Sustaining open research resources – a funder perspective

This is the second in a series of three blog posts which set out the perspectives of researchers, funders and universities on support for open resources. The first was Open Resources, who should pay? In this post, David Carr from the Open Research team at the Wellcome Trust provides the view of a research funder on the challenges of developing and sustaining the key infrastructures needed to enable open research.

As a global research foundation, Wellcome is dedicated to ensuring that the outputs of the research we fund – including articles, data, software and materials – can be accessed and used in ways that maximise the benefits to health and society. For many years, we have been a passionate advocate of open access to publications and data sharing.

I am part of a new team at Wellcome which is seeking to build upon the leadership role we have taken in enabling access to research outputs. Our key priorities include:

- developing novel platforms and tools to support researchers in sharing their research – such as the Wellcome Open Research publishing platform which we launched last year;
- supporting pioneering projects, tools and experiments in open research, building on the Open Science Prize which with the NIH and Howard Hughes Medical Institute;
- developing our policies and practices as a funder to support and incentivise open research.

We are delighted to be working with the Office of Scholarly Communication on the Open Research Pilot Project, where we will work with four Wellcome-funded research groups at Cambridge to support them in making their research outputs open. The pilot will explore the opportunities and challenges, and how platforms such as Wellcome Open Research can facilitate output sharing.

Realising the long-term value of research outputs will depend critically upon developing the infrastructures to preserve, access, combine and re-use outputs for as long as their value persists. At present, many disciplines lack recognised community repositories and, where they do exist, many cannot rely on stable long-term funding. How are we as a funder thinking about this issue?

Meeting the costs of outputs sharing

In July 2017, Wellcome published a new policy on managing and sharing data, software and materials. This replaced our long-standing policy on data management and sharing – extending our requirements for research data to also cover original software and materials (such as antibodies, cell lines and reagents). Rather than ask for a data management plan, applicants are now asked to provide an outputs management plan setting out how they will maximise the value of their research outputs more broadly.
Wellcome commits to meet the costs of these plans as an integral part of the grant, and provides guidance on the costs that funding applicants should consider. We recognise, however, that many research outputs will continue to have value long after the funding period comes to an end. Further, while it not appropriate to make all research data open indefinitely, researchers are expected to retain data underlying publications for at least ten years (a requirement which was recently formalised in the UK Concordat on Open Research Data). We must accept that preserving and making these outputs available into the future carries an ongoing cost.

Some disciplines have existing subject-area repositories which store, curate and provide access to data and other outputs on behalf of the communities they serve. Our expectation, made more explicit in our new policy, is that researchers should deposit their outputs in these repositories wherever they exist. If no recognised subject-area repository is available, we encourage researchers to consider using generalist repositories – such as Dryad, FigShare and Zenodo – or if not, to use institutional repositories. Looking ahead, we may consider developing an orphan repository to house Wellcome-funded research data which has no other obvious home.

Recognising the key importance of this infrastructure, Wellcome provides significant grant funding to repositories, databases and other community resources. As of July 2016, Wellcome had active grants totalling £80 million to support major data resources. We have also invested many millions more in major cohort and longitudinal studies, such as UK Biobank and ALSPAC. We provide such support through our Biomedical Resource and Technology Development scheme, and have provided additional major awards over the years to support key resources, such as PDB-Europe, Ensembl and the Open Microscopy Environment.

While our funding for these resources is not open-ended and subject to review, we have been conscious for some time that the reliance of key community resources on grant funding (typically of three to five years’ duration) can create significant challenges, hindering their ability to plan for the long-term and retain staff. As we develop our work on Open Research, we are keen to explore ways in which we adapt our approach to help put key infrastructures on a more sustainable footing, but this is a far from straightforward challenge.

**Gaining the perspectives of resource providers**

In order to better understand the issues, we did some initial work earlier this year to canvas the views of those we support. We conducted semi-structured interviews with leaders of 10 resources in receipt of Wellcome funding – six database and software resources, three cohort resources and one materials stock centre – to explore their current funding, long-term sustainability plans and thoughts on the wider funding and policy landscape.

We gathered a wealth of insights through these conversations, and several key themes emerged:
All of the resources were clear that they would continue to be dependent on support from Wellcome and/or other funders for the long-term.

While cohort studies (which provide managed access to data) can operate cost recovery models to transfer some of the cost of accessing data onto users, such models were not appropriate for data and software resources who commit to open and unrestricted access.

Several resources had additional revenue-generation routes – including collaborations with commercial entities – and these had delivered benefits in enhancing their resources. However, the level of income was usually relatively modest in terms of the total cost of sustaining the resource. Commitments to openness could also limit the extent to which such arrangements were feasible.

Diversification of funding sources can give greater assurance and reduce reliance on single funders, but can bring an additional burden. There was felt to be a need for better coordination between funders where they co-fund resources. Europe PMC, which has 27 partner funders but is managed through a single grant is a model which could be considered.

Several of the resources were actively engaged in collaborations with other resources internationally that house related data – it was felt that funders could help further facilitate such partnerships.

We are considering how Wellcome might develop its funding approaches in light of these findings. As an initial outcome, we plan to develop guidance for our funded researchers on key issues to consider in relation to sustainability. We are already working actively with other funders to facilitate co-funding and make decisions as streamlined as possible, and wish to explore how we join forces in the future in developing our broader approaches for funding open resources.

**Coordinating our efforts**

There is growing recognition of the crucial need for funders and wider research community to work together develop and sustain research data infrastructure. As the first blog in this series highlighted, the scientific enterprise is global and this is an issue which must be addressed international level.

In the life sciences, the ELIXIR and US BD2K initiatives have sought to develop coordinated approaches for supporting key resources and, more recently, the European Open Science Cloud initiative has developed a bold vision for a cloud-based infrastructure to store, share and re-use data across borders and disciplines.

Building on this momentum, the Human Frontiers Science Programme convened an international workshop last November to bring together data resources and major funders in the life sciences. This resulted in a call for action (reported in Nature) to coordinate efforts to ensure long-term sustainability of key resources, whilst supporting resources in providing access at no charge to users. The group proposed an international mechanism to prioritise core data resources of global importance, building on the work undertaken by ELIXIR to define criteria for such resources. It was proposed national funders could
potentially then contribute a set proportion of their overall funding (with initial proposals suggesting around 1.5 to 2 per cent) to support these core data resources.

**Grasping the nettle**

Public and charitable funders are acutely aware that many of the core repositories and resources needed to make research outputs discoverable and useable will continue to rely on our long-term funding support. There is clear realisation that a reliance on traditional competitive grant funding is not the ideal route through which to support these key resources in a sustainable manner.

But no one yet has a perfect solution and no funder will take on this burden alone. Aligning global funders and developing joint funding models of the type described above will be far from straightforward, but hopefully we can work towards a more coordinated international approach. If we are to realise the incredible potential of open research, it’s a challenge we must address

*Published 26 July 2017 via https://unlockingresearch-blog.lib.cam.ac.uk/?p=1520*

*Written by David Carr, Wellcome Trust ([d.carr@wellcome.ac.uk](mailto:d.carr@wellcome.ac.uk))*