

UCI Department of Computer Science

DISTINGUISHED LECTURE SERIES

RENÉE J. MILLER

Professor of Computer Science, University of Toronto

“Evaluating Data Integration and Data Curation Systems”

Friday - May 12, 2017 - 11 a.m.

Donald Bren Hall, Room 6011

*No cost to attend — Open to the public
Seating is on a first-come, first-served basis*

ABSTRACT

In this talk, I present some important innovations in data integration research over the last two decades. These include data exchange, which provides a foundation for reasoning about the correctness of transformed data, and the use of declarative mappings in integration, cleaning and more broadly data curation. I discuss how curation research has expanded from the enterprise to consider how to add value to open data, and how big data is changing the nature of curation solutions. While data integration and cleaning are very mature fields, rigorous empirical evaluations of systems are still relatively scarce. I discuss how developments in information integration over the last two decades enable evaluation. I also identify major roadblocks for empirical work including the lack of tools that aid in generating the inputs and gold standard outputs for integration or cleaning tasks in a controlled, effective, and repeatable manner. I give an overview of our efforts to develop such tools and highlight how our tools have been used for streamlining the empirical evaluation of a variety of systems.

BIO

Renée J. Miller is a Professor of Computer Science at the University of Toronto. She has been named fellow of the Royal Society of Canada (Canada’s National Academy), ACM Fellow, and Bell Canada Chair of Information Systems. She received the US Presidential Early Career Award for Scientists and Engineers (PECASE), the highest honor bestowed by the United States government on outstanding scientists and engineers beginning their careers. She received an NSF CAREER Award, the Premier’s Research Excellence Award, and an IBM Faculty Award. Her research is in the area of data integration and data curation. She and her co-authors received the ICDT Test-of-Time Award for their influential 2003 paper establishing the foundations of data exchange. She has served on the Board of Trustees of the VLDB Endowment and as President of the Endowment. Her research is funded by NSERC, NSF, IBM, SAP, and Bell Canada among others. She received her PhD in Computer Science from the University of Wisconsin, Madison and Bachelor’s degrees in Mathematics and in Cognitive Science from MIT.

More info: www.cs.toronto.edu/~miller

For further information, please contact
hbyrnes@ics.uci.edu or go to www.cs.uci.edu