

## Editorial

### Data Citation - Moving to New Norms

**H**ow many times have you heard platitudes about the value of data, about how it is the most important resource of the institution, about how people can be replaced but data is unique, about how data must be properly managed and curated? Resource investments suggest otherwise. Compare the costs of obtaining the data and the sums spent on curating and managing it. Somehow the two do not add up. But open access and transparency are demanding changes in the way we manage and resource data so let us consider what needs to be done to protect data and make it available.

Capturing the data is the first task. The pattern varies across the world, and even across institutions. Whilst some countries have insisted that receiving a research grant is tied to data deposit others still seem to find it impossible, especially with university scientists, to adopt such a measure. The problem is not simply a lack of commitment from management to build the right infrastructure or employ the right number of data managers, although that can be a problem. It goes deeper and is centred on the individual scientist's assessment of ownership, the importance and value of data and on the way management rewards effort. Many scientists, especially in large programmes, have a real interest in the quality and curation of data, but too many others still see the cleaning and organizing of data and the provision of metadata as a housekeeping chore too far. One important reason for this is the lack of credit they would get for what could be quite time consuming work. But all this is changing as the value of data citation is being recognized and, with it, the efforts by individuals to prepare the data for open access. Discussions on this began almost a decade ago but progress has been slow. Not least were the problems in reaching agreement on standards and formats, including providing each data set with a unique Digital Object Identifier (DOI).

Two recent initiatives in Germany show progress is being made. Earth System Science Data (<http://earth-system-science-data.net/>) is a new journal which aims to publish papers on original research datasets to draw attention to their extent and quality as a way of encouraging re-use. With papers encompassing everything from comparative reviews of datasets to methods for cleaning and normalizing data, and an open access review policy, this journal is a welcome new step. This will clearly allow those who develop methodologies, who plan experiments or collect data to share their ideas and developments in a new way. To complement this is the development of DataCite, (<http://www.datacite.org/>), a new international consortium of nine countries intent on allowing datasets to be more easily shared and repurposed, improving access and giving datasets DOIs to make them citable and a part of the record of any project. The International Polar Year saw data management as a key issue and is urging scientists to use the existing tools to cite data (<http://ipydis.org/data/citations.html>). SCAR continues to urge all Antarctic scientists to register their data in the Master Directory (<http://scadm.scar.org/adms.html>).

This journal will also be moving towards asking authors to begin using data citations. There are still many unknowns to deal with - how will citations be identified (agreed standard format), who will identify them (publishers or data centres), where should they be in the paper (text hyperlink or reference list) and what search terms should be used. But data preparation for curation should not be optional and data citation should be part of any publication - these are the new norms.

DAVID W.H. WALTON