

Systematic review of data quality research

ABSTRACT

Data quality drawn a major concern when dealing with data especially in the event that insightful outputs is needed. Research in data quality emerged in various topics and diversification in known knowledge and used approach is inevitable. In this paper, we apply systematic review study to explain the landscape of data quality and to identify available research gap by using categorization and mapping. Our search scope is limited to research articles from journals, conference proceedings and magazine published between 2010 until 2016. We defined three types of main categorization to map the selected research articles and to answer our research questions. These categorization focus on research topics, research type and contribution type. On average, fifty-four research articles related to data quality were published every year. This number shows the importance of data quality research in various research topics such as online users, database, web information, sensors and big data. This study also indicates that almost half of the selected articles proposed a novel solution or an essential extension of an existing data quality technique. Moreover, most of the selected research articles belongs to the model type in the contribution category. Our mapping also suggests that obvious contribution disparity happen between contribution in metric type and model type category.

Keyword: Data quality; Information quality; Systematic review