

MINH PHAM

Los Angeles, California

☎ 385-238-5583 ✉ minhpham1811@gmail.com Website: <https://minhptx.github.io/>

RESEARCH INTERESTS

- Knowledge Graph Construction
- Automatic Source Modeling
- Automatic Data Cleaning
- Applied Deep Learning

EDUCATION

University of Southern California

Sep 2015 – Now

Ph.D. Candidate in Computer Science

- Relevant Courses: Machine Learning, Building Knowledge Graphs, Advanced NLP, Representation Learning
- GPA: 3.88/4.0
- Advisors: Craig A. Knoblock and Muhao Chen

Ho Chi Minh City University of Technology

Sep 2009 – Jan 2014

Bachelor of Engineering in Computer Science

- Thesis Topic: Wikipedia-based Entity Disambiguation
- GPA: 8.5/10
- Advisors: Tru Hoang Cao

SELECTED PUBLICATIONS

M. Pham, C. Knoblock, M. Chen, B. Vu, and J. Pujara. *SPADE: A Semi-supervised Probabilistic Approach for Detecting Errors in Tables*. In 30th International Joint Conference on Artificial Intelligence (IJCAI 2021).

M. Pham, C. Knoblock and J. Pujara. *Learning Data Transformation with Minimal User Effort*. In 2019 IEEE International Conference on Big Data (IEEE BigData 2019).

M. Pham, S. Alse, C. Knoblock, and P. Szekely. *Semantic Labeling: A Domain-Independent Approach*. In 15th International Semantic Web Conference (ISWC 2016).

M. Pham, T. H. Cao and H. M. Huynh. *Candidate Searching and Key Coreference Resolution for Wikification*. In 10th International Conference on Ubiquitous Information Management and Communication 2016.

RESEARCH EXPERIENCE

Robust and Proactive Error Detection and Correction

Sep 2015 – Now

Research Assistant, Center on Knowledge Graphs, USC/ISI

- Develop an table verification approach that exploits open-domain question generation and answering to verify and correct semantic facts in tabular data
- Developed a semi-supervised error detection approach that combines Probabilistic Soft Logic and deep neural networks to detect syntactic errors in tables (IJCAI 2021)

MINT Project – <http://mint-project.info>

Sep 2015 – Now

Research Assistant, Center on Knowledge Graphs, USC/ISI

- Designed and developed a modular transformation pipeline to unify domains-specific data from different formats for scientific modeling
- Developed a novel unsupervised approach to transform string values between different formats (IEEE BigData 2019)

PRINCESS Project – <https://cra.com/projects/princess>

Sep 2015 – Now

Research Assistant, Center on Knowledge Graphs, USC/ISI

- Developed a machine learning approach for domain-independent semantic labeling, which can effectively analyze and map attributes/columns in tables to their correct properties in domain ontologies (ISWC 2016)

Nuance Communications Inc.

Jun 2019 - Aug 2019

Research Intern, Artificial Intelligence and Language Lab

- Developed an unsupervised entity resolution approach using Probabilistic Soft Logic and automatic data profiling

John von Neumann Institute, Vietnam National University

Sep 2014 – May 2015

Research Assistant

- Improved an existing learning-based entity linking system by candidate searching and rule-based coreference resolution.

PROFESSIONAL EXPERIENCE

DataFirst JSC Co.

Feb 2015 - Aug 2015

Research Programmer

- Extracted real estates' information from millions of Vietnamese online listings with high accuracy for market analysis.

East Agile

Jun 2014 – Aug 2014

Software Engineer Intern

- Developed and maintained an in-house video sharing platform using Ruby on Rails, CoffeeScript, HTML5 and CSS.

TEACHING EXPERIENCE

University of Southern California

2016 - 2021

Teaching Assistant, DSCI 558: Building Knowledge Graphs

Los Angeles, California

- Designed and evaluated course examinations, written assignments, and weekly quizzes
- Presented several sessions of lectures & research seminars to the class

AWARDS AND SCHOLARSHIPS

Best Paper Award, ISI Graduate Student Symposium, University of Southern California

2019

Title: Learning Data Transformations with Minimal User Effort

Vietnam Education Foundation (VEF) Fellowship for Ph.D. study in US

2015

\$54,000 for 35 selected Fellows in the whole country

TECHNICAL SKILLS

- **Machine Learning:** PyTorch, Tensorflow, Keras, Scikit-learn, Snorkel
- **Languages:** Python, Java, C++, SQL
- **Semantic Web:** RDF, Turtle, SPARQL
- **High Performance Computing:** Dask, Spark
- **Databases:** ElasticSearch, MongoDB

LEADERSHIP / EXTRACURRICULAR

Vietnam Education Foundation Fellows and Scholars Association (VEFFA)

2016 - 2017

Board of Executives

- Organized VEFFA annual conferences and events.
- Organized mock interviews for more than 60 Vietnam Education Foundation (VEF) Scholarship applicants.
- Led a group of 20 mentors to support VEF Scholarship applicants in preparing their applications.

PAKDD 2015 Conference & ACML 2014 Conference

2014 - 2015

Website Administrator and Volunteer

- Designed and managed conference website.
- Monitored presentation sessions in the conference.