

Taming the Data Deluge to Unravel the Mysteries of the Universe

Melanie Johnston-Hollitt
School of Chemical & Physical Sciences, Victoria
University of Wellington & Peripety Scientific Ltd.
Wellington, New Zealand
Melanie@PeripetySci.com

ABSTRACT

Modern Astrophysics is one of the most data intensive research fields in the world and is driving many of the required innovations in the ‘big data’ space. Foremost in astronomy in terms of data generation is radio astronomy, and in the last decade an increase in global interest and investment in the field had led to a large number of new or upgraded facilities which are each currently generating petabytes of data per annum. The peak of this so-called ‘radio renaissance’ will be the Square Kilometre Array (SKA) – a global observatory designed to uncover the mysteries of the Universe. The SKA will create the highest resolution, fastest frame rate movie of the evolving Universe ever and in doing so will generate 160 terrabytes of data a day, or close to 5 zettabytes of data per annum.

Furthermore, due to the extreme faintness of extraterrestrial radio signals, the telescope elements for the SKA must be located in radio quiet parts of the world with very low population density. Thus the project aims to build the most data intensive scientific experiment ever, in some of the most remote places on Earth. Generating and serving scientific data products of this scale to a global community of researchers from remote locations is just the first of the ‘big data’ challenges the project faces. Coordination of a global network of tiered data resources will be required along with software tools to exploit the vast sea of results generated.

In fact, to fully realize the enormous scientific potential of this project, we will need not only better data distribution and coordination mechanisms, but also improved algorithms, artificial intelligence and ontologies to extract knowledge in an automated way at a scale not yet attempted in science. In this keynote I will present an overview of the SKA project, outline the ‘big data’ challenges the project faces and discuss some of the approaches we are taking to tame the astronomical data deluge we face.

Author Keywords

Square Kilometre Array; Radio Astronomy; Big Data; Design and Analysis of Algorithms; Pattern Matching; Machine Learning; Data Compression

© 2017 International World Wide Web Conference Committee (IW3C2), published under Creative Commons CC BY 4.0 License.

WWW 2017, April 3-7, 2017, Perth, Australia.

ACM 978-1-4503-4913-0/17/04.

DOI: <http://dx.doi.org/10.1145/3038912.3050769>



BIOGRAPHY

Melanie Johnston-Hollitt is an internationally prominent radio astronomer working in the space between astrophysics, computer science and big data. She is Director of Astronomy & Astrophysics at Victoria University of Wellington and CEO of Peripety Scientific Ltd., an astrophysics and data analytics research company based in Wellington, New Zealand. In addition she is currently Chair of the Board of the 60 million dollar Murchison Widefield Array (MWA) radio telescope and a founding member of the Board of Directors of the Square Kilometre Array (SKA) Organisation Ltd, which is tasked with building the world’s largest radio telescope. In her nearly 20 year career, Prof. Johnston-Hollitt has been involved in design, construction and operation of several major radio telescopes including the Low Frequency Array in the Netherlands, the MWA in Australia and the SKA which will be hosted in both Australia and South Africa. These instruments produce massive quantities of data, requiring new and disruptive technologies to allow value to be extracted from the data deluge. As a result Prof. Johnston-Hollitt’s recent interests span the intersection between radio astronomy, signal processing and big data analytics and she leads a multi-disciplinary team in Wellington, NZ investigating how best to meet the science challenges of these next generation instruments in the big data era.

