

Learned Publishing, 28: 151–155
doi:10.1087/20150211

Most researchers who have co-authored articles for publication, whether with one collaborator or twenty, have a story about wrangling over the order in which author names appear in the byline. And every journal publisher, large or small, deals regularly with cases of author dispute.¹

In the 1930s, the average number of collaborators on scientific papers was roughly two, and this number remained steady for four decades.² Authorship and collaboration have changed dramatically since the 1970s, and growth in multi-authorship has accelerated, driven both by academic reward systems and the ease of collaboration in the Internet age. By 2000, the average number of authors in articles published in high-ranking medical journals was seven. Before 1975, the maximum number of authors associated with any article in MEDLINE was 38,³ whereas it is not unusual today for scientific publications to list hundreds or thousands of authors.⁴ At the same time, interdisciplinary collaboration has increased, and other forms of scholarly output, including data and software, are now published in citable form.⁵ Some of these new scholarly collaborations, in particular citizen-science projects such as the Sloan Galaxy Zoo, can attract hundreds of thousands of named contributors.⁶

As the average number of authors on scientific articles grows, authorship-related problems, ranging from disputes to outright misconduct, mount. Why, then, do we persist with a practice of attributing scientific contribution that fails to capture the true nature of the underlying collaboration – or, more precisely, to capture who did what? It's not as though the stakes here

INDUSTRY UPDATE

Beyond authorship: attribution, contribution, collaboration, and credit

Amy Brand *Digital Science*
Liz Allen *Wellcome Trust*
Micah Altman *MIT Libraries*
Marjorie Hlava *Access Innovations*
Jo Scott *Wellcome Trust*

Key points

- As the number of authors on scientific publications increases, ordered lists of author names are proving inadequate for the purposes of attribution and credit.
- A multi-stakeholder group has produced a contributor role taxonomy for use in scientific publications.
- Identifying specific contributions to published research will lead to appropriate credit, fewer author disputes, and fewer disincentives to collaboration and the sharing of data and code.



Amy Brand



Liz Allen



Micah Altman



Marjorie Hlava



Jo Scott

© Amy Brand, Liz Allen, Micah Altman, Marjorie Hlava, and Jo Scott 2015

are inconsequential. Who gets credit for discovery or creation has a tremendous impact on people's lives. It affects career advancement and tenure in the academic sphere, and the transparency and integrity of the permanent research record.

When there are multiple authors, we tend to rely on the order in which names are listed to infer lead contribution, but in fact there are no consistent name ordering practices from one field to the next.^{7,8} Even weak conventions around ordering break down in multidisciplinary collaborations when field-specific practices conflict. In fields such as economics, in which the order of names defaults to alphabetical, and it is typically assumed that all authors contribute equally, it has even been shown that you are somewhat more likely to get tenure or win a prestigious prize if your last name begins with a letter earlier in the alphabet.⁹

In fields without the alphabetical order convention, decisions about lead authorship can be especially contentious. How we apportion credit for collaborative works today is highly subjective, open to abuse, and often determined more by laboratory politics or seniority than by actual effort or contribution.¹⁰ In these situations, junior researchers, for whom the reputational stakes are especially high, and those making non-traditional research contributions, such as in the form of data or software code, tend to lose out most on the recognition they deserve.

A separate but related question concerns what qualifies an individual contributor for authorship status, and this, too, is often contested. Within the biomedical community, the authorship guidelines produced by broadly recognized organizations such as the International Committee of Medical Journal Editors (ICMJE) and the Committee on Publication Ethics (COPE) focus primarily on who should be listed as author, and what processes should be used for authorship disputes and corrections. Within the field of medicine, conventions vary across publication venues and institutions.¹¹ For example, Harvard Medical School's authorship rules specify that, 'Everyone who has made substantial intellectual contributions to the work should be an author' (see <http://hms.harvard.edu/about-hms/integrity-academic-medicine/hms-policy/faculty-policies-integrity-science/authorship-guidelines>). Similarly, the *Proceedings of the National Academy of Sciences (PNAS)* require that 'authorship must be limited to those who have contributed substantially to the work' (see <http://www.pnas.org/site/misc/iforc.pdf>).

The ICMJE policy, on the other hand, limits authorship to those who make

substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpre-

tation of data for the work; *and* drafting the work or revising it critically for important intellectual content; *and* final approval of the version to be published; *and* agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. (See <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html> emphasis added)

Awareness of problems with the conventional authorship model is by no means new. The topic received a great deal of attention in the late 1990s, for example, in the context of accountability in medical journal publishing. The work of Drummond Rennie on this subject was particularly influential.¹² In a 1997 article in the *Journal of the American Medical Association (JAMA)*, Rennie and co-authors wrote:

The system of authorship, while appropriate for articles with only one author, has become inappropriate as the average number of authors of an article has increased; as the work of coauthors has become more specialized and relationships between them have become more complex; and as both credit and, even more, responsibility have become obscured and diluted. Credit and accountability cannot be assessed unless the contributions of those named as authors are disclosed to readers, so the system is flawed. We argue for a radical conceptual and systematic change, to reflect the realities of multiple authorship and to buttress accountability. We propose dropping the outmoded notion of author in favor of the more useful and realistic one of contributor. This requires disclosure to readers of the contributions made to the research and to the manuscript by the contributors, so that they can accept both credit and responsibility.

Whether from the perspective of credit or accountability, clearly we need a better system for representing collaborative contribution to published works – something more akin to film credits.¹³ In the intervening years since Rennie's radical call to action, several medical and life science publishers have started to collect contribution statements for multi-authored works. Some publishers, such as the American Association for Cancer Research (AACR) and the Public Library of Science (PLOS), ask authors to select roles from a predefined list. Others, such as Nature, invite or require free-text contribution statements, yet many publishers who collect role information from authors do not even publish it.

What the scholarly publishing community still lacks is coordination among contributorship efforts. In the absence of standardization and coordination, the infor-

Table 1. CRediT – contributor role taxonomy

Header: This taxonomy provides a high-level classification of the diverse roles performed in the work leading to a published research output in the sciences. Its purpose is to provide transparency in contributions to scholarly published work, to enable improved systems of attribution, credit, and accountability.

The classification includes, but is not limited to, traditional authorship roles. That is, these roles are not intended to define what constitutes authorship. Rather, the roles are intended to apply to all those who contribute to research that results in scholarly published works, and it is recommended that all tagged contributors be listed, whether they are formally listed as authors or named in acknowledgements.

An individual contributor may be assigned multiple roles, and a given role may be assigned to multiple contributors. When there are multiple people serving in the same role, a degree of contribution may optionally be specified as ‘lead’, ‘equal’, or ‘supporting’. It is recommended that corresponding authors assume responsibility for role assignment, and that all contributors be given the opportunity to review and confirm assigned roles.

Term	Definition
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims
Methodology	Development or design of methodology; creation of models
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components
Validation	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs
Formal Analysis	Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools
Data curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse
Writing – Original Draft	Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation)
Writing – Review & Editing	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team
Project Administration	Management and coordination responsibility for the research activity planning and execution
Funding acquisition	Acquisition of the financial support for the project leading to this publication.

mation about contribution that publishers may already be collecting is not making its way into the metadata network that underpins our citation and credit systems. What we need is a controlled vocabulary of contributor roles and mechanisms for capturing contribution tags within the scholarly metadata ecosystem.

Imagine publishers collecting structured information about contribution in a standard format. Imagine, further, that this information is associated with the article DOI, via CrossRef, and with ORCID author identifiers. We would then have the infrastructure in place to track not only who authored which publications, but also who contributed what to each publication that names the individual as a contributor. With this infrastructure in place, it would eventually be possible to devise more precise, author-centric credit and impact tracking tools, on which the byline order of author names would have no bearing.

In May 2012, we hosted a workshop at Harvard University to explore this topic with representatives of the publishing, funding, and academic worlds.¹⁴ A key outcome of this workshop was the commitment by a sub-group of attendees to devise a high-level contributor role taxonomy for the sciences. We drafted a preliminary taxonomy by analyzing acknowledgments and free-text contribution statements, and conducted a survey study in partnership with several publishers to gauge the feasibility of asking corresponding authors to assign the roles. The results of the study were overwhelmingly positive. These efforts are described in a Nature commentary article published last year.¹⁵

Based on the success of the 2013 study, we are partnering with two information industry standards organizations, the Consortia Advancing Standards in Research Administration Information (CASRAI) and the US-based National Information Standards Organization (NISO), to achieve broader community consultation in refining the taxonomy and testing its fit with a range of scientific fields. During the latter half of 2014, a 17-person working group composed of representatives from several publishers, funders, and universities met monthly under the auspices of CASRAI to review and refine each of the roles and role descriptions. This effort adopted the name Project CRediT, and the project overview and the taxonomy itself are available at <http://projectcredit.net>. Once we reached consensus on the 14-term taxonomy – see Table 1 – we opened the project up for public comment and received over 100 responses using an online feedback form. Researchers constituted 75% of the respondents, and the group was fairly diverse in terms of geographic make-up, with the biological sciences being more strongly represented than other researcher areas.

As with the earlier study, the results of the feedback process for this version of the taxonomy were highly encouraging. A clear majority of respondents agreed with all of the proposed terms. Most of the questions that arose concerned confusion over whether the taxonomy was explicitly intended to specify which types of contribution qualify for authorship status, when in fact that was never the intention. As stated in the taxonomy header:

The classification includes, but is not limited to, traditional authorship roles. That is, these roles are not intended to define what constitutes authorship. Rather, the roles are intended to apply to all those who contribute to research that results in scholarly published works, and it is recommended that all tagged contributors be listed, whether they are formally listed as authors or named in acknowledgements.

Among the other recommendations to emerge from the public consultation process were: (1) to adopt a coarse-grained degree of contribution, as an optional tag to be used in conjunction with a contributor role when more than one contributor serves in the same role; and (2) to have corresponding authors be responsible for role assignment, but only with review and confirmation by all contributors. The taxonomy header captures both of these recommendations as follows:

An individual contributor may be assigned multiple roles, and a given role may be assigned to multiple contributors. When there are multiple people serving in the same role, a degree of contribution may optionally be specified as 'lead', 'equal', or 'supporting'. It is recommended that corresponding authors assume responsibility for role assignment, and that all contributors be given the opportunity to review and confirm assigned roles.

We are currently working with several publishing partners and providers of manuscript-tracking and author-submission systems to undertake early implementations of the taxonomy. We expect a number of these implementations to be up and running later in 2015. An early pilot of the taxonomy by Mozilla Science Labs is already underway. It uses the taxonomy in a set of digital contributorship badges maintained at the browser level, as described in a recent *Science Magazine* news piece.¹⁶

As word about the Project CRediT taxonomy has spread through ongoing conference presentations and coverage in leading community blogs,¹⁷ new efforts are focused on implementation pathways, including integrating the taxonomy into the National Library of Medicine's Journal Article Tag Suite (JATS) DTD (<http://jats.nlm.nih.gov/about.html>). The Force11 com-

munity (<https://www.force11.org/about>) has created a new working group to look at attribution implementation for all research products.

With these aligned efforts, and the groundswell of interest among researchers, funding agencies, academic institutions, and editors in increasing the transparency of research contributions, standardized contribution tagging, while still early-stage, is gaining firm footing in scholarly journal publishing. If this initiative is ultimately successful, there will be far fewer author disputes, and fewer disincentives to collaboration and the sharing of data and code, for example, because those contributions will be more reliably recognized. Among the less obvious benefits is enhanced mineable information on research expertise, for the purposes of research networking and peer-reviewer identification.¹⁸ Hence these efforts could positively influence both the cooperative culture of research, and academic incentive structures more generally. We invite authors and publishers alike to follow the example below in describing contribution using the CRediT taxonomy.

Author statement and acknowledgements

The first author named is lead and corresponding author. All other authors are listed in alphabetical order. We describe contributions to the paper using the taxonomy provided above. *Writing – Original Draft*: A.B. and M.A.; *Writing – Review & Editing*: M.A., A.B., and M.H.; *Conceptualization*: L.A. and A.B.; *Investigation*: L.A., A.B., M.A., and M.H.; *Methodology*: M.A. and J.S.; *Formal Analysis*: M.A. and J.S.; *Project Administration*: L.A. and A.B.; *Funding Acquisition*: L.A. and A.B.

The work described in this article was supported by the Wellcome Trust and Digital Science.

The authors would like to acknowledge the other members of the CASRAI working group who provided critical review of the taxonomy, but are not responsible for the content of this article: Helen Atkins, David Baker, Monica Bradford, Todd Carpenter, Jon Corsant-Rikert, Jeffrey Doyle, Melissa Haendel, Daniel S. Katz, Veronique Kiemer, Nettie Lagace, Emile Marcus, Walter Schaeffer, Gene Sprouse, and Victoria Stodden.

References

- Levsky, M.E, Rosin, A., Coon, T.P, Enslow W.L., and Miller, M.A. 2007. A descriptive analysis of authorship within medical journals, 1995–2005. *Southern Medical Journal*, 4(1): 371–375. <http://dx.doi.org/10.1097/01.smj.0000257537.51929.4b>
- Clarke, B.L. 1964. Multiple authorship trends in scientific papers. *Science*, 143(3608): 822–824. <http://dx.doi.org/10.1126/science.143.3608.822>
- Number of Authors per MEDLINE®/PubMed® Citation, <http://www.nlm.nih.gov/bsd/authors1.html>
- King, C. 2012. Multiauthor papers: onward and upward. *ScienceWatch Newsletter*. http://archive.sciencewatch.com/newsletter/2012/201207/multiauthor_papers/
- Altman, M. and Crosas, M. 2013. The evolution of data citation: from principles to implementation. *IASSIST Quarterly*, 37: 62.
- Raddick, M.J., Georgia, B., Pamela, L.G., Lintott, C.J., Murray, P., Schawinski, K., Szalay, A.S., and Jan Vandenberg, J. 2010. Galaxy zoo: exploring the motivations of citizen science volunteers. *Astronomy Education Review*, 9(1): 010103. <http://dx.doi.org/10.3847/AER2009036>
- Academic authorship, 22 January 2015. In *Wikipedia, The Free Encyclopedia*. Retrieved 15.47, 7 February 2015, from http://en.wikipedia.org/wiki/Academic_authorship
- Wager, E. 2007. Do medical journals provide clear and consistent guidelines on authorship? *Medscape General Medicine*, 9(3): 16.
- Einav, L. and Yariv, L. 2006. What's in a surname? the effects of surname initials on academic success. *Journal of Economic Perspectives*, 20(1): 175–187. <http://dx.doi.org/10.1257/089533006776526085>
- Dance, A. 2012. Who's on first? *Nature*, 489: 591–593. <http://dx.doi.org/10.1038/nj7417-591a>
- Bosnjak, L. and Marusic, A. 2012. Prescribed practices of authorship: review of codes of ethics from professional bodies and journal guidelines across disciplines. *Scientometrics*, 93(3): 751–763.
- Rennie, D., Yank, V., and Emanuel, L. 1997. When authorship fails. A proposal to make contributors accountable. *JAMA*, 278(7): 579–585. <http://dx.doi.org/10.1001/jama.1997.03550070071041>
- Patterson, M. 2007. Roll credits: sometimes the authorship byline isn't enough. Guest PLoS Blog by Michael Molla and Tim Gardner. Posted 6 November 2007. <http://blogs.plos.org/plos/2007/11/roll-credits-sometimes-the-authorship-byline-isnt-enough/>
- IWCSA Report 2012. Report on the International Workshop on Contributorship and Scholarly Attribution, 16 May 2012. Harvard University and the Wellcome Trust. http://projects.iq.harvard.edu/attribution_workshop.
- Allen, L., Scott, J., Brand, A., Hlava, M., and Altman, M. 2014. Credit where credit is due. *Nature*, 508: 312–313. <http://dx.doi.org/10.1038/508312a>
- Dalmeert, S.C. 2014. Could digital badges clarify the roles of co-authors. *Science*, published online 3 November 2014. <http://news.science.org/scientific-community/2014/11/could-digital-badges-clarify-roles-co-authors>
- Meadows, A. 2014. An interview with Amy Brand on a proposed new contributor taxonomy initiative. *Scholarly Kitchen*, published online 20 August 2014. <http://scholarlykitchen.sspnet.org/2014/08/20/an-interview-with-amy-brand-on-a-proposed-new-contributor-taxonomy-initiative/>
- Academy of Medical Sciences. 2015. Team Science Working Group – project update February 2015. <http://www.acmedsci.ac.uk/download.php?f=file&i=30834>

Amy Brand (corresponding author)

Vice President

Digital Science

25 First Street Suite 104

Cambridge MA 02141, USA

a.brand@digital-science.com

Liz Allen

Wellcome Trust

Micah Altman

MIT Libraries

Marjorie Hlava

Access Innovations

Jo Scott

Wellcome Trust