the Maven Ecosystem	300
<i>Taha Draoui (University of Michigan-Flint, USA), Faten Jebari (University of Michigan-Flint, USA), Chawki Ben Slimen (University of Michigan-Flint, USA), Munjaap Uppal (University of Michigan-Flint, USA), and Mohamed Wiem Mkaouer (University of Michigan-Flint, USA)</i>	
Dependency Dilemmas: A Comparative Study of Independent and Dependent Artifacts in Maven Central Ecosystem	304
<i>Mehedi Hasan Shanto (Khulna University, Bangladesh), Muhammad Asaduzzaman (University of Windsor, Canada), Manishankar Mondal (Khulna University, Bangladesh), and Shaiful Chowdhury (University of Manitoba, Canada)</i>	
Cascading Effects: Analyzing Project Failure Impact in the Maven Central Ecosystem	309
<i>Mina Shehata (Belmont University, USA), Saidmakhmud Makhkamjonov (Belmont University, USA), Mahad Syed (Belmont University, USA), and Esteban Parra (Belmont University, USA)</i>	
Do Developers Depend on Deprecated Library Versions? A Mining Study of Log4j	314
<i>Haruhiko Yoshioka (Nara Institute of Science and Technology), Sila Lertbanjongngam (Nara Institute of Science and Technology), Masayuki Inaba (Nara Institute of Science and Technology), Youmei Fan (Nara Institute of Science and Technology), Takashi Nakano (Nara Institute of Science and Technology), Kazumasa Shimari (Nara Institute of Science and Technology), Raula Gaikovina Kula (Osaka University), and Kenichi Matsumoto (Nara Institute of Science and Technology)</i>	

Mining for Lags in Updating Critical Security Threats: A Case Study of Log4j Library	319
<i>Hidetake Tanaka (Nara Institute of Science and Technology, Japan), Kazuma Yamasaki (Nara Institute of Science and Technology, Japan), Momoka Hirose (Nara Institute of Science and Technology, Japan), Takashi Nakano (Nara Institute of Science and Technology, Japan), Youmei Fan (Nara Institute of Science and Technology, Japan), Kazumasa Shimari (Nara Institute of Science and Technology, Japan), Raula Gaikovina Kula (Osaka University, Japan), and Kenichi Matsumoto (Nara Institute of Science and Technology, Japan)</i>	
On the Evolution of Unused Dependencies in Java Project Releases: An Empirical Study	324
<i>Nabhan Suwanachote (Nara Institute of Science and Technology, Japan), Yagut Shakizada (Nara Institute of Science and Technology, Japan), Yutaro Kashiwa (Nara Institute of Science and Technology, Japan), Bin Lin (Hangzhou Dianzi University, China), and Hajimu Iida (Nara Institute of Science and Technology, Japan)</i>	
Out of Sight, Still at Risk: The Lifecycle of Transitive Vulnerabilities in Maven	329
<i>Piotr Przymus (Nicolaus Copernicus University in Toruń, Poland), Mikołaj Fejzer (Nicolaus Copernicus University in Toruń, Poland), Jakub Narębski (Nicolaus Copernicus University in Toruń, Poland), Krzysztof Rykaczewski (Nicolaus Copernicus University in Toruń, Poland), and Krzysztof Stencel (University of Warsaw, Poland)</i>	
Popularity and Innovation in Maven Central	334
<i>Nkiru Ede (Victoria University of Wellington, New Zealand), Jens Dietrich (Victoria University of Wellington, New Zealand), and Ulrich Zuelicke (Victoria University of Wellington, New Zealand)</i>	
Software Bills of Materials in Maven Central	339
<i>Yogya Gamage (Université de Montréal), Nadia Gonzalez Fernandez (Université de Montréal), Martin Monperrus (KTH Royal Institute of Technology), and Benoit Baudry (Université de Montréal)</i>	
The Ripple Effect of Vulnerabilities in Maven Central: Prevalence, Propagation, and Mitigation Challenges	344
<i>Ehtisham Ul Haq (York University, Canada), Song Wang (York University, Canada), and Robert S. Allison (York University, Canada)</i>	
Tracing Vulnerabilities in Maven: A Study of CVE Lifecycles and Dependency Networks	349
<i>Corey Yang-Smith (University of Calgary, Canada) and Ahmad Abdellatif (University of Calgary, Canada)</i>	
Understanding Abandonment and Slowdown Dynamics in the Maven Ecosystem	354
<i>Kazi Amit Hasan (Queen's University, Canada), Jerin Yasmin (Queen's University, Canada), Huizi Hao (Queen's University, Canada), Yuan Tian (Queen's University, Canada), Safwat Hassan (University of Toronto, Canada), and Steven H. H. Ding (McGill University, Canada)</i>	
Characterizing Packages for Vulnerability Prediction	359
<i>Saviour Owolabi (University of Calgary, Canada), Francesco Rosati (University of Calgary, Canada), Ahmad Abdellatif (University of Calgary, Canada), and Lorenzo De Carli (University of Calgary, Canada)</i>	

Understanding the Popularity of Packages in Maven Ecosystem	364
<i>Sadman Jashim Sakib (University of Windsor, Canada), Muhammad Asaduzzaman (University of Windsor, Canada), Curtis Bright (University of Windsor, Canada), and Cole Morgan (University of Windsor, Canada)</i>	
Navigating and Exploring Software Dependency Graphs Using Goblin	369
<i>Damien Jaime (Sorbonne Université, CNRS, LIP6, Université Paris Nanterre), Joyce El Haddad (Université Paris Dauphine - PSL, CNRS, LAMSADE), and Pascal Poizat (Sorbonne Université, CNRS, LIP6, Université Paris Nanterre)</i>	

Software evolution and analysis

50 Years of Programming Language Evolution through the Software Heritage Looking Glass	372
<i>Adèle Desmazières (Sorbonne University, France), Roberto Di Cosmo (Inria and University Paris Cité, France), and Valentin Lorentz (Inria Foundation and Inria, France)</i>	
It Works (only) on My Machine: A Study on Reproducibility Smells in Ansible Scripts	384
<i>Ghazal Sobhani (Dalhousie University, Canada), Israat Haque (Dalhousie University, Canada), and Tushar Sharma (Dalhousie University, Canada)</i>	
Are the Majority of Public Computational Notebooks Pathologically Non-Executable?	396
<i>Tien Nguyen (Virginia Tech, USA), Waris Gill (Virginia Tech, USA), and Muhammad Ali Gulzar (Virginia Tech, USA)</i>	
Understanding Test Deletion in Java Applications	408
<i>Suraj Bhatta (North Dakota State University, USA), Frank Kendemah (North Dakota State University, USA), and Ajay Kumar Jha (North Dakota State University, USA)</i>	
A Public Benchmark of REST APIs	421
<i>Alix Decrop (University of Namur, Belgium), Sara Eraso (University of Valle, Colombia), Xavier Devroey (University of Namur, Belgium), and Gilles Perrouin (University of Namur, Belgium)</i>	
What Do Contribution Guidelines Say About Software Testing?	434
<i>Bruna Falcucci (UFMG, Brazil), Felipe Gomide (UFMG, Brazil), and Andre Hora (UFMG, Brazil)</i>	
Measuring InnerSource Value	439
<i>Chamindra de Silva (Citibank, United Kingdom) and Daniel Izquierdo-Cortázar (Bitergia, Spain)</i>	
CoUpJava: A Dataset of Code Upgrade Histories in Open-Source Java Repositories	441
<i>Kaihang Jiang (University of Waterloo, Canada), Bihui Jin (University of Waterloo, Canada), and Pengyu Nie (University of Waterloo, Canada)</i>	
EvoChain: A Framework for Tracking and Visualizing Smart Contract Evolution	446
<i>Ilham Qasse (Reykjavik University, Iceland), Mohammad Hamdaqa (Polytechnique Montreal, Canada), and Björn Þór Jónsson (Reykjavik University, Iceland)</i>	
CoDocBench: A Dataset for Code-Documentation Alignment in Software Maintenance	451
<i>Kunal Pai (University of California, Davis, USA), Premkumar Devanbu (University of California, Davis, USA), and Toufique Ahmed (University of California, Davis, USA)</i>	

RefExpo: Unveiling Software Project Structures through Advanced Dependency Graph Extraction	456
<i>Vahid Haratian (Bilkent University, Turkey), Pouria Derakhshanfar (JetBrains Research, The Netherlands), Vladimir Kovalenko (JetBrains Research, The Netherlands), and Eray Tuzün (Bilkent University, Turkey)</i>	
HyperAST: Incrementally Mining Large Source Code Repositories	461
<i>Quentin Le Dilairec (Delft University of Technology, The Netherlands) and Andy Zaidman (Delft University of Technology, The Netherlands)</i>	

LLMs for Code

How Much Do Code Language Models Remember? An Investigation on Data Extraction Attacks Before and After Fine-Tuning	465
<i>Fabio Salerno (Delft University of Technology, The Netherlands), Ali Al-Kaswan (Delft University of Technology, The Netherlands), and Maliheh Izadi (Delft University of Technology, The Netherlands)</i>	
Can LLMs Generate Higher Quality Code Than Humans? An Empirical Study	478
<i>Mohammad Talal Jamil (Lahore University of Management Sciences, Pakistan), Shamsa Abid (National University of Computer and Emerging Sciences, Pakistan), and Shafay Shamil (Lahore University of Management Sciences, Pakistan)</i>	
Prompt Engineering or Fine-Tuning: An Empirical Assessment of LLMs for Code	490
<i>Jiho Shin (York University, Canada), Clark Tang (n/a), Tahmineh Mohati (University of Calgary, Canada), Maleknaz Nayebe (York University, Canada), Song Wang (York University, Canada), and Hadi Hemmati (York University, Canada)</i>	
Drawing Pandas: A Benchmark for LLMs in Generating Plotting Code	503
<i>Timur Galimzyanov (JetBrains Research), Sergey Titov (JetBrains Research), Yaroslav Golubev (JetBrains Research), and Egor Bogomolov (JetBrains Research; Delft University of Technology)</i>	
SnipGen: A Mining Repository Framework for Evaluating LLMs for Code	508
<i>Daniel Rodriguez-Cardenas (William & Mary, VA), Alejandro Velasco (William & Mary, VA), and Denys Poshyvanyk (William & Mary, VA)</i>	

Software ecosystems and humans

The Ecosystem of Open-Source Music Production Software – A Mining Study on the Development Practices of VST Plugins on GitHub	513
<i>Bogdan Andrei (University Of Amsterdam, The Netherlands), Mauricio Verano Merino (Vrije Universiteit Amsterdam, The Netherlands), and Ivano Malavolta (Vrije Universiteit Amsterdam, The Netherlands)</i>	
Can LLMs Replace Manual Annotation of Software Engineering Artifacts?	526
<i>Toufique Ahmed (University of California, USA; IBM Research, Yorktown Heights, USA; University of Stuttgart, Germany), Premkumar Devanbu (University of California, USA), Christoph Treude (Singapore Management University, Singapore), and Michael Pradel (University of Stuttgart, Germany)</i>	

Investigating the Understandability of Review Comments on Code Change Requests	539
<i>Shamimur Rahman (University of Saskatchewan, Canada), Zadia Codabux (University of Saskatchewan, Canada), and Chanchal K. Roy (University of Saskatchewan, Canada)</i>	
Mining a Decade of Event Impacts on Contributor Dynamics in Ethereum: A Longitudinal Study..	552
<i>Matteo Vaccarigu (University of Cagliari, Italy; Brunel University of London, UK), Sabrina Aufiero (University College London, UK), Cheik Ba (Queen Mary University of London, UK), Silvia Bartolucci (University College London, UK), Richard Clegg (Queen Mary University of London, UK), Daniel Graziotin (University of Hohenheim, Germany), Romyana Neykova (Brunel University of London, UK), Roberto Tonelli (University of Cagliari, Italy), and Giuseppe Destefanis (Brunel University of London, UK)</i>	
Is it Really Fun? Detecting Low Engagement Events in Video Games	564
<i>Emanuela Guglielmi (University of Molise, Italy), Gabriele Bavota (Università della Svizzera Italiana, Switzerland), Nicole Novielli (University of Bari), Rocco Oliveto (University of Molise, Italy), and Simone Scalabrino (University of Molise, Italy)</i>	
A Dataset of Software Bill of Materials for Evaluating SBOM Consumption Tools	576
<i>Rio Kishimoto (The University of Osaka, Japan), Tetsuya Kanda (Notre Dame Seishin University, Japan), Yuki Manabe (The University of Fukuchiyama, Japan), Katsuro Inoue (Nanzan University, Japan), Shi Qiu (Toshiba Corporation, Japan), and Yoshiki Higo (The University of Osaka, Japan)</i>	
Jupyter Notebook Activity Dataset	581
<i>Tomoki Nakamaru (The University of Tokyo, Japan), Tomomasa Matsunaga (The University of Tokyo, Japan), and Tetsuro Yamazaki (The University of Tokyo, Japan)</i>	
CoPhi - Mining C/C++ Packages for Conan Ecosystem Analysis	586
<i>Vivek Sarkar (University of Washington, USA), Anemone Kampkötter (Technische Universität Dortmund, Germany), and Ben Hermann (Technische Universität Dortmund, Germany)</i>	
MARIN: A Research-Centric Interface for Querying Software Artifacts on Maven Repositories	591
<i>Johannes Düsing (TU Dortmund University, Germany), Jared Chiaramonte (Arizona State University, USA), and Ben Hermann (TU Dortmund University, Germany)</i>	
GitProjectHealth: An Extensible Framework for Git Social Platform Mining	596
<i>Nicolas Hlad (Berger-Levrault, France), Benoit Verhaeghe (Berger-Levrault, France), and Kilian Bauvent (Berger-Levrault, France)</i>	
MYRIAD PEOPLE Open Source Software for New Media Arts	601
<i>Benoit Baudry (Université de Montréal, Canada), Erik Natanael Gustafsson (Independent Artists, Sweden), Roni Kaufman (Independent Artists, Sweden), and Maria Kling (Independent Artists, Sweden)</i>	

OpenMent: A Dataset of Mentor-Mentee Interactions in Google Summer of Code	606
<i>Erfan Raoofian (University of British Columbia, Canada), Fatemeh H. Fard (University of British Columbia, Canada), Ifeoma Adaji (University of British Columbia, Canada), and Gema Rodríguez-Pérez (University of British Columbia, Canada)</i>	
Under the Blueprints: Parsing Unreal Engine's Visual Scripting at Scale	611
<i>Kalvin Eng (University of Alberta, Canada) and Abram Hindle (University of Alberta, Canada)</i>	

Build systems and DevOps

Build Code Needs Maintenance Too: A Study on Refactoring and Technical Debt in Build Systems	616
<i>Anwar Ghammam (Oakland University, USA), Dhia Elhaq Rzig (University of Michigan- Dearborn, USA), Mohamed Almukhtar (University of Michigan- Flint, USA), Rania Khalsi (University of Michigan- Flint, USA), Foyzul Hassan (University of Michigan- Dearborn, USA), and Marouane Kessentini (Grand Valley State University, USA)</i>	
LLMSecConfig: An LLM-Based Approach for Fixing Software Container Misconfigurations	629
<i>Ziyang Ye (CREST - The Centre for Research on Engineering Software Technologies, Australia; The University of Adelaide, Australia), Triet Huynh Minh Le (CREST - The Centre for Research on Engineering Software Technologies, Australia; The University of Adelaide, Australia), and M. Ali Babar (CREST - The Centre for Research on Engineering Software Technologies, Australia; The University of Adelaide, Australia)</i>	
How Do Infrastructure-as-Code Practitioners Update Their Dependencies? An Empirical Study on Terraform Module Updates	642
<i>Mahi Begoug (University of Quebec, Canada), Ali Ouni (University of Quebec, Canada), and Moataz Chouchen (Concordia University, Canada)</i>	
TerraDS: A Dataset for Terraform HCL Programs	654
<i>Christoph Bühler (University of St. Gallen, Switzerland), David Spielmann (University of St. Gallen, Switzerland), Roland Meier (Armasuisse, Switzerland), and Guido Salvaneschi (University of St. Gallen, Switzerland)</i>	
OSPtrack: A Labeled Dataset Targeting Simulated Execution of Open-Source Software	659
<i>Zhuoran Tan (University of Glasgow, United Kingdom), Christos Anagnostopoulos (University of Glasgow, United Kingdom), and Jeremy Singer (University of Glasgow, United Kingdom)</i>	
CARDS: A Collection of Package, Revision, and Miscellaneous Dependency Graphs	664
<i>Euxane Tran-Girard (Uni. Gustave Eiffel, France), Laurent Bulteau (Uni. Gustave Eiffel, France), and Pierre-Yves David (Octobus S.c.o.p., France)</i>	
GHALogs: Large-Scale Dataset of GitHub Actions Runs	669
<i>Florent Moriconi (EURECOM, AMADEUS), Thomas Durieux (TU Delft), Jean-Rémy Falleri (University of Bordeaux, LaBRI, UMR 5800), Raphaël Troncy (EURECOM), and Aurélien Francillon (EURECOM)</i>	

AI for SE (2)

Automatic High-Level Test Case Generation using Large Language Models	674
<i>Navid Bin Hasan (Bangladesh University of Engineering and Technology, Bangladesh), Ashraful Islam (Bangladesh University of Engineering and Technology, Bangladesh), Junaed Younus Khan (Bangladesh University of Engineering and Technology, Bangladesh), Sanjida Senjik (Bangladesh University of Engineering and Technology, Bangladesh), and Anindya Iqbal (Bangladesh University of Engineering and Technology, Bangladesh)</i>	
Prompting in the Wild: An Empirical Study of Prompt Evolution in Software Repositories	686
<i>Mahan Tafreshipour (University of California, Irvine, USA), Aaron Imani (University of California, Irvine, USA), Eric Huang (University of California, Irvine, USA), Eduardo Santana de Almeida (Federal University of Bahia, Brazil), Thomas Zimmermann (University of California, Irvine, USA), and Iftekhar Ahmed (University of California, Irvine, USA)</i>	
Towards Detecting Prompt Knowledge Gaps for Improved LLM-Guided Issue Resolution	699
<i>Ramtin Ehsani (Drexel University, USA), Sakshi Pathak (Drexel University, USA), and Preetha Chatterjee (Drexel University, USA)</i>	
Intelligent Semantic Matching (ISM) for Video Tutorial Search using Transformer Models	712
<i>Ahmad J. Tayeb (King Abdulaziz University, Saudi Arabia) and Sonia Haiduc (Florida State University, USA)</i>	
Language Models in Software Development Tasks: An Experimental Analysis of Energy and Accuracy	725
<i>Negar Alizadeh (Utrecht University, The Netherlands), Boris Belchev (University of Twente, The Netherlands), Nishant Saurabh (Utrecht University, The Netherlands), Patricia Kelbert (Fraunhofer IESE, Germany), and Fernando Castor (University of Twente, The Netherlands)</i>	
TriGraph: A Probabilistic Subgraph-Based Model for Visual Code Completion in Pure Data	737
<i>Anisha Islam (University of Alberta, Canada) and Abram Hindle (University of Alberta, Canada)</i>	
Inferring Questions from Programming Screenshots	750
<i>Faiz Ahmed (York University, Canada), Xuchen Tan (York University, Canada), Folajinmi Adewole (York University, Canada), Suprakash Datta (York University, Canada), and Maleknaz Nayeibi (York University, Canada)</i>	
Human-In-The-Loop Software Development Agents: Challenges and Future Directions	756
<i>Jirat Pasuksmit (Atlassian, Australia), Wannita Takerngsaksiri (Monash University, Australia), Patanamon Thongtanunam (The University of Melbourne, Australia), Chakkrit Tantithamthavorn (Monash University, Australia), Ruixiong Zhang (Atlassian, Australia), Shiyang Wang (Atlassian, Australia), Fan Jiang (Atlassian, Australia), Jing Li (Atlassian, Australia), Evan Cook (Atlassian, Australia), Kun Chen (Atlassian, Australia), and Ming Wu (Atlassian, Australia)</i>	
FormalSpecCpp: A Dataset of C++ Formal Specifications Created using LLMs	758
<i>Madhurima Chakraborty (University of California, Riverside), Peter Pirkelbauer (Lawrence Livermore National Laboratory, USA), and Qing Yi (Lawrence Livermore National Laboratory, USA)</i>	

Software quality

PyExamine: A Comprehensive, Un-Opinionated Smell Detection Tool for Python	763
<i>Karthik Shivashankar (University of Oslo, Norway) and Antonio Martini (University of Oslo, Norway)</i>	
Does Functional Package Management Enable Reproducible Builds at Scale? Yes.	775
<i>Julien Malka (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), Stefano Zacchiroli (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), and Théo Zimmermann (LTCI, Télécom Paris, Institut Polytechnique de Paris, France)</i>	
Refactoring for Dockerfile Quality: A Dive into Developer Practices and Automation Potential	788
<i>Emna Ksontini (University of Michigan - Flint, USA), Meriem Mastouri (University of Michigan - Flint, USA), Rania Khalsi (University of Michigan - Flint, USA), and Wael Kessentini (DePaul University, USA)</i>	
Smells-Sus: Sustainability Smells in IaC	801
<i>Seif Kosbar (Polytechnique Montréal, Canada) and Mohammad Hamdaqa (Polytechnique Montréal, Canada)</i>	
Evidence is All We Need: Do Self-Admitted Technical Debts Impact Method-Level Maintenance?..	813
<i>Shaiful Chowdhury (University of Manitoba, Canada), Hisham Kidwai (University of Manitoba, Canada), and Muhammad Asaduzzaman (University of Windsor, Canada)</i>	
DPy: Code Smells Detection Tool for Python	826
<i>Aryan Boloori (Dalhousie University, Canada) and Tushar Sharma (Dalhousie University, Canada)</i>	
CoMRAT: Commit Message Rationale Analysis Tool	831
<i>Mouna Dhaouadi (Université de Montréal, Canada), Bentley James Oakes (Polytechnique Montréal, Canada), and Michalis Famelis (Université de Montréal, Canada)</i>	
E2EGit: A Dataset of End-to-End Web Tests in Open Source Projects	836
<i>Sergio Di Meglio (University of Naples Federico II, Italy), Luigi Libero Lucio Starace (University of Naples Federico II, Italy), Valeria Pontillo (Vrije Universiteit Brussel, Belgium), Ruben Opdebeeck (Vrije Universiteit Brussel, Belgium), Coen De Roover (Vrije Universiteit Brussel, Belgium), and Sergio Di Martino (University of Naples Federico II, Italy)</i>	
TestMigrationsInPy: A Dataset of Test Migrations from Unittest to Pytest	841
<i>Altino Alves (UFMG, Brazil) and Andre Hora (UFMG, Brazil)</i>	
pyMethods2Test: A Dataset of Python Tests Mapped to Focal Methods	846
<i>Idriss Abdelmadjid (University of Nebraska-Lincoln, USA) and Robert Dyer (University of Nebraska-Lincoln, USA)</i>	
DataTD: A Dataset of Java Projects Including Test Doubles	851
<i>Mengzhen Li (University of Minnesota, USA) and Mattia Fazzini (University of Minnesota, USA)</i>	

JPerfEvo: A Tool for Tracking Method-Level Performance Changes in Java Projects	856
<i>Kaveh Shahedi (Polytechnique Montréal, Canada), Maxime Lamothe</i>	
<i>(Polytechnique Montréal, Canada), Foutse Khomh (Polytechnique</i>	
<i>Montréal, Canada), and Heng Li (Polytechnique Montréal, Canada)</i>	
Author Index	861