In recent years, the availability of big data has resulted in a growing number of users who are interested in interpreting the trends and anomalies for large datasets. This presents an imminent requirement of sophisticated data analysis tools that can provide qualitative information based on query answers on such datasets. In this talk, I will describe my current research on developing a principled framework for explaining query answers inspired by the theory of causality and intervention from the area of Artificial Intelligence. I will present our solutions to core challenges in this task such as obtaining concise descriptions of explanations, handling inherent dependencies of database tuples, and achieving real-time efficiency in large explanation spaces. I will conclude the talk with several exciting future research directions spanning database theory and systems, algorithms, and user interactions with a graphical interface.

Bio:
Sudeepa Roy is an Assistant Professor in Computer Science at Duke University since Fall 2015. She works in the area of databases and data management, with a focus on foundational aspects of big data analysis, which includes causality and explanations for big data, data provenance, probabilistic databases, and applications of database techniques in other domains. Prior to Duke, she did a postdoc at the University of Washington, and obtained her Ph.D. from the University of Pennsylvania. She is a recipient of the NSF CAREER award and a Google PhD Fellowship.

Friday, February 3 @ 10 AM
University at Buffalo – North Campus – Davis 113A