

Researchers argue for standard format to cite lab resources

Research Resource Identifier (RRID) aims to clean up poorly referenced data.

Dalmeet Singh Chawla

29 May 2015

A single, standard format to identify resources such as model organisms and reagents used in published experiments is catching on with researchers and journals, its originators claim. The biosciences would benefit from a universal way of citing research materials, they say.

An article published before peer review at *F1000Research*¹ describes the early fortunes of the Research Resource Identifier (RRID) — a citation format used in papers for referencing antibodies, model organisms and software. The RRID was launched in February 2014 as a pilot study sponsored by the US National Institutes of Health (NIH) in Bethesda, Maryland, and has since made its way into hundreds of papers, says Anita Bandrowski at the University of California, San Diego, who is coordinating the project.

The scientific literature is full of references to reagents, antibodies, tissue samples and software tools, but these are often referenced so poorly that it is hard for other researchers to pinpoint the exact materials used. One study² found that 54% of resources are not uniquely identifiable in publications.

A standard citation format does exist for many types of resource. Researchers can reference the type of *Drosophila* they used by citing the correct ID from the Bloomington *Drosophila* Stock Center at Indiana University; for mice, a similar identifier can be found at the database Mouse Genome Informatics (MGI). But in practice, Bandrowski says, researchers get confused when they have to go to different repositories to find identifiers for each resource, and often cite resources improperly.

One format to rule them all

The new ID format simply adds the prefix 'RRID:' to existing identifiers. Each RRID is stored at a [central website](#), where authors can search for their resource in any of 10 repositories or databases (including the BDSC and the MGI), retrieve the relevant ID, and click a 'cite this' button, which presents correctly formatted text that authors can copy into their papers.

If identifiers don't yet exist for a resource, the portal does not mint fresh ones itself, but helps researchers to create one at in a suitable outpost (such as the *Drosophila* database FlyBase), by clicking on 'Add a Resource' at the central portal. Once created, the portal presents the identifier to the author with the RRID wrapper.

After more than a year, the number of published papers with RRIDs has reached more than 350 — but almost 90% of them are in neuroscience journals. Bandrowski says that the identifiers will break out into other disciplines, and are being adopted by broad-scope journals such as *PLoS ONE*. "They'll be useful if they catch widespread attention and uptake," says Dan MacLean, a bioinformatician at the Sainsbury Laboratory in Norwich, UK.

Because RRIDs are computer-readable, it is possible to automatically pull out lists of resources cited in articles that use the identifiers. Publishing giant Elsevier, for example, has included a box showing 'Antibody data cited for this article' alongside the online version of J.G. Doria *et al.*³. The publisher is looking to extend the idea to cover software tools and model organisms too.

"The concept behind the RRID project is nothing revolutionary, but is critical for science to be done accurately and reproducibly," says Joanne Berghout, who coordinates outreach for MGI.

Whereas RRIDs are focused on standardizing citations of antibodies, model organisms and software, it is only one system in a sea of different standard formats for permanently identifying online scholarly content — acronyms for citation that include DOI, URI, PURL, and hdl (Handle), for example. "There is no single authority that is bringing these together," says Susanna Assunta-Sansone, who works on ways to curate and share reproducible research at the University of Oxford e-Research Centre, UK. Sansone says that heterogeneity should be embraced, and that any identifiers for research objects with aspirations of being universal should build on existing community practices.

Bandrowski believes that introducing RRIDs for biological cell lines is “the next logical step”. She says the team is currently liaising with the NIH to map out all entities that are worth giving identifiers to. The team may also turn to governmental bodies or the commercial sector for support with specific resource types, she says.

Nature | doi:10.1038/nature.2015.17652

References

1. Bandrowski, A. *et al.* *F1000Research* <http://dx.doi.org/10.12688/f1000research.6555.1> (2015).
2. Vasilevsky, N. A. *et al.* *PeerJ* **1**, e148 (2013).
3. Doria, J. G. *et al.* *Neurobiol. Dis.* **73**, 163–173 (2015).