CoE-MaSS weekly seminar series

THE DST-NRF CENTRE OF EXCELLENCE IN MATHEMATICAL AND STATISTICAL SCIENCES (CoE-MaSS) WOULD LIKE TO PRESENT A SEMINAR BY

Prof Bhekisipho Twala

Director: Institute for Intelligent Systems, UJ, South Africa

"Big Data Quality Issues"

Friday 22 September 2017
10h30-11h30

Broadcast live from:
Videoconferencing Facility, 1st Floor
Mathematical Sciences Building, Wits West Campus

How to connect to this seminar remotely:
You can connect remotely via Vidyo to this research seminar by clicking on this link:
http://wits-vc.tenet.ac.za/flex.html?roomdirect.html&key=y0SSOwFsvsidbzg4qFdWXvvQtyI
and downloading the Vidyo software before the seminar.
You must please join in the virtual venue (called *"CoE Seminar Room (Wits)"* on Vidyo) strictly between **10h00-10h15**. No latecomers will be added.

Important videoconferencing netiquette:
Once the seminar commences, please mute your own microphone so that there is no feedback from your side into the virtual room. During the Q&A slot you can then unmute your microphone if you have a question to ask the speaker.

Title:
Big Data Quality Issues

Presenter:
Bhekisipho Twala, Ph.D.
DIRECTOR Institute for Intelligent Systems
Professor in Artificial Intelligence & Statistical Science
University of Johannesburg
btwala@uj.ac.za
T +27 11 559-4404 | F +27 11 559-2344 |

Abstract:
Recent advances in technology and increasing demands for accountability have allowed organizations in industries to create, store, and process massive amounts of data. To be useful, data in registries or databases must be of good quality and statistical ideas are essential part of this. Thus, the importance of data quality (DQ) and statistical inference (SI) in the information age cannot be overestimated, yet most databases contain errors and inconsistencies, and the role of statistical inference in the analysis of Big Data has not been addressed. Errors in data cannot only cost a company millions of dollars, alienate customers, and make implementing new strategies difficult or impossible while wrong statistical inference and lead to wrong scientific conclusions. This talk gives overviews on the salient features of Big Data and how these features impact on paradigm change on statistical and computational methods as well as computing architectures. We also provide various new perspectives on Big Data Quality and Statistical Inference.