

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

- Program Comprehension ✓ (Pre-trained) Language Model of Code
- ★ Software Text Analytics • Vulnerability/Defect Prediction

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

- ASE'22
CORE-A* [C15] ✓ 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**
- ASE'22
CORE-A* [C14] ★ 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* [C13] ✓ 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)*.
- ICPC'22
CORE-A [C12] ★ 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)*. **Invited to Empirical Software Engineering Journal**
- SANER'22
CORE-A [C11] ✓ 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 [C10] ★ 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)*.
- TSE'21
CORE-A* [J3] ★ **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 [J2] ■ **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	[C9] ★ 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?". <i>36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020)</i> .
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". <i>The 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2019)</i> .
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". <i>AAAI 2019 Reasoning and Complex QA Workshop</i> .
PAKDD'19 CORE-A	[C8] ■ 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". <i>The 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2019)</i> .
ESEM'18 CORE-A	[C7] ✓ Bowen Xu , Amirreza Shirani, David Lo and Mohammad Amin Alipour. "Prediction of Relatedness in Stack Overflow: Deep Learning vs. SVM - A Reproducibility Study". <i>12th International Symposium on Empirical Software Engineering and Measurement (ESEM 2018)</i> . Highly Commended Full Paper Award
ASE'17 CORE-A*	[C6] ★ Bowen Xu , Zhenchang Xing, Xin Xia, David Lo. "AnswerBot - Automated Generation of Answer Summary to Developers' Technical Questions". <i>32nd IEEE/ACM International Conference on Automated Software Engineering (ASE 2017)</i> .
EMSE'17 CORE-A	[J1] ★ Bowen Xu , Zhenchang Xing, Xin Xia, David Lo, Shanping Li. "Domain-Specific Cross-Language Relevant Question Retrieval". <i>Empirical Software Engineering (EMSE)</i> .
ESEC/FSE'17 CORE-A*	[T1] ★ Bowen Xu , Zhenchang Xing, Xin Xia, David Lo. "XSearch: A Domain Specific Cross-Language Relevant Question Retrieval Tool". <i>The 25th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2017)</i> .
ASE'16 CORE-A*	[C5] ★ Bowen Xu , Deheng Ye, Zhenchang Xing, Xin Xia, Guibin Chen, Shanping Li. "Predicting Semantically Linkable Knowledge in Developer Online Forums via Convolutional Neural Network". <i>31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)</i> .
ASE'16 CORE-A*	[C4] ★ Guibin Chen, Chunyang Chen, Zhenchang Xing, Bowen Xu . "Learning Dual-Language Vector Space for Domain-specific Cross-Lingual Question Retrieval". <i>31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)</i> .
MSR'16 CORE-A	[C3] ★ Bowen Xu , Zhenchang Xing, Xin Xia, David Lo, Qingye Wang, Shanping Li. "Domain-Specific Cross-Language Relevant Question Retrieval". <i>13th International Conference on Mining Software Repositories (MSR 2016)</i> . Invited to Empirical Software Engineering Journal
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". <i>20th International Conference on Engineering of Complex Computer Systems (ICECCS 2015)</i> .
APSEC'15 CORE-B	[C1] • Bowen Xu , David Lo, Xin Xia, Ashish Sureka, Shanping Li. "EFSPredictor: Predicting Configuration Bugs With Ensemble Feature Selection". <i>22th Asia-Pacific Software Engineering Conference (APSEC 2015)</i> .

Teaching

- 2022 Guest Instructor, Singapore Management University, IS706 - Software Mining and Analysis
- 2020 Teaching Assistant, Singapore Management University, IS706 - Software Mining and Analysis
Students' Feedback: <https://tinyurl.com/1t9xklsh>
- 2019 Graduate Instructor Certification, Singapore Management University, Singapore.

Education

- 2017-2021 **Ph.D Degree** in School of Computing and Information Systems, Singapore Management University, Singapore
Supervised by Prof. David Lo
- 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

Working Experience

- August, 2021 - Present **Research Scientist (Post-Doc)** in Software Analytics Research Group, Singapore Management University, Singapore
Work with Prof. David Lo
- Feb - July, 2016 **Research Assistant** in Software Engineering Lab, Nanyang Technological University, Singapore
Work with Prof. Zhenchang Xing

Research Mentoring

- Ting Zhang, PhD on the work ICSME'20
- 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

Professional Services

- Journal Referee** for TOSEM'22 and EMSE'22, TOIS'20 and JSS'20.
- Co-Organizer** for the workshop MaLTaSQuE'22 (collocated with ESEC/FSE'22).
- Program Committee** for MSR'22.

Certifications, honors & awards

- 2020 Dean's List (top3%), Singapore Management University, Singapore.
- 2020 Presidential Doctoral Fellowship (top3%), Singapore Management University, Singapore.
- 2017 Excellent Graduate of Zhejiang Province, Ministry of Education of the People's Republic of Zhejiang Province (top3%), China.