



Ph.D. Dissertation Defense

Candidate:	Lun Li
Degree:	Doctor of Philosophy
School/Department.:	Interdisciplinary / Data Science
Date:	Tuesday, November 5, 2024
Time:	11:00 – 12:30 pm
Location:	Babbio 601
Title:	Workforce Analytics in the Era of Big Data
Chairperson:	Dr. Emily Liu, Information Systems, School of Business Dr. Theodoros Lappas, Department of Marketing & Communication, Athens University of Economics and Business
Committee Members:	Dr. Jingyi Sun, Information Systems, School of Business Dr. Jeffrey Nickerson, Information Systems, School of Business Dr. Justine Herve, Economics, School of Business

Abstract

This dissertation comprises three essays to develop innovative and systematic methods for workforce analysis. By utilizing large-scale online resume data and advanced deep learning technologies, we examine one of the most prevalent topics in workforce analysis – attrition and retention – at both the organizational and individual levels.

• **Organizational Level:** The first essay introduces a novel algorithm, *TalentRank*, to assess and track employers’ attractiveness. This algorithm offers a dynamic ranking system based on the changes in talent inflow and outflow, and it provides detailed rankings across various job titles. As a by-product, *TalentGrouper* is devised to address the significant variability encountered in online resume databases, making it a valuable tool for other workforce analysis.

• **Individual Level:** The main thrust of the second and third essays is to explain individuals’ job-hopping behavior via a theory-driven deep-learning paradigm. Under the guidance of category theory, we refine the definition of typicality to account for the temporal axis, which captures the evolving expertise of an individual and thus their ability to acquire a job that they are most experienced in. Although dynamic typicality plays a significant role in explaining occupational mobility, it does not reconcile the fact that individuals do span the boundaries for “untypical” opportunities. To this end, we introduce two forms of category interdependencies: **hierarchical category logic** and **social comparison logic**. The former one is taken account by establishing a bi-level title hierarchy, and the latter one is modeled via a novel spatial-temporal graph guided by a theory-driven meta-path. The hybrid representation of an employee is interpreted by a bi-resolution probabilistic structure that enables next occupation prediction. In summary, this work holds significant implications for employers to optimize their human capital strategies and provide a data-driven solution to retain talents more effectively.