Geophysical Research Abstracts
Vol. 14, EGU2012-5175-1, 2012
EGU General Assembly 2012

The Geoscience Data Journal: collaboration between data repositories and publishing houses in data publishing

F.L. Murphy (1), S.A. Callaghan (2), P.J. Hardaker (3), and R.J Allan ()

(1) Wiley-Blackwell, Earth Science Journals, Chichester, West Sussex, UK (fmurphy@wiley.com), (2) British Atmospheric Data Centre, STFC - Rutherford Appleton Laboratory, Didcot, Oxfordshire, UK (sarah.callaghan@stfc.co.uk), (3) Royal Meteorological Society, Reading, UK (chiefexec@rmets.org)

It is becoming increasingly important to the scientific and wider non-academic communities that the data that underpins key scientific results should be made available to allow for the testing and confirmation of those results. Historically, publishing data has been so difficult as to be prohibitive, and those cases where it has been possible, the raw data has had to have been converted to other formats; for example, instead of raw numbers being published in a (lengthy) table, it has been converted to a graph.

As scientists' ability to create and collect new data has been growing, so too has our ability to store it. A dataset can be stored on any digital medium that is convenient, but future-proofing the data so that it is readable and understandable in 20 years' time remains a time-consuming and difficult job. Yet, if the results drawn from that data are to stand up to scrutiny in the future, the data must be curated and archived properly.

Openly sharing data is often proposed as a method for ensuring that data underpinning the scientific record is kept. There are issues with this in that sharing data in an unstructured way often results in the provenance of the dataset (and often the dataset itself) being changed as it passes from one "owner" to another, thereby reducing any chances of using that data to test the reproducibility of results originally made from it. Also, the present mechanism for academic recognition revolves around the production and publication of peer-reviewed papers. The production of high-quality datasets takes time and effort, and is often insufficiently recognised as an activity worthy of prestige, even though the papers that result from that dataset may be considered of significant scientific importance. Simple sharing of data is unlikely to provide the data creators with the academic recognition they deserve.

It is for these reasons that a partnership has been developed between the British Atmospheric Data Centre, the Royal Meteorological Society and the academic publishers Wiley-Blackwell, in order to develop a mechanism for the formal publication of data in the (soon to be launched) Geoscience Data Journal. This journal builds on the work funded by JISC in the OJIMS (Overlay Journal Infrastructure for Meteorological Sciences) project, and parallels with work done by the NERC Science Information Strategy Data Citation and Publication project team, which brings all the NERC environmental data centres together.

The aim of the Geoscience Data Journal is to provide a platform where scientific data can be formally published, in a way that includes scientific peer-review. This will provide the dataset creator with full credit for their efforts, while also improving the scientific record, and allowing major datasets to be fully described, cited and discovered.