Bowen Xu

Research Scientist (Postdoc)

School of Computing and Information Systems

Singapore Management University

✆ https://shorturl.at/hzBC5

✉ maxxbw54@gmail.com

⌂ https://www.bowenxu.me

Research Interests and Past Works

- ♣ Machine Learning for Code
- ★ Software Text Analytics for Software Automation
- ■ Understanding and Assistance for Software Practice

[J]: Journal Papers   [C]: Conference Papers   [T]: Tool Demo Papers   [W]: Workshop Papers

Highlights of Publications (Awarded, Highly Cited, Invited)

ASE'22
CORE-A*

[C16] ♣ Jieke Shi, Zhou Yang, **Bowen Xu**, Hong Jin Kang, David Lo. "Compressing Pre-trained Models of Code into 3 MB". *The 37th IEEE/ACM International Conference on Automated Software Engineering. (ASE 2022).*

♟ Nominated for ACM SIGSOFT Distinguished Paper Award

ICPC'22
CORE-A

[C15] ★ Junda He, **Bowen Xu**, Zhou Yang, DongGyun Han, Chengran Yang and David Lo. "PTM4Tag: Sharpening Tag Recommendation of Stack Overflow with Pre-trained Models". *30th ACM/IEEE International Conference on Program Comprehension (ICPC 2022).*

👍 High peer review scores, Invited to Empirical Software Engineering Journal

ACSAC'22
CORE-A

[C14] Chen Gong, Zhou Yang, Yunpeng Bai, Jieke Shi, Arunesh Sinha, **Bowen Xu**, David Lo, Xinwen Hou, Guoliang Fan. "Curiosity-Driven and Victim-Aware Adversarial Policies". *The Annual Computer Security Applications Conference.*

♟ Won Honorable Mention Award

TSE'21
CORE-A*

[J4] ★ **Bowen Xu**, Hoang Van Duc Thong, Abhishek Sharma, Xin Xia, and David Lo. "Post2Vec: Learning Distributed Representations of Stack Overflow Posts". *IEEE Transactions on Software Engineering (TSE).*

📣 Presented during ICSE 2022 as part of the Journal First Paper Track

EMSE'20
CORE-A

[J3] ■ **Bowen Xu**, Le An, Ferdian Thung, Foutse Khomh, David Lo. "Why Reinventing the Wheels? An Empirical Study on Library Reuse and Re-implementation". *Empirical Software Engineering (EMSE).*

📣 Presented during ICSE 2020 as part of the Journal First Paper Track

ICSME'20
CORE-A

[C13] ★ Ting Zhang, **Bowen Xu**, Ferdian Thung, Stefanus Agus Haryono, David Lo, Lingxiao Jiang. "Sentiment Analysis for Software Engineering: How Far Can Pre-trained Transformer Models Go?". *36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020).*

🔥 Most cited paper (54 citations) among all the accepted papers at the time of writing

ESEM'18
CORE-A

[C12] ♣ **Bowen Xu**, Amirreza Shirani, David Lo and Mohammad Amin Alipour. "Prediction of Relatedness in Stack Overflow: Deep Learning vs. SVM - A Reproducibility Study". *12th International Symposium on Empirical Software Engineering and Measurement (ESEM 2018).*

♟ Highly Commended Full Paper Award

ASE'17 **CORE-A**[*]	[C11] ★ **Bowen Xu**, Zhenchang Xing, Xin Xia, David Lo. "AnswerBot - Automated Generation of Answer Summary to Developers' Technical Questions". *32nd IEEE/ACM International Conference on Automated Software Engineering (ASE 2017).* 🔥 *100 citations at the time of writing*
ASE'16 **CORE-A**[*]	[C10] ★ **Bowen Xu**, Deheng Ye, Zhenchang Xing, Xin Xia, Guibin Chen, Shanping Li. "Predicting Semantically Linkable Knowledge in Developer Online Forums via Convolutional Neural Network". *31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016).* 🔥 *156 citations at the time of writing*
MSR'16 **CORE-A**	[C9] ★ **Bowen Xu**, Zhenchang Xing, Xin Xia, David Lo, Qingye Wang, Shanping Li. "Domain-Specific Cross-Language Relevant Question Retrieval". *13th International Conference on Mining Software Repositories (MSR 2016).* 👍 *High peer review scores, Invited to Empirical Software Engineering Journal*

Other Publications

TOSEM'22 **CORE-A**[*]	[J2] ★ Ting Zhang, DongGyun Han, Venkatesh Vinayakarao, Ivana Irsan, **Bowen Xu**, Ferdian Thung, David Lo, Lingxiao Jiang. "Duplicate Bug Report Detection: How Far Are We?". *ACM Transactions on Software Engineering and Methodology. (TOSEM 2022).*
ASE'22 **CORE-A**[*]	[C8] ★ Chengran Yang, **Bowen Xu**, Ferdian Thung, Yucen Shi, Ting Zhang, Zhou Yang, Xin Zhou, Jieke Shi, Junda He, DongGyun Han, David Lo. "Answer Summarization for Technical Queries: Benchmark and New Approach". *The 37th IEEE/ACM International Conference on Automated Software Engineering. (ASE 2022).*
ESEC/FSE'22 **CORE-A**[*]	[C7] ♣ Yucen Shi, Ying Yin, Zhengkui Wang, David Lo, Tao Zhang, Xin Xia, Yuhai Zhao, **Bowen Xu**. "How to Better Utilize Code Graphs in Semantic Code Search?". *The 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2022).*
SANER'22 **CORE-A**	[C6] ♣ Jieke Shi, Zhou Yang, Junda He, **Bowen Xu** and David Lo. "Can Identifier Splitting Improve Open-Vocabulary Language Model of Code?". *29th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2022).* Accepted as a ERA track Paper.
SANER'22 **CORE-A**	[C5] ★ Chengran Yang, **Bowen Xu**, Junaed Younus Khan, Gias Uddin, Donggyun Han, Zhou Yang and David Lo. "Aspect-Based API Review Classification: How Far Can Pre-Trained Transformer Model Go?". *29th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2022).*
ESEC/FSE'19 **CORE-A**[*]	[T2] ★ Liang Cai, Haoye Wang, **Bowen Xu**, Qiao Huang, Xin Xia, David Lo, Zhenchang Xing. "AnswerBot: An Answer Summary Generation Tool Based on Stack Overflow". *The 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2019).*
AAAI'19 **RCQA**	[W1] ★ Amirreza Shirani, **Bowen Xu**, David Lo, Thamar Solorio and Mohammad Amin Alipour. "Question Relatedness on Stack Overflow: The Task, Dataset, and Corpus-inspired Model". *AAAI 2019 Reasoning and Complex QA Workshop.*

PAKDD'19 **CORE-A**	[C4] ■ Heng-Yi Li, Shu-Ting Shi, Ferdian Thung, Xuan Huo, **Bowen Xu**, Ming Li, and David Lo. "DeepReview: Automatic Code Review using Deep Multi-Instance Learning". *The 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2019).*
EMSE'17 **CORE-A**	[J1] ★ **Bowen Xu**, Zhenchang Xing, Xin Xia, David Lo, Shanping Li. "Domain-Specific Cross-Language Relevant Question Retrieval". *Empirical Software Engineering (EMSE).*
ESEC/FSE'17 **CORE-A**[*]	[T1] ★ **Bowen Xu**, Zhenchang Xing, Xin Xia, David Lo. "XSearch: A Domain Specific Cross-Language Relevant Question Retrieval Tool". *The 25th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2017).*
ASE'16 **CORE-A**[*]	[C3] ★ Guibin Chen, Chunyang Chen, Zhenchang Xing, **Bowen Xu**. "Learning Dual-Language Vector Space for Domain-specific Cross-Lingual Question Retrieval". *31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016).*
ICECCS'15 **CORE-B**	[C2] ■ Yun Zhang, David Lo, Xin Xia, **Bowen Xu**, Jianling Sun, Shanping Li. "Combining Software Metrics and Text Features for Vulnerable File Prediction". *20th International Conference on Engineering of Complex Computer Systems (ICECCS 2015).*
APSEC'15 **CORE-B**	[C1] ■ **Bowen Xu**, David Lo, Xin Xia, Ashish Sureka, Shanping Li. "EFSPredictor: Predicting Configuration Bugs With Ensemble Feature Selection". *22th Asia-Pacific Software Engineering Conference (APSEC 2015).*

Teaching

2023	Guest Instructor, Dalhousie University, Code Embedding (**Postgraduate Course**)
2022	Guest Instructor, Singapore Management University, IS706 - Software Mining and Analysis (**Postgraduate Course**)
2020	Teaching Assistant, Singapore Management University, IS706 - Software Mining and Analysis (**Postgraduate Course**) ⚇ **Students' Feedback**: https://tinyurl.com/1t9xklsh
2019	⚇ **Certification for passing Graduate Instructor Foundations in Teaching (GIFT) programme** British Council & Singapore Management University. ▶ Includes lesson planning, content development, instructional strategies, peer observation, and portfolio preparation, etc.
2019	Teaching Observation, Singapore Management University, IS113 - Web Application Development (**Undergraduate Course**)

Education

2017-2021	**Ph.D Degree** in School of Computing and Information Systems, Singapore Management University, Singapore Supervised by Prof. David Lo
2016	**Visiting Student** in Software Engineering Lab, Nanyang Technological University, Singapore Supervised by Prof. Zhenchang Xing

2014-2017	**Master Degree** in School of Software Technology, Zhejiang University, China
	Supervised by Prof. Shanping Li and Prof. Xin Xia
2010-2014	**Bachelor Degree** in College of Computer Science and Technology, Hunan City University, China

Grant Proposal Experience

2022-2027	**NRF Investigatorship** from National Research Foundation (NRF), Singapore
USD$2.4M	**Title**: "TrustedSEERs: Trusted Intelligent Work Bots for Engineering Better Software Faster"
Awarded	**Role**: Coordinator and key contributor for writing 1 of 3 core working packages
	Highlight: Only 3 awards given to CS projects in Singapore before
2022-2025	**MOE Tier-2** from Ministry of Education (MOE), Singapore
USD$770k	**Title**: "CCGuru: Learning Representation of Code Changes as Compressed Models to Automate
Under re-	Software Engineering Tasks"
view	**Role**: Coordinator and key contributor for writing 1 of 3 core working packages
2020-2023	**Industry Alignment Fund** from National Research Foundation (NRF), Singapore
USD$440K	**Title**: "Making Big Code Active: From Billions of Code Tokens to Automation"
Ongoing	**Role**: KPI contributor and responsible for managing manpower

Professional Services

Journal Referee for TR'22, TOSEM'22 and EMSE'22, TOIS'20 and JSS'20.

Co-Organizer for the workshop MaLTeSQuE'22 (collocated with ESEC/FSE'22).

Program Committee for ICSE'24, MSR'22.

Research Mentoring

Ting Zhang, PhD on the work ICSME'20, TOSEM'22

Chengran Yang, PhD on the work TSE'21, SANER'22, ASE'22

Jieke Shi, PhD on the work SANER'22, ASE'22

Junda He, PhD on the work ICPC'22

References

Dr. David Lo	Dr. Zhenchang Xing	Dr. Foutse Khomh
Full Professor	Senior Principal Research Scientist	Full Professor
Singapore Management University	CSIRO's Data61 & Australian National University	Polytechnique Montréal
davidlo@smu.edu.sg	zhenchang.xing@data61.csiro.au	foutse.khomh@polymtl.ca