



## Spring 2021 Colloquium

Department of Computer and Information Sciences

### *Title: Learning-Based Methods for Understanding and Linking Entities*

**Dr. Lucian Popa**

Principal Research Staff Member and Manager  
IBM Research - Almaden

**Thursday, April. 8th, 11 AM**

**Zoom Link: <https://temple.zoom.us/j/94687622618>**

**Abstract:** Entity resolution is a key form of knowledge discovery that allows to understand, connect, and group different representations of the same entity across various types of data. It can manifest itself in multiple forms, ranging from understanding the string representations of an entity name, to structured entity matching in databases, and to linking of entities mentioned in text to knowledge bases such as DBpedia or Wikidata. Along with other ubiquitous operations such as information extraction, data transformation and fusion, entity resolution is a crucial step for building high-value, domain-specific knowledge from raw data.



One of the key research challenges that we address, in this space, is the development of human-in-the-loop, learning based tools that can help domain experts build high-accuracy entity resolution algorithms that are scalable and explainable, from small amounts of labeled data. Towards this goal, I will describe various types of learning techniques that we have been developing. These include combinations of active learning with weak supervision and rules, as well as a neuro-symbolic approach for entity linking that combines the performance of neural learning with the benefits of explainable, symbolic representations, while reaching state-of-the-art performance on several benchmark datasets.

**Bio:** Lucian Popa is a Principal Research Staff Member and Manager at IBM Research - Almaden. He is known for his work on data exchange and schema mapping, for which he received two Test-of-Time Awards in ICDT 2013 and PODS 2014, and for work on human-in-the-loop tools and foundations for entity resolution, for which he received a Best Paper Award in ICDT 2015. He is also a winner of the 2020 Alonzo Church Award for Outstanding Contributions to Logic and Computation. Lucian has contributed to several IBM products in the area of information integration and entity resolution, and he is an ACM Distinguished Member.