



RDM in Natural History

Professor Malcolm Kennedy is a researcher who believes the research process should be as open and transparent as possible, for the benefit of other researchers, society and the integrity of scientific research. Here he discusses the ways in which he endeavours to make his research open.

Research Area

Malcolm is Professor of Natural History and eclectic may be the most accurate description of his research interests. He notes on his webpage that the direction in which his research might move can't be anticipated because he doesn't know what might spark his interest next, but the following list gives an idea of where his interests have led him in the past:

- Evolution of viviparity in mammals
- Lactation biology – from giant pandas to seals and beyond
- Allergen proteins
- The molecular biology and biochemistry of frog nest foams
- Nesting by marine turtles
- The structure and function of unusual lipid binding proteins of free-living and parasitic nematodes
- The structure and function of intrinsically-active, cell compatible surfactant proteins of frogs and horses

Funding and Partners

Malcolm's work is currently unfunded, but in the past, he has had projects funded by the Wellcome Trust; The Carnegie Trust for the Universities of Scotland; The Percy Sladen Memorial Fund; The Horserace Betting Levy Board; the European Commission and the BBSRC.

Of these, Wellcome have done much in recent years to drive firstly the open access to research publications agenda, and more recently the open data \ open research agenda.

Malcolm does not have a research group, but collaborates widely with colleagues, both at Glasgow and worldwide.

Open Research

Malcolm initially moved into Open Research by ensuring that his publications are published with open access whenever possible. Over the last two years, he has started engaging in other open research practices:

On noticing that the supporting data for some articles was available via a link included in the paper, Malcolm discovered repositories for research data which facilitate the sharing of research data. He now makes a point of making his own data available when he publishes papers. He has used both Figshare and Dryad to make research data available.



Figshare is a generic research data repository that accepts data in any format from any field of research.



Dryad is a research data repository that accepts data predominantly from the life sciences. Data deposited in Dryad must be associated with a peer-reviewed article or other reputable source such as a dissertation or book. Dryad curates all datasets deposited to check for technical issues prior to the dataset being made public. There is an ingest charge

for using Dryad (which is occasionally waived or covered if the data is sponsored by the publisher of the associated article).

One of the common concerns researchers have about sharing their data is that they will have to spend a lot of time getting the data into a form which would be easily understandable by other researchers. Malcolm notes that he does spend quite a bit of time annotating his excel files with notes and comments to aid understanding, but that the time he spends on this is primarily for his own benefit – his analysis of large datasets can take a very long time and good annotations made in real-time help him recall his thought processes and ideas. These annotations then serve the double purpose of helping others to understand his research data.

Malcolm is not aware of any re-use of data that he's shared yet, but he anticipates that due to the nature of his work, it might take some time before this happens.



In addition to publishing with open access and sharing the data underpinning publications, Malcolm also makes preprints of his papers available on bioRxiv whenever possible. bioRxiv is an online preprint server which allows researchers the opportunity to make their findings immediately available to the scientific community, and to receive feedback on draft manuscripts before they are submitted to journals.

Most journals allow posting of manuscripts to preprint servers prior to publication and preprint servers can facilitate the submission of manuscripts to journals, streamlining a researcher's workflow, in addition to promoting open research.

Critical Issues

Recently, Malcolm has been involved in collaborative proteomics research with colleagues in Glasgow. The research is in the field of proteomics and their data goes directly into a proprietary system called Mascot. While this system facilitates good analysis of very large and complex datasets, Malcolm notes that it does not support simple sharing of the (raw) research data.

Proteomics datasets are generally more complex than genomics datasets. This is exacerbated by the wide variety of analytical approaches, tools and pipelines available for analysis. Initiatives (for example the ProteomeXchange consortium) are now attempting to enable better sharing of proteomics data through the implementation of standardized submission and dissemination pipelines for proteomics information. However, this is not yet mature and previously widely-used resources (eg Peptidome and Tranche) have already been lost due to lack of funding.

Leadership

In addition to being an active researcher, Malcolm is also on the management committee of the journal *Parasite Immunology*. He has proposed that the journal should ask authors to share the data which underpins articles they publish. The journal has not yet taken this step due to concerns over

the practicalities of sharing research data, but other journals have moved in this direction and Malcolm hopes that Parasite Immunology will join them.

Malcolm comments on this process:

...when comparing the process of submission of data to repositories like Figshare and Dryad with electronic submission of a paper to a journal, the latter is more tricky and time-consuming. I do not understand why some editors are reluctant to require data submission 'because it is more difficult'...

More Information

Malcolm Kennedy (staff profile page)

<https://www.gla.ac.uk/schools/lifesciences/staff/malcolmkennedy/>

Figshare (generic research data repository) <https://figshare.com/>

Dryad (research data repository for ecological data) <https://datadryad.org/>

bioRxiv (preprint server for Biological Sciences) <https://www.biorxiv.org/>

Wellcome Open Research policy <https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-and-sharing>

This case study was written by Mary Donaldson for the University of Glasgow Research Data Management service.

©MMD October 2018

This resource is made available with a CC-BY 4.0 licence.