Major challenges on authorship and concept of authorship - why is something more needed on contributorship?

Authors

- **Nicole Vasilevsky**
  ⓘ [0000-0001-5208-3432](#) · ○ [nicolevasilevsky](#) · 🐦 [n_vasilevsky](#) · ✖ 5 roles
  Oregon Clinical and Translational Research Institute, Oregon Health & Science University; Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University · Funded by Grant# NCATS U24TR002306

- **Julien Colomb**
  ⓘ [0000-0002-3127-5520](#) · ✖ 3 roles
  Humboldt-Universität zu Berlin, SFB1315, Institut für Biologie, Charitéplatz 1, 10117 Berlin · Funded by The German Research Foundation, Grant Number 327654276: SFB 1315

- **Scott Edmunds**
  ⓘ [0000-0001-6444-1436](#) · ✖ 2 roles
  GigaScience

- **Karen Gutzman**
  ⓘ [0000-0001-6331-4451](#) · ✖ 2 roles
  Galter Health Sciences Library and Learning Center, Northwestern University Feinberg School of Medicine · Funded by Grant# UL1TR001422

- **Melissa Haendel**
  ⓘ [0000-0001-9114-8737](#) · ○ [mellybelly](#) · ✖ 4 roles
  Oregon Clinical and Translational Research Institute, Oregon Health & Science University; Linus Pauling Institute, Oregon State University · Funded by Grant# NCATS U24TR002306

- **Daniel Himmelstein**
  ⓘ [0000-0002-3012-7446](#) · ○ [dhimmel](#) · ✖ 2 roles
  University of Pennsylvania

- **Mohammad Hosseini**
  ⓘ [0000-0002-2385-985X](#) · ✖ 2 roles
  Dublin City University

- **Violeta Ilik**
  ⓘ [0000-0003-2588-3084](#) · ○ [vioil](#) · ✖ 3 roles
  Library, Columbia University

- **Barbara Kern**
  ⓘ [0000-0002-1184-911X](#) · ✖ 1 role
  University of Chicago

- **Ehsan Mohammadi**
  ⓘ [0000-0003-3481-6991](#) · ✖ 1 role

University of South Carolina

- **Lisa O'Keefe**
  · ⚒ 1 role
  Galter Health Sciences Library and Learning Center, Northwestern University Feinberg School of Medicine

- **Juliane Schneider**
  ⓘ 0000-0002-7664-3331 · ○ pitviper6 · ⚒ 2 roles
  Harvard University

- **Britton Smith**
  ⓘ 0000-0002-6804-630X · ⚒ 1 role
  University of Edinburgh

- **Samantha Teplitzky**
  ⓘ 0000-0001-7071-332X · ○ samteplitzky · ⚒ 2 roles
  University of California, Berkeley

- **Marijane White**
  ⓘ 0000-0001-5059-4132 · ⚒ 1 role
  Library, Oregon Health & Science University · Funded by Grant# U24 TR002306

- **Kristi Holmes**
  ⓘ 0000-0001-8420-5254 · ○ kristiholmes · ⚒ 6 roles
  Department of Preventive Medicine, Northwestern University Feinberg School of Medicine; Galter Health Sciences Library and Learning Center, Northwestern University Feinberg School of Medicine · Funded by Grant# U24TR002306, UL1TR001422

Is authorship sufficient for today's collaborative research? A call for contributor roles

Introduction

Background perspectives on authorship

Scientific authorship generally consists of publishing academic findings via journal articles, book chapters, and monographs [1]. In academic collaborations within science and engineering where co-authorship is the norm, authorship status is attributed to those who have made a significant contribution to certain tasks within the project [2]. Beyond being used as an instrument to recognize contributions, authorship is also used to hold contributors accountable for the accuracy and integrity of published claims [3].

Receiving recognition through authorship has long been the entrenched currency of the scholarly realm. Even so, it has long been recognized that assigning authorship credit is neither a fair nor uniform process [4]. Historically, concerns about authorship credit expressed by academics in the literature centered around what was seen as an egregious problem of awarding authorship credit to those who did not deserve it, and consequently diminishing the contributions of the first, or primary authors. Terms such as profligate, honorary, and courtesy authorship describe various forms of authorship abuse. Some of the proposed solutions to address these problems include defining criteria for authorship (e.g. by the Vancouver group since 1987), providing details of contributions [5], and assigning a rating to authors' efforts [6]. These solutions often stemmed from a desire to narrow the

criteria for authorship, and to clarify roles or the extent of contributions to prevent awarding author status to those who did not deserve it. Nevertheless, applying these solutions in practice may contribute to other tensions.

Contributorship: making the contribution the focus, rather than the resulting paper

Authorship versus contribution

While the definition and the exact role of an author is somewhat ambiguous, tracking contributorship in publications is intended to more explicitly define and give credit to contributors to a work. Contributors can contribute to a study and/or publication in various ways, and may not necessarily be involved in the writing or revision of the manuscript. Traditional roles of contributors may include the planning, conducting, and reporting the work. Non-traditional roles may be more varied. For example, a primary technician may perform assays such as human brain autopsies, whole brain hemisphere sectioning, immunohistochemical staining, stereological analysis and and assist with publication preparation. Or a librarian may deliver invited lectures and served as a consultant for libraries to help them establish clinical and translational research support services. These non-traditional roles can be essential to the success of scholarship, but are often not credited with authorship as often as more traditional roles.

> **Assigning authorship credit can easily go awry, damaging the reputation of authors, institutions, journals and science in general, as exemplified in [7] where a published work was retracted because of an authorship dispute. Ongoing questions also persist across disciplines regarding credit for the staff who performed most, if not all experiments that lead to knowledge and breakthroughs, as demonstrated in the debate on "Who really made Dolly?" in the Guardian [8]: "You get some papers where the authors haven't done a scrap of work themselves, it's all down to the technicians acknowledged at the back."**

Modern research is interdisciplinary, reflecting a team science approach where the skills needed to conduct reliable science are often specialized [9]. In this dynamic where various contribution-types are required, revamping our understanding of authorship, credit, and recognition of individual efforts in academia seems necessary [10]. Rather than coming from a place of censure, we propose a continuum in which contributions from a team of people could be welcomed and recognized.

Challenges of authorship

Ethical challenges

As authorship credit remains the single most important form of recognizing individual contributions, tensions around its definition and enforcement remain challenging to address. Many guidelines such as those provided by the Council of Science Editors [11] and The International Committee of Medical Journal Editors [12] suggest that authors should have made a 'significant contribution' to the study. Nevertheless, what constitutes a 'significant contribution' is ambiguous and difficult to formally define [13]. Furthermore, while a strict application of suggested authorship criteria within these guidelines may exclude contributors who made important but non-standard contributions including dataset management, software and protocol development [14], a lax attitude towards authorship criteria might lead to inflated bylines and hyperauthorship [15].

While modern science needs the participation of a range of contributors, in recent decades a steady increase in the average number of co-authors per publication [16] has contributed to major ethical issues. For instance, in the presence of more co-authors, addressing ethical challenges in the distribution of authorship, acknowledgment credit [17], and handling authorship order [18]would be

more challenging. Similarly, with more authors in the byline, ambiguities in relation to individual and shared responsibilities are much more pronounced [19]. As such, questions about the attribution of authorship status to various contributors remain difficult to answer. For example, it is not clear whether Principal Investigators always deserve authorship status [20], or, should graduate students, research technicians, project/program managers, and core lab scientists be included as authors or only mentioned in the acknowledgment section. Moreover, the role of non-academic contributors such as citizen scientists [21], [22] and community-based partnerships seems difficult to recognize [23]. Within interdisciplinary projects, other issues such as dissimilar norms in the distribution of authorship credit and author's order may be present as well. Some fields list authors in alphabetical order and others based on the degree of contribution. It is common in certain disciplines, such as physics, to have hundreds of authors on a paper, whereas in other fields like humanities, one or very few authors may contribute to publications.

Social challenges and Authorship Criteria

Authorship practices have real consequences, as observed when applying authorship credit for tenure and promotion. While distribution of authorship credit is not straightforward, principles and standards that are suggested for articles involving one or two individuals are similar to articles published by team science endeavours with hundreds or thousands of contributors [24]. To prevent tensions, it is often advised that roles and duties of individuals should be agreed upon and discussed at the outset of a study [17]. However, this can be a challenge as research personnel and the work may change over the course of a project. Furthermore, in most cases explicit discussions about awarding credit occur in response to issues that arise, hence, minimizing the usefulness of discussions [25].

> **Longer authorship lists complicate measuring individual contributions [26], further disincentivizing authorship practices that recognize more than the most involved researchers on a project.**

Additionally, the participation of junior and senior contributors with unequal authority and institutional influence, contribute to other forms of authorship abuse [27]. "Honorary" and "gift" authorship, involve "naming as an author, an individual who does not meet authorship criteria" [28]. In severe cases, individuals are listed without having made any contributions and are included as authors to add perceived prestige or credibility to the research [13]. In contrast, sometimes it is the lack of giving due credit to those who deserve it (so-called ghost authorship) that raises concerns. Junior scholars or researchers from the industry who made notable contributions to a project are among common ghost-authors [29,30]. Gender disparity in the distribution of authorship credit is another social challenge. Underrepresentation and lower visibility of women in publications is reported in male-dominate research areas such as Computer Sciences [31], Political Sciences [32], and Neurosurgery [33]. Even in fields such as Higher Education where the gender composition of scholars is more balanced, gender inequity is still noticeable [34]. Women publish fewer articles, tend to receive fewer citations and are less likely to occupy important positions of the byline. When it comes to contribution types and labor roles, women across all academic ages are more often involved in performing experiments, which are associated with academically younger scholars [35]. Even in cases where authors made equal contributions, female authors are often not receiving first authorship positions [36].

There are a number of guidelines on authorship and scholarly works. In 1985 the International Committee of Medical Journal Editors (ICMJE) outlined guidelines on authorship, which have evolved and been updated since [12]. The ICMJE list specific criteria that must be met for authorship including conceptualization of the work, acquisition or analysis or interpretation of the data, drafting the text, approval of the draft, and responsibility for the published content. With respect to authorship versus contributorship, the ICMJE classifies project members who do not participate in the four authorship criteria above as "non-author contributors". Such an approach works for authorship decisions unless

one who makes "substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work is not included in drafting the work or revising it critically for important intellectual content" [37]. The guidelines describe work that alone qualify a contributor for authorship, such as acquisition of funding, leadership of a research group, administrative support, and writing support. The ICMJE recommends that such non-author contributors be acknowledged and their contributions to the work specified. In addition to the ICMJE, the Committee on Publications Ethics has played a significant role in this area, contributing guidelines on "authorship and contributorship" [38]. Yet another important work in this area is the 2006 "White Paper on Publication Ethics" the Council of Science Editors which is updated on a rolling basis. [39].

Technical challenges

Measuring research contributions in a systematic way is an important issue not only for authors but also universities and scientific institutions [40] [41]. However, institution and author name disambiguation have been a challenge, including proper assignment of authorship credit with the use of machine-readable data. In response, some unique identifiers have been proposed to overcome these challenges such as ORCID [43] for authors and Research Organization Registry [44] for institutions, among others. As academics move through their careers, their name, position and affiliations may change. Tracking these changes so that their entire output can be discovered easily is made difficult through proprietary publishing models requiring different formats for names and citations, multiple profiles systems and the proliferation of persistent identifiers (PIDs) attached to a person, affiliation or citation. Authorship information that is siloed or suffers from multiple PIDs can negatively affect metrics, which is crucial to academic promotion, and puts a burden on authors to try and track multiple sites through varying formats to accurately represent their output. In addition, as research becomes more interdisciplinary, and multi-site studies are encouraged by funders, the discipline and the role of one person may change depending on the project.

These issues could be mitigated by the adoption of standards and formats across disciplines and institutions, and allowing at least the personal data from any type of institutional profile system (proprietary or open) to be harvested and used by their researchers to create consistent, comprehensive views of their work. For a better understanding of their contribution to research, adoption of a standard vocabulary for types of attribution would be useful. Persistent identifiers are a critical component to linking persons to their research artifacts and are a critical component of the research process as well as the overall knowledge graph. PIDs should be created with care, or they add to the burden of disambiguation between people, versions of papers, and institutions. Several resources aggregate information about scholars and researchers, and sometimes provision their own PIDs and sometimes reuse existing PIDs. A detailed look at a subset of such resources is outlined in Table 1; the highlighting indicates the openness of the data, from completely open resources (green), to various flavors of partially open data (yellow), to closed data (red).

TABLE GOES HERE

Table 1. Constructing a scholarly graph. A non-comprehensive list of resources in use that can contribute to the graph of scholarship. The colors indicate whether the data are easily available for reuse via API: green - the data are open and freely available; orange - the data is partially closed; and red - the data is closed/inaccessible. The function column describes the primary function of the resource. The data license is indicated, if it could be found, with the accompanying URL to the webpage that describes the license for the respective resource. The final column indicates which Persistent IDs (PIDs) are used by the respective resource: author/contributor, organizational affiliation, artifacts (manuscripts and other scholarly products), and funding source. Note that wikidata scholia is using wikidata as a data source, and that ORCID information can be sent to wikidata automatically, although there is no "statement" for funding yet.

Shifting the focus to contributorship

Authorship versus contributorship

Scholars contribute knowledge by developing various research outputs such as publications, datasets, software, and protocols. Increasingly, large research funders like the US National Science Foundation [44] and the US National Institutes of Health [45] consider nontraditional research outputs as important research products to communicate and track. Although there is a real lack of understanding about how best to address authorship for non-published works, such as datasets [46] [47], that can aid translation of knowledge.

The definition and exact role of authors in traditional publications can be ambiguous, and therefore. tracking contributorship enables more explicit definition and credit to contributors for their role on a given work. Contributors can make contributions to a study and/or publication in various ways, and may not necessarily be involved in the writing or revision of the manuscript. Traditional roles of contributors may include the planning, conducting, and reporting of work. Non-traditional roles may be more varied. For example in a basic science research lab, a technician may write and track the protocols, care for the animals and prepare the lab reagents that are needed for experiments that are ultimately published as figures. A librarian may provide expert search services, guide research data management and preservation in the institutional repository. These non-traditional roles can be essential to the success of a project, but can often not be credited with authorship as more traditional roles.

Making contributorship work in systems

More nuanced characterization and contextualization of contributions is a recognized need by the scholarly community and a number of efforts are underway. Perhaps most well-known is the CRediT taxonomy, a high level standardized vocabulary that contains 14 roles for use in representing scholarly contributions to research outputs [48], [49], [50]. This taxonomy has been incorporated into several workflows, including journal submission and review systems (e.g., PubSweet, Scholar One, ReView) and other tools such as Rescognito and OpenConf [51]. The Contributor Role Ontology (CRO) was developed as an extension of the CRediT taxonomy, and consumes and expands the contributor roles to provide a structured representation of contribution roles in research and scholarship, which is designed for crediting persons or organizations. The CRO is an open-source, community-developed ontology containing over 50 terms [52]. The first iteration of the CRO was developed by the FORCE11 Attribution Working Group (https://www.force11.org/group/attributionwg) and implemented into the OpenVIVO scholar profile system for the 2016 Force11 conference. As noted in the paper by Ilik et al. "this ontology extends the contributions to scholarship beyond manuscript authorship to capture the broadening of researchers' participation in scientific discoveries that have not been previously recognized by traditional measures of scholarly impact" [53]. The work done by the FORCE11 Attribution Working Group along with the OpenVIVO Task Force members at the time included reviewing existing scholarly contribution taxonomies and exploring ways to extend the CRediT taxonomy to create a prototype contributorship model that covers a wide selection of fields of research. The CRO is a component of the Contributor Attribution Model (CAM), an ontology-based specification for representing information about contributions made to research-related artifacts. The CAM refines earlier work and has been expanded to include the information model and tools and straightforward guidance for implementation [54].

> **Individual incentives**: increased recognition and validation of skills and expertise, easy pathways to complete routine workflows > >(e.g., building a CV, completing annual reporting and review activities, reporting to funders); professional branding such as tools to describe or explaining expertise and skills to the public, and increased ease of identifying and engaging a professional network of colleagues and experts.
> **Organizational incentives**: better understanding of expertise capacity on campus, more equitable decisions related to promotion and/or tenure, identification of collaborative people for team-based science, easier methods for recognizing and describing non-publication impacts, illumination and interventions to address disparities in research roles (select references include [35] [55], [56]).
> **Public incentives**: recognition of community partners and their roles on research projects, better understanding and identification of local experts across different organizations, clearer communication of what roles are required in research.
> **Scholarly community incentives**: identification of experts in a field for a specific purpose (clinical specialty, speaking at a conference or to the media, etc.), impartial selection and nomination for awards and committees, highlight research projects and teams that can directly demonstrate implementation and translation to impact.

Table 2. Incentivization of contributorship. Regardless of whether people want to better credit a range of contributor roles, successful incorporation of contributor roles will require culture change and incentives to make this easier for a wide range of relevant stakeholders.

***Expanding measures of success It should be noted that improving the characterisation and contextualisation of contributions will not automatically improve person-level assessment processes. However, incentives clearly exist across stakeholder groups, as highlighted in Table 2. The reward system of science has for so long been solely reliant on authorship. This has disincentivized innovation in approach and the systems used in routine academic workflows, such as publishing, reporting to funders, annual faculty reporting, hiring, and promotion and tenure. As long as scientists are being hired and promoted based on the number of publications, author order, and impact factor of journals, more accurate identifiers of contributions would have limited impact on scientific evaluation and promotion processes. Even researchers based in non-academic institutions report similar patterns in evaluation and promotion [57]. In other words, as long as institutions have not integrated accurate models of contribution into their workflows, journals' adoption alone is not going to benefit the scientific community. Increasingly, there are examples of contributor roles being incorporated into academic assessment workflows through reporting and promotion processes. One such example of this is the Team Scientist Track at Northwestern University Feinberg School of Medicine. Team Scientists on the track "make substantial contributions to the research and/or educational missions of the medical school [...] engage in team science. Their skills, expertise and/or effort play a vital role in obtaining, sustaining and implementing programmatic research." [58]

Making contributorship work: what's needed?

Addressing the challenges (past & current)

A number of strategies to give credit to authors have been developed and implemented (Table 3) to help ensure that everyone receives fair and transparent credit for their contributions, including giving specialist contributors (e.g. data or software development roles) more weighting within their communities. Badges to acknowledge open science practices have been used by the Open Science Foundation to provide incentives for researchers to share data, materials, or to preregister [59]. A similar approach was adopted by the Mozilla Science Lab, with ORCiD, the Wellcome Trust, Digital Science, PLOS and BioMed Central, to create the Paper Badger widget to use open badges to assign digital credentials to contributions on academic papers. The 14 different badges describing the 14 contribution types appear on the article as well as on the author's ORCiD page, and are JSON packages containing metadata validating the badge. Two journals, GigaScience and Journal of Open Research

Software from Ubiquity Press trialled adding the Paper Badger widget to their papers, and although development currently is stalled, this open source project is available for anyone to reuse and take it further [60]. The Author Contribution Index (ACI) [61] aims to circumvent the issue of author order by allowing authors to quantify their contribution through a contribution percentage.

INSERT TABLE 3

What is required for adoption of a contribution-based approach?

A key aspect of adoption of any strategy for greater incorporation of contributor recognition is to lower the barrier of use. Researchers have a number of everyday challenges [62], [63], [64], [65] that can be frustrating and time consuming; the production of scholarly works among those contributing significantly to this burden [66]. Authoring tools like Overleaf (Overleaf) or Manubot (Manubot) (used in the production of this work) create files which could be exported in different formats depending on the publisher's request. However, non-publication research artifacts (datasets, software, materials, protocols, etc.) have less well-established workflows to collect and present structured metadata about the work, and their authors, to ensure that they are part of the scholarly commons.

Ideally, each research product or artifact (e.g., manuscripts, datasets, software, grant applications, reagents, and protocols, to name a few) should have a way to list contributors and their contributions, with many reflecting traditional authorship roles. This information should be held in a machine operable format and linked to the researcher PID. To advance this, technical and social advancements are required and must reflect the diversity of stakeholders who will use such an approach. Perhaps paramount is to define this format together with the different stakeholders, especially publishers and data aggregators, to ensure the information can be linked back to researcher profiles in a trusted and automated way. Only then can other stakeholders be asked to integrate strategies to collect and present contributions. Important to consider, is demonstrating the usefulness of the contributors list before its large-scale adoption. Finally, to support incorporation of contributor roles into academic workflows, tools to make the creation of these contributor lists as easy and re-usable as possible must be developed, and make sure the information is collected in an appropriate format.

Clearly, significant effort has been dedicated to the creation and acculturation of the CRedIT taxonomy (now available as an OWL implementation file [67] to facilitate incorporation into information systems) and the subsequent CRO ontology. But only what can be counted counts, and contribution information must be measured on a large scale. To this end, practical use of these ontologies should be defined and guidance created [68]. Publication information leverages an XML format technical standard called the Journal Article Tag Suite [69] to describe elements of a journal article. The National Information Standards Organization (NISO) is currently incorporating CRedIT into the JATS standard. The current recommendations [70] needs to be further enhanced with the incorporation of a resolvable URI for the CRedIT roles, as well as expansion of contributor role types to reflect roles related to data or other critical activities in modern research. Moreover, different research outputs use a variety formats for their author list, which were designed for better human writability and simplicity (for example YAML in Manubot or JSON in Zenodo). Therefore, it may be more efficient to establish mechanisms to translate the information from one format to another. As an example, one can get inspiration from the integration between Overleaf and F1000Research, where the author list written in the Latex format is automatically imported in the publisher's workflow. Ultimately, information must be accessible and computer readable to incorporate in information systems (e.g., research profiling systems, aggregators, and institutional or funder statistics). Because the ecosystem of research scholarly communication is complex, the process of defining best practices takes time and effort.

Global aspects of adoption

A number of cultural aspects must be addressed for broad adoption of contributor roles. Currently, systems that allow for annotation of contribution roles only do so as the result of an assertion on the part of the individual. Researchers may be unaware of the advantages (or existence) of contributorship approaches such as CRedIT and/or lack straightforward ways to incorporate them into their workflow. This will likely change over time as funders champion efforts to make research results and data more available. While pressure from funders and publishers can trigger change, incentives on the individual level can lead to better engagement and adoption. Such reward strategies, like badges, have been modestly successful, suggesting that further changes in the funding schemes and publisher mandates can help trigger necessary cultural changes to enable a more nuanced understanding and establishment of contributor roles and credit.

For instance, some countries like China, Mexico and Vietnam offer cash-per-publication rewards to authors that are directly linked to impact factor of the journal published in. In China these can be extremely lucrative, with reports of Universities offering $45,000 USD for publications in the highest ranked journals [71]. This is on top of local and central government rewards. As an example In Shenzhen in 2014, the updated "National Leading Talent" and "Peacock" scheme for recruiting overseas high-level talent offered 3M RMB (about $430,000 USD) to first and corresponding authors of papers published in Nature or Science. This extreme commoditization of authorship has increased pressure to inflate the number of joint-first and joint corresponding authors, as well as gift authorship and ghost-writing of fake papers [72]. The ICJME guidelines state the role of the corresponding author is to take care of all the administrative requirements and communication with the journal, but there is a misunderstanding that the most senior authors should have this position, possibly because this role is awarded with financial and other benefits. Unfortunately, confusion of the senior author role and the guidance and pressure authors are under to be a corresponding author is an example that directly contradicts ICMJE guidelines. To help tackle this some journals have been strictly limiting numbers of joint-first and corresponding authorship, as well offering to highlight senior authors with a separate designation on the paper [73]. Contributorship has the potential to solve these problems, which could be a high motivation for funders and researchers alike.

Conclusion

Adding contribution information to research outputs has the potential to inspire innovation to help catalyze improved workflows in scholarly communication. More precise information on a researcher's contributions to research outputs allows the precise, standardized human-readable and machine-operable expressions of researchers' contributions to be better represented, allowing for a more comprehensive and transparent view of what roles and actions power research forward [74], For this to occur, technical and cultural challenges must be addressed to lower the burden on the individual and system level to include this information, provide easy ways to collect and measure this information, and enable downstream opportunities for this information to have a real impact on the academic (and non-academic) reward system. The adoption of contributor roles can make it easier to identify and credit the whole team, catalyzing the necessary cultural shift to evolve scholarship to grow toward open knowledge infrastructures [75].

Acknowledgements

OLD CONTENT:

Making contributorship work in systems

To address the need for a more nuanced characterization and contextualization of contributions, a couple efforts have developed standardized terminologies for contributor roles. The CRediT taxonomy is a high level standardized vocabulary that contains 14 roles for use in representing scholarly contributions to research outputs [1,2]. The Contributor Role Ontology (CRO) was developed as an extension of the CRediT taxonomy, and consumes and expands the contributor roles to provide a structured representation of contribution roles in research and scholarship, which is designed for crediting persons or organizations. The CRO is an open-source, community-developed ontology that currently contains over 50 classes [3]. Contribution Role Ontology was developed by the FORCE11 Attribution Working Group. As noted in the paper by Ilik et al. "this ontology extends the contributions to scholarship beyond manuscript authorship to capture the broadening of researchers' participation in scientific discoveries that have not been previously recognized by traditional measures of scholarly impact" [4]. The work done by the FORCE11 Attribution Working Group along with the OpenVIVO Task Force members at the time included reviewing existing scholarly contribution taxonomies and exploring ways to extend the CRediT taxonomy to create a prototype contributorship model that covers a wide selection of fields of research. The CRO can be used with the Contributor Attribution Model (CAM) - an ontology-based specification for representing information about contributions made to research-related artifacts [5].

Incentives to use contributions

[JC: are these potential incentives, or does it really works this way, if the latter, is there data to back it up, if the former, this should be rephrased]

Benefits to the community could include: better discovery and selection of team members for team-based science, the creation of a directory of experts that can be used for a variety of needs (treating patients, speaking at a conference or to the media, etc.), impartial selection and nomination for awards and committees, more equitable decisions related to promotion and/or tenure, and the ability to illuminate and eliminate gender disparities in research roles (see citation: 6). Benefits to the Individual include: increased recognition and validation of current skills and expertise, motivation to build upon skills or gain new skills, ease of building your professional brand (i.e. describing or explaining expertise or skills to community), and increased ease of finding and engaging your network of colleagues and experts.

Expanding measures of success

It should be noted that improving the characterisation/contextualisation of contributions, will not automatically improve the evaluation processes. The reward system of science has for so long been solely reliant on authorship and the current infrastructure that is used for hiring and promotion is not yet fully prepared to use, and benefit from descriptions of contributions. As long as scientists are being hired and promoted based on the number of publications, impact factor of journals where they published and positions in the byline, even more accurate identifiers of contributions would not improve scientific evaluation and promotion processes. Even researchers based in non-academic institutions report similar patterns in evaluation and promotion [7]. In other words, as long as institutions have not integrated accurate models of contribution such as CRediT and CRO into their workflows, journals' adoption alone is not going to benefit the scientific community.

Manubot Experience and Evaluation

In an ideal world, we would have frictionless workflows that make it easy to introduce credit and attribution to authors and beyond. In addition, workflows would allow seamless ways to promote preservation and discoverability of work. Tools are being developed that help address this issue, such a Manubot [8], a workflow and set of tools that promotes open and automated workflows for writing manuscripts. This commentary was generated using manubot at https://github.com/data2health/contributorship.

A great advantage of using Manubot was that is was designed to track author contributions and made the contributions more transparent. Manubot tracks author metadata in a YAML file, which can be transfered to other manuscripts, saving time when collecting author details each time you write a manuscript. We extended the metadata format to include a list of contritubotor roles for each author. We modified the frontmatter template to note each authors roles along with their identifiers and affiliations, such that readers of the online manuscript have quick access to this information. The discussion and implementation of adding contributor roles are available on GitHub, highlighting another way workflows like Manubot can assist with attribution. When projects are implemented with full transparency in an online venue, there is a public record of who did what and the intellectual contributions sourrouding decisions.

Manubot incorporates a simple workflow for generating citations from persistent identifiers, to automatically create a reference list and remove the need to use a reference manager. By removing dependencies on proprietary software and empowering scholars with a cost-free personal preprint server, Manubot can help improve access to scholarly publishing. On the other hand, tools like Manubot can introduce new barriers that should be considered. Manubot is more technical to use compared to more commonly used authoring tools like Microsoft Word or Google Docs. All authors are required to have a GitHub account and some basic familiarity with using GitHub as well as with writing in Markdown. While Manubot tracked the contributions, the contributions are skewed towards users who are more familiar and comfortable with using GitHub and writing in Markdown.

[**Add something here about using CRO? Manubot should be be able to track our contributions in the yaml file too?**] [**JC: I am not sure that it actually works, github functionality is only counting commits number, not importance, and manubot would only track work done during manuscript writing. To get something automated we would need something much more complicated, where (for example) data would be published with a list of contributor, that list being imported into the manuscript while linking the manuscript to the data.**] [**Can it be imported into author submission systems too, so we don't have to fill out the authorship sections in journals? That would be amazing!**] [**JC: very little portability of the yaml actually, not even to other system using yaml to enter author information. One would need a codec to translate the list from one format to another, see github.com/open-science-promoters/contibutor_manager**] [**putting citation just with doi has also its disadvantage, I do prefer a solution using a bib file (produced by a reference manager) and an author-date call to the bib file, otherwise one always need to click to see what the publication actually is**]

Some features that would be nice to see in Manubot: [NV: not sure if these already exist? `@dhimmel`] - A way to tag other users. In google docs, you can +email someone, and they will get an email that you mentioned them in a comment, and can assign them tasks in the doc - is there spell check integrated somewhere? - can the yaml file with author metadata be converted to a spreadsheet, for easier viewing? Can the authors be reordered/alphabetized?

References

1. **CRediT - Contributor Roles Taxonomy**
   CASRAI
   https://casrai.org/credit/

2. **Contributorship, Not Authorship: Use CRediT to Indicate Who Did What**
   Alex O. Holcombe
   Publications (2019-07-02) https://doi.org/gghnrr
   DOI: 10.3390/publications7030048

3. **Contributor Role Ontology** https://data2health.github.io/contributor-role-ontology/

4. **OpenVIVO: Transparency in Scholarship**
   Violeta Ilik, Michael Conlon, Graham Triggs, Marijane White, Muhammad Javed, Matthew Brush,
   Karen Gutzman, Shahim Essaid, Paul Friedman, Simon Porter, … Kristi L. Holmes
   Frontiers in Research Metrics and Analytics (2018-03-01) https://doi.org/gc5ptw
   DOI: 10.3389/frma.2017.00012

5. **Welcome to the Contributor Attribution Model — Contributor Attribution Model
   documentation** https://contributor-attribution-model.readthedocs.io/en/latest/

6. **Is Science Built on the Shoulders of Women? A Study of Gender Differences in
   Contributorship**
   Benoit Macaluso, Vincent Larivière, Thomas Sugimoto, Cassidy R. Sugimoto
   Academic Medicine (2016-08) https://doi.org/f8z4p5
   DOI: 10.1097/acm.0000000000001261 · PMID: 27276004

7. **Authors' opinions on publication in relation to annual performance assessment**
   Robin L Walker, Lindsay Sykes, Brenda R Hemmelgarn, Hude Quan
   BMC Medical Education (2010-03-09) https://doi.org/c29h2t
   DOI: 10.1186/1472-6920-10-21 · PMID: 20214826 · PMCID: PMC2842280

8. **Open collaborative writing with Manubot**
   Daniel S. Himmelstein, Vincent Rubinetti, David R. Slochower, Dongbo Hu, Venkat S. Malladi, Casey
   S. Greene, Anthony Gitter
   PLOS Computational Biology (2019-06-24) https://doi.org/c7np
   DOI: 10.1371/journal.pcbi.1007128 · PMID: 31233491 · PMCID: PMC6611653