**INTRODUCTION**

This paper addresses the problem of query-aware data cleaning, wherein the needs of the query dictates which parts of the data should be cleaned. Query-aware cleaning is emerging as a new paradigm for data cleaning to support today's increasing demand for (near) real-time analytical applications. Modern enterprises have access to potentially limitless data sources, e.g., web data repositories, social media posts, clickstream data, etc. Analysts usually wish to integrate one or more such data sources (possibly with their own data) to perform joint analysis and decision making. As a result of merging data from different sources, a given real-world object may often have multiple representations, resulting in data quality challenges. In this paper, we focus on the Entity Resolution (ER) challenge [16], [19], [29].