

**Title: Balancing privacy with access in big data analytics**

- Christine M O'Keefe,  
Senior Principal Research Scientist, CSIRO

Vast amounts of data are now being generated from census and surveys, scientific research, observational projects, instruments, and sensors of many kinds. The need to protect the privacy of individuals in the context of health, social and survey databases is widely-recognised, however confidentiality can also be an issue with business datasets of commercially sensitive information. The increasingly common practice of geo-coding datasets can substantially increase the confidentiality issues.

In this talk, I will outline current approaches to balancing use and analysis of data with confidentiality protection, for both personal and business data. These approaches were generally developed in the context of national statistical agencies and other data custodians releasing social and survey data, but are increasingly being adapted to other contexts. The challenge, in all cases, is to ensure that the statistical outputs sufficiently protect the confidentiality of the data while yielding correct statistical inferences.

I will indicate some of the issues that need to be considered in the increasingly important context of big data analytics.