

XIAOYI DUAN

◇ +1 408 470 0153 ◇ xiaoyid@andrew.cmu.edu ◇ Mountain View, CA

EDUCATION

PhD student/Master in Software Engineering

Department of Electrical and Computer Engineering

Bachelor/Master in Software Engineering

Software Engineering Institute

Sep. 2016 - May 2019 expected

Carnegie Mellon University, Silicon Valley

Aug. 2009 - June 2016

East China Normal University, Shanghai, China

SKILLS

Knowledge

REST service, Data analytics, Text mining, Database

Technical Strengths

Python, Java, Javascript, Ajax, SQL, Cypher

Tools

Git, Linux, AWS, Jupyter, Docker

PROJECTS

Text Analysis and Evaluation System

May 2018 - Now

NASA project, human-involved domain words classification and customized lex parser

Individual Project

- Designed incremental classification algorithms to classify 12,023 domain words with limited human evaluations, by utilizing word definition, WordNet, and ontology
- Implemented a customized lex parser based on Stanford NLTK by integrating linguistic rules
- Built a flask-based web application to support multi-user evaluation on machine classification results and submit sentences to run customized lex parser

Data Analytics Powered Knowledge Network System in Earth Science

July 2017 - Now

NSF/NASA project, a scientific knowledge network for earth science papers

Student Leader

- Developed flask APIs for external systems and data analytical tasks, including pipelines for data extraction and processing and data retrieval from mongodb or neo4j
- Implemented algorithms for advanced tasks, such as semantic entity identification from unstructured data, topic modeling and word disambiguation
- Designed the structure of scientific knowledge network and functionalities for visualization and query

Collaborative Scientific Workflow

Sep. 2016 - June 2017

NSF/NASA project, a real-time collaboration system to design and manage scientific workflow

Student Leader

- Designed data schema to store design-time workflow provenance
- Developed back-end node.js APIs to query the provenance in graph database

PUBLICATIONS

(Best Student Paper Award) Xiaoyi Duan, Jia Zhang, Rahul Ramachandran, et al. "A Neural Network-Powered Cognitive Method of Identifying Semantic Entities in Earth Science Papers." In IEEE International Conference on Cognitive Computing (ICCC), 2018. [slides](#)

Xiaoyi Duan, Jia Zhang, Qihao Bao, Rahul Ramachandran, et al. "Linking Design-Time and Run-Time: A Graph-Based Uniform Workflow Provenance Model." In 2017 IEEE International Conference on Web Services, pp. 97-105, 2017. [slides](#)

Qihao Bao, Jia Zhang, **Xiaoyi Duan**, Rahul Ramachandran, Tsengdar J. Lee, et al. "A Fine-Grained API Link Prediction Approach Supporting Mashup Recommendation." In IEEE International Conference on Web Services, pp. 220-228, 2017.

Xiaoyi Duan, Cheqing Jin, Xiaoling Wang, et al. "Real-time Personalized Taxi-Sharing." In International Conference on Database Systems for Advanced Applications, pp. 451-465, 2016.

Keqiang Wang, **Xiaoyi Duan**, Jiansong Ma, Chaofeng Sha, et al. "Linking Design-Time and Run-Time: A Graph-Based Uniform Workflow Provenance Model." In International Conference on Database Systems for Advanced Applications, pp. 381-395, 2016.

ACTIVITIES AND AWARDS

Student Volunteer, 2018/2017 IEEE Services Congress

July 2-7, 2018/June 25-30, 2017

Second Place Award, China Big Data Techniques Innovation and Entrepreneurship Contest

Fall 2014

Excellent Award, Microsoft Imagine Cup Contest

Spring 2014