



Contribution ID: 155

Type: **Talk (invited)**

KEYNOTE 6: Big Data Analytics: Does Data Quality Really Matter?

Wednesday, 5 December 2018 09:45 (40 minutes)

Big Data analytics is the often complex process of examining large and varied data sets – or Big Data – to uncover information including hidden patterns, unknown correlations, market trends and customer preferences that can help organizations make informed business decisions, including new revenue opportunities, more effective marketing, better customer service, improved operational efficiency and competitive advantages over rivals. Numerous industrial and research databases have quality issues including outliers, noise, missing values, and so on. In fact, it is not uncommon to encounter databases that have up to a half of the entries missing, making it very difficult to mine them using data analysis methods that can work only with complete data. Currently, comprehensive analysis and research of quality standards and quality assessment methods for big data are lacking. First, this talk summarizes reviews of data quality research. We then analyze the data characteristics of the Big Data environment, and present quality challenges faced by Big Data. Finally, we construct a dynamic assessment process for data quality.

Presenter Biography

Primary author: TWALA, Bhekisipho (University of South Africa)

Presenter: TWALA, Bhekisipho (University of South Africa)

Session Classification: KEYNOTE 5 DIRISA: Prof Bhekisipho Twala, University of South Africa

Track Classification: SA NREN