Query Evaluation on Schematic Correspondences in the Context of Dataspaces

By Lu Mao -- lmao@cs.man.ac.uk

In collaboration with: Khalid Belhajjame, Norman W. Paton and Alvaro A. A. Fernandes

Background:
- Information management systems seek to provide low (labour) cost information integration services over large-scale and heterogeneous information spaces, namely dataspaces.
- The integration is realised by query translation between heterogeneous data sources.
- Example applications: scientific research, enterprise information management and personal information management.

Motivations:
- Conventional data integration features resolve heterogeneities heavily rely on labour and tool investments to elicit semantic correspondences between data instances residing at pair of data sources.
- Automating such process can generate significant amount of incomplete and inconsistent description between heterogeneous data sources which will subsequently impair the quality of the information integration service.
- Another type of correspondence, Schematic correspondences, are known to be more informative on schema elements which describe stored data instances, yet it requires less cost to elicit than those semantic correspondences.

Aim:
- To develop automated mechanisms to translate queries between heterogeneous data sources given information on their schematic correspondences.