

# Information Integration: a Vision for Integration Independence and Linking Open Data

Renee J. Miller

University of Toronto  
miller@cs.toronto.edu

To integrate information, data in different formats, from different, potentially overlapping sources, must be related and transformed to meet the users' needs. Ten years ago, Clio introduced a new paradigm for creating declarative schema mappings to describe the relationship between data in heterogeneous schemas. This enabled powerful tools for mapping discovery and integration code generation, greatly simplifying the integration process. In this talk, I take a look at where our field was a decade ago and where it is now in terms of support for data integration. I share a vision for raising the level of abstraction further, to better isolate applications from the details of how the integration is accomplished. Integration independence allows applications to be independent of how, when, and where information integration takes place, making materialization and the timing of transformations an optimization decision that is transparent to applications. I identify a number of research challenges that remain to be addressed in order to ultimately achieve this vision.