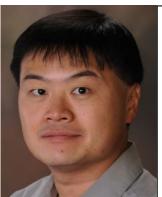# Department of Data Science

Data Science Seminar Series

A Natural Language Interface for Databases

Wei-Shinn Ku, Ph.D.

Professor
Auburn University

Date: Tuesday, February 15th, 2022
Time: 11:00 AM – 12:00 PM EST
Location: Zoom Virtual Room
Web Link: [Zoom Meeting Room Link](#)

Relational database management systems (RDBMSs) are powerful because they are able to optimize and execute queries against relational databases. However, when it comes to NLIDB (natural language interface for databases), the entire system is often custom-made for a particular database. Overcoming the complexity and expressiveness of natural languages so that a single NLI can support a variety of databases is an unsolved problem. In this research, we show that it is possible to separate data-specific components from latent semantic structures in expressing relational queries in a natural language. With the separation, transferring an NLI from one database to another becomes possible. We develop a neural network classifier to detect data-specific components and an adversarial mechanism to locate them in a natural language question. We then introduce a general-purpose transfer-learnable NLI that focuses on the latent semantic structure. We devise a deep sequence model that translates the latent semantic structure to an SQL query. In this talk, I will also cover my background and my vision for the Department of Data Science at NJIT.

Wei-Shinn Ku received his Ph.D. degree in computer science from the University of Southern California (USC) in 2007. He also obtained both the M.S. degree in computer science and the M.S. degree in electrical engineering from USC in 2003 and 2006, respectively. He is a professor with the Department of Computer Science and Software Engineering at Auburn University. He had been a program director at the National Science Foundation from 2019 to 2022. His research interests include data science, data management, mobile computing, and cybersecurity. He has published more than 140 research papers in refereed international journals and conference proceedings. He is a senior member of the IEEE and a member of the ACM.