Nghi D. Q. **Bui**

Ph.D. in Computer Science

(+65) 85080361 | dqnbui.2016@phdis.smu.edu.sg | bdqnghi.github.io | bdqnghi | quocnghi | @QuocNghi91

Summary

My research interests lie at the intersection of Software Engineering, Natural Language Processing, and Machine Learning. With experience across industry and academia, I bring practical experience and theoretical expertise to this domain. The focus of my research is to build intelligent software engineering tools that enhance the development process of software developers.

My skills and achievements can be summarized as below:

- Have industry experience in software engineering (5 years) and data science (3 years)
- Familiar with software engineering development process and software engineering tools
- Familiar with back-end system design
- Good knowledge in program analysis techniques, such as static analysis, dynamic analysis, control-flow, and data-flow analysis
- Familiar with machine learning techniques
- Hands on experience with neural networks implementation and design
- Be able to design experiments for data science projects
- Hands on experience to manage an engineering and data science team
- Published papers in top-tier conferences across different domains in Computer Science , such as Software Engineering (ICSE, ESEC/FSE, ASE), Artificial Intelligence (NeurIPS, AAAI).

Education

Singapore Management University — *Singapore*

Ph.D. in Computer Science — *August 2016 - August 2020*

Ho Chi Minh University of Science — *Ho Chi Minh, Viet Nam*

Bachelor in Computer Science — *August 2009 - August 2013*

- Graduated from Advanced Program in Computer Science, which is an honor program

Skills

Distributed Computing & Streaming	Kafka, RabbitMQ, Spark, Hadoop
Machine Learning & Analytics	Tensorflow, Pytorch, Keras, Pandas, Scikit-learn, SciPy, Numpy, Pyplot
Programming Languages	Python, Java, Scala, Javascript, Matlab, C/C++
Source Code Version Control	Git, SVN
Web Development	Spring Framework, Play Framework, NodeJS, Flask
Database	MySQL, PostgreSQL, MongoDB, Redis, ElasticSearch
DevOps	Docker, Ansible

Work Experience

Huawei Research Center, Ireland — *Dublin, Ireland*

Research Scientist — *September 2020 - Now*

- Conduct multi-disciplinary cutting edge research across machine learning, software engineering and programming languages.
- Lead AI4Code, a project with the aims to combine deep learning and compiler theory to translate C++ into Rust.

Noggin Asia — *Singapore*

Head of Data Science — *March 2017 - Now*

- Supervising and manage the tasks for junior member of the team.
- Designing ML system architecture to process and collect a large volume of user's personal data, such as geolocation, website activities, etc.
- Using statistical modeling to model user's behavior.
- Developing recommendation system for item recommendation on client's website.

University of Bristol - The Open University — *United Kingdom*

Visiting Researcher — *October 2019 - December 2020 (Expected)*

- Collaborate with the researchers in the two universities to work on the topic: "Interpretability Method for Program Representation Learning". The goal for this research is to propose techniques that can interpret the output of neural networks that are being applied to solve a wide range of problems in software systems

Munchee *United States*

CTO *January 2017 - March 2018*

- Designing software architecture and monitoring all of the development processes for back-end, front-end and mobile applications
- Researching on Blockchain technology to serve for the business purpose
- Running the ICO

701Search *Ho Chi Minh City, Viet Nam*

Senior Software Engineer *March 2015 - July 2016*

- Worked with the team to redesign and migrate the existed monolith architecture into the micro-service architecture
- Developed the push notification service for the mobile app
- Used ElasticSearch as the information retrieval tool for geo-location queries
- Researched and applied Kafka as an event-sourced system to communicate between internal micro-services

Atlassian *Ho Chi Minh City, Viet Nam*

Software Engineer *January 2014 - March 2015*

- Worked as a software engineer in the JIRA team to develop useful libraries and features for internal usages, such as logging management dashboards to measure the quality of services
- Ensured the quality of source code by writing test cases and doing code review

Publications

[Under Review] **Efficient Framework for Learning Code Representations through Semantic-Preserving Program Transofmation**, by Nghi D. Q. BUI, Yijun YU, Lingxiao JIANG.

[AAAI'21] **TreeCaps:Tree-based Capsule Networks for Source Code Processing**, by Nghi D. Q. BUI, Yijun YU, Lingxiao JIANG, Full Paper, Main Track in Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI), Vancouver, Canada

[NeurIPS'19] **TreeCaps:Tree-Structured Capsule Networks for Program Source Code Processing**, by Vinoj Jayasundara, Nghi D. Q. Bui, Lingxiao Jiang, David Lo , in NeurIPS Workshop on ML for Systems, 2019, Vancouver, Canada.

[ASE'19] **AutoFocus: Interpreting Attention-based Neural Networks by Code Perturbation**, by Nghi D. Q. BUI, Yijun YU, Lingxiao JIANG, in Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019), Research Track, New Ideas Papers, San Diego, California, United States, 2019.

[ESEC/FSE'19] **SAR: Learning Cross-Language API Mappings with Little Knowledge** , by Nghi D. Q. Bui, Yijun Yu, Lingxiao Jiang , accepted at the IEEE/ACM 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), Research Track, Tallinn, Estonia, 2019.

[ICSE'19] **Towards Zero Knowledge Learning for Cross-Language API Mappings** , by Nghi D. Q. BUI, in Proceedings of the IEEE/ACM 41th International Conference on Software Engineering (ICSE): ACM Student Research Competition Track (SRC), Montreal, Canada, 2019 - (**Bronze Medal**).

[SANER'19] **Bilateral Dependency Neural Networks for Cross-Language Algorithm Classification**, by Nghi D. Q. Bui, Yijun Yu, Lingxiao Jiang, in the 26th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Research Track, Zhejiang University in Hangzhou, February 24-27, 2019.

[ICSE'18] **Hierarchical Learning of Cross-Language Mappings through Distributed Vector Representations for Code** , by Nghi D. Q. Bui, Lingxiao JIANG, in Proceedings of the IEEE/ACM 40th International Conference on Software Engineering (ICSE): New Ideas and Emerging Technologies Results Track (NIER), Gothenburg, Sweden, 2018 - (**ACM SIGSOFT Distinguished Paper Award**).

[NL4SE-AAAI'17] **Cross-Language Learning for Program Classification Using Bilateral Tree-Based Convolutional Neural Networks** , by Nghi D. Q. Bui, Lingxiao JIANG, and Yijun YU . In the proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI) Workshop on NLP for Software Engineering, New Orleans, Lousiana, USA, 2018.

Honors & Awards

2019	**SMU Dean's List,** Ph.D. Program	*Singapore*
2019	**SMU Presidential Doctoral Fellowship,** Ph.D. Program	*Singapore*
2019	**ACM SIGSOFT Travel Grant Award,** ESEC/FSE 2019	*Tallinn, Estonia*
2019	**Bronze Medal,** ACM Student Research Competition, ICSE 2019	*Montreal, Canada*
2019	**ACM SIGSOFT Travel Grant Award,** ICSE 2019	*Montreal, Canada*
2018	**ACM SIGSOFT Distinguished Paper Award,** ICSE 2018	*Gothenburg, Sweden*
2016	**Ph.D. Scholarship from the Ministry of Education,** Ph.D Program	*Singapore*

Teaching Experience

IS706: Software Mining and Analysis: Teaching Assistant

IS212: Software Project Management: Teaching Assistant