Binh Vu

https://binh-vu.github.io
https://www.linkedin.com/in/binh-v-3828a16a/

Updated: November 7, 2019
binhvu@isi.edu
+1-213-269-9961

RESEARCH INTERESTS

Machine Learning, Data Integration, Knowledge Graph Construction

EDUCATION

- **Ph.D. Student in Computer Science** Aug 2016 - Present
  University of Southern California, Computer Science Department, Los Angeles, CA
  Advisor: Professor Craig Knoblock

- **Bachelor of Engineering in Computer Science** Aug 2010 - Jan 2015
  HCMC University of Technology, Ho Chi Minh, Vietnam
  Thesis: *Wikipedia-based Entity Disambiguation using Deep Learning.*
  Advisor: Professor Tru Cao. (Graduated in top 1%)

RESEARCH EXPERIENCE

- **University of Southern California, ISI**, Marina del Rey, CA Aug 2016 - Present
  Graduate Research Assistant, *Center on Knowledge Graph*

  - **Learning Semantic Models of Data Sources** Oct 2017 - Present
    Using probabilistic graphical models and deep learning to understand and automatically build semantic models of data sources. The semantic models can be used to publish data to knowledge graphs

  - **Creating a Data Catalog to Support Scientific Modeling** 2018 - Present
    Developing techniques for querying heterogeneous datasets and semi-automated data transformation.

  - **From sketches to HTML/CSS code** Fall 2018
    Applied deep neural network to synthesize HTML/CSS programs from mock-up UIs

  - **Identifying potential company hardware/software vulnerabilities** Oct 2016 - Sep 2017
    Auto-crawling online sources to retrieve expertise of employees of a company, then predicting the software and hardware used in the company. A list of vulnerabilities is obtained by linking the software and hardware to the CVE database.

- **HCMC University of Technology**, Ho Chi Minh, Vietnam Jun 2014 - Dec 2015
  Undergraduate Research Assistant, *Computer Science Department*

  - **Wikipedia-based Entity Disambiguation using Deep Learning**
    Using autoencoders to extract latent features of entities in Wikipedia articles for the entity linking problem

WORKING EXPERIENCE

- **Rakuten Inc.**, Tokyo, Japan Jun 2015 - May 2016
  Research Engineer, *Big Data Department*

  - **Fraud Detection in ID Hijacking and Payment**
    Developing a near real-time distributed streaming system using Apache Storm and Cassandra to analyzes time-series data for fraud detection. The system is designed to run models that use related historical events up to the past 60 days to flag fraudulent transactions within seconds.
    Implementing a machine learning model for payment fraud detection.

TEACHING EXPERIENCE

- **University of Southern California**, Los Angeles, CA   Fall 2017, Spring 2018
  Graduate Teaching Assistant for graduate-level course INF 558, "Building Knowledge Graph"

- **HCMC University of Technology**, Ho Chi Minh, VN   2015
  Teaching Assistant for courses: Artificial Intelligence, Introduction to Programming

HORNORS AND AWARDS

NSF sponsored Student Travel Awards for ISWC	2017, 2019
ISI Distinguished Top-Off Fellowship	April 2016
Vietnam Education Foundation Fellowship to pursue Ph.D. degree in the U.S	2016
Outstanding Honor Student Award	2011 - 2014

PUBLICATIONS

- **Binh Vu**, Craig Knoblock, and Jay Pujara. 2019. *Learning Semantic Models of Data Sources Using Probabilistic Graphical Models.* In The World Wide Web Conference, pp. 1944-1953.

- **Binh Vu**, Jay Pujara, and Craig Knoblock. 2019. *D-REPR: A Language for Describing and Mapping Diversely-Structured Data Sources to RDF.* In The Tenth International Conference on Knowledge Capture (K-CAP).

- B. Shbita, **B. Vu**, D. Feldman, M. Pham, A. Rajendran, C. Knoblock, J. Pujara, and Y. Chiang. 2019. Creating a FAIR Data Catalog to Support Scientific Modeling, in Workshop on Advanced Knowledge Technologies for Science in a FAIR World (AKTS).

TECHNICAL SKILLS

- **Languages**: Python, Rust, C/C++, Java, HTML, CSS, Javascript (Full-stack Web Developer)

- **Technologies**: Pytorch, Tensorflow, Docker, ElasticSearch, Databases (MySQL, Postgres, Redis, Cassandra), Apache Storm, Apache Hadoop, Apache Spark, ReactJS.

- **Operating Systems**: Linux, MacOS, Windows

PROFESSIONAL ACTIVITIES

Leader of a freelancer software development group	2012 - 2015
Organizing Volunteer in the Pacific-Asia Conference on Knowledge Discovery and Data Mining	2015
Organizing Volunteer in the Asian Conference on Machine Learning	2014
Organizing Volunteer in the Asian Conference on Information Systems	2014

REFERENCES

- **Professor Craig Knoblock**
  Interim Michael Keston Executive Director of the Information Sciences Institute
  University of Southern California, Information Science Institute, Marina del Rey, CA
  Email: knoblock@isi.edu

- **Professor Tru Cao**
  Department of Computer Science and Engineering HCMC University of Technology, Vietnam National University, Ho Chi Minh, VN
  Email: tru.cao@jvn.edu.vn