**Credit Where Credit Is Due**

*Promoting open, inclusive, transparent, and traceable science requires that scientists revamp the ways in which we acknowledge all manner of contributions to research.*

Credit is the currency of science. Scientists are evaluated and promoted in their jobs and professional communities on the basis of their *recognized* contributions to science. Unlike a financial contribution, a scientific contribution is difficult to measure. Traditionally, credit for scientific contributions has been given through authorship and citations in scientific literature as well as awards and the naming of geographic features, instruments, and methods and other honorifics. However, these practices do not capture the breadth and depth of the contributions by all actors in modern, open science.

As science becomes more complex, it is increasingly challenging to recognize (and hold accountable) the many people taking part in projects that may involve detailed planning and funding efforts, sophisticated data collection and analysis techniques, custom software development, [integration of data from multiple sources](https://eos.org/opinions/data-sets-are-foundational-to-research-why-dont-we-cite-them), and complex workflows involving machine learning and other sophisticated methods.

Today, there is increasing recognition among many scientists and scientific institutions that fostering science that is more [inclusive](https://eos.org/articles/open-and-fair-data-in-tropical-forests), [transparent](https://eos.org/opinions/the-importance-of-data-set-provenance-for-science), and [reproducible](https://eos.org/opinions/improving-reproducibility-in-earth-science-research) requires that we support wider designation of credit *everywhere* it is due—and to do that, authorship roles and other contributions must be more clearly delineated.

In 2017, AGU adopted the Contributor Roles Taxonomy ([CRediT](https://doi.org/10.3789/ansi.niso.z39.104-2022)) for use in its peer-reviewed journals, in part to increase “transparency around contributions in scholarly research” [[*Hanson and Webb*](https://doi.org/10.1029/2018EO104827), 2018]. This was reinforced in a leading opinion in the *Proceedings of the National Academy of Sciences* [[McNutt et al.](https://doi.org/10.1073/pnas.1715374115), 2018]. Earlier this year, CRediT, which “describes 14 roles that represent the typical range of contributors to scientific scholarly outputs, and that can be used to enable recognition and facilitate transparency” was [published as a national standard](https://www.niso.org/press-releases/contributor-roles-taxonomy-credit-formalized-ansiniso-standard). This standardization is a great advance in recognizing the many roles that factor into the production of publications, but what about authorship and credit for other research artifacts, such as data, code, algorithms, methods, and samples?

The Earth Science Information Partners ([ESIP](https://www.esipfed.org/))—a community of Earth and data scientists focused on the collection, stewardship, and use of Earth science data, information, and knowledge in response to societal needs—created the Research Artifact Citation Cluster to examine all aspects of referencing and crediting the artifacts of research. The cluster examined whether the CRediT taxonomy, and others, could be applied across a broad range of research artifacts. Through guided sessions and structured discussions, we learned valuable lessons about these taxonomies—including that no one taxonomy suffices in all cases—and about how different communities approach crediting and attribution.

**Mapping Out Who Does What**

Citation is a credit mechanism that developed for publications, but various contributions may need to be credited differently for other research artifacts. People often suggest using a model styled after film credits, in which certain key roles—lead actors, directors, and producers, for example—are listed at the beginning of the movie, and numerous supporting roles, from camera operators to caterers, are listed at the end. What may not be evident to viewers not in the movie industry is that the order, prominence, number, and categories of the roles listed are determined in a highly negotiated process involving agents, unions, contracts, and other factors. In much the same way, finding an effective way to designate credit and capture the complexities of different roles in scientific research—and to share this information in a simple way—continues to be a challenge.

Over about 18 months beginning in spring 2020, the ESIP cluster conducted multiple working meetings that brought together dozens of researchers, software developers, and data professionals for discussions of how the myriad roles involved in producing research artifacts could be appropriately credited. We took an exploratory approach in which early meetings refined the questions explored in subsequent meetings.

Initially, participants examined whether the 14 roles included in CRediT (and listed in the acknowledgments at the end of this article) could be applied to appropriately denote credit for other research artifacts, including data, software, learning resources, physical samples, and semantic resources (e.g., scientific categorizations, taxonomies, and ontologies). We also considered the prominence of each role: whether it is akin to authorship of a publication (like top billing in movie credits) or more like a contribution that would be included in formal acknowledgments or in metadata or documentation connected to the artifact (like the fine print at the end of the movie).

From that effort, we determined that defining the research artifact in question is essential for understanding all its component roles and for properly assigning credit, yet even that can be quite difficult. For example, a “model” can be many different things—a conceptual diagram, a climate simulation, a machine learning approach, or something else—each of which can involve different sets of contributors and types of recognition. Even when one can define the artifact clearly, our discussions revealed that CRediT does not apply well for artifacts other than articles. Some roles may be missing from the taxonomy, whereas others may be too vague. For example, CRediT includes a very broad category for software contributions, which is not very helpful when defining distinct roles within software development.

**Different Communities, Different Vocabularies**

We began to explore what roles are missing from CRediT and what other credit taxonomies might be useful. This exploration was partially based on the work of [*Habermann*](https://doi.org/10.5281/zenodo.4767798) [2021], who provides comparisons, or “crosswalks,” of how concepts and contributor roles are described differently depending on context and across several approaches, including CRediT, the [Contributor Role Ontology](https://obofoundry.org/ontology/cro.html), the [Data Documentation Initiative](https://ddialliance.org/), and others. For example, roles like “data quality control” or “data validation” are interpreted very differently with respect to simulation data compared with direct observational data. Similarly, the role of “collector” is critical for physical samples but irrelevant for software.

Although we initially thought we would be able to generalize credit mechanisms for certain types or classes of research artifacts, we found that not only does production of different types of artifacts involve different roles but also different research communities have distinct cultures and approaches for recognizing those roles. For example, the [semantic web](https://www.w3.org/standards/semanticweb/) community, which works to describe web content formally in a machine-interpretable way, has established consistent ways to credit contributions to ontologies and definitions by explicitly labeling terms with [persistent identifiers](https://github.com/ESIPFed/sweet/wiki/Recognition-of-SWEET-Contributors) for the authors and editors who work on them. On the other hand, communities who collect and curate physical samples—such as ice cores or biological specimens—often have distinct approaches to acknowledging credit because from discipline to discipline, these activities can involve very different methods and roles in the field and the lab.

It was becoming clear that no one basic credit approach or taxonomy would suffice for even a plurality of artifacts. So we decided to narrow our focus strictly to data citation to see whether we could identify a relevant taxonomy or guideline describing the primary role of data authorship.

This task also proved difficult. We asked participants to assign weights to the importance of 36 roles listed in the [Contributor Role Ontology](https://obofoundry.org/ontology/cro.html), which expands on the roles included in CRediT, for different types of data. We found that the importance of various roles could vary significantly depending on the type of data. For example, participants weighted study design and protocols very highly for data collected during a field campaign, whereas data integration and quality assurance rose to the top for satellite remote sensing data.

There were, however, some commonalities across data types. For example, in almost all cases, participants agreed about the importance of those who develop the initial idea for, or conceptualize, a data collection effort. This is in keeping with the general ESIP definition of data authorship: “The people or organizations responsible for the intellectual work to develop a data set” [[*ESIP Data Preservation and Stewardship Committee*](https://doi.org/10.6084/m9.figshare.8441816.v1), 2019]. There was also general agreement that some roles, although important, do not rise to the level of data authorship. These include roles like providing funding and designing instruments used in data collection and analysis, as well as the many unseen roles of infrastructure support and maintenance.

**A Case-by-Case Basis**

Our main takeaway from the meetings and discussions was that designating credit and attribution is extremely situational and contextual.

This observation is not entirely new. Who should be considered an author of a scientific article and the order in which authors are listed are issues that have been debated for centuries—and approaches vary immensely across disciplines. However, our work revealed that when we consider the wide range of research activities conducted and artifacts produced today, the complexities multiply. Our work also reinforced that citation is but one of multiple mechanisms that scientists should consider in recognizing the contributions and roles of everyone involved in producing valuable scientific artifacts.

Although we did not find a taxonomy that applies well across scientific disciplines or types of research artifacts for designating credit, several consistent lessons emerged from our work that serve as recommendations. It is very helpful, for example, for a broad group of team members to identify the various roles in a project carefully and deliberately and to assess each role’s significance early in the research process. Teams should think about the following questions: Who is contributing to the project and how? Are those contributions significant enough to warrant authorship? If not, how else could people and organizations be recognized for their work?

Second, taxonomies can provide useful guidance when addressing those questions, potentially helping key participants recognize more fully the breadth and impact of supporting roles, but they are never definitive. The specific context of a project must be considered, including whether roles are being parsed or aggregated to provide a fair representation of contributions. As part of this effort, it is important to decide how roles will be formally recorded, whether, for example, as citations or acknowledgments in a publication or in the documentation of an artifact. Someone who developed data collection protocols for a large, complex field campaign might be considered an “author” of the resulting data sets, whereas in a relatively simple field experiment using established methods (i.e., the method could be cited), that person may only be listed in the acknowledgments. Meanwhile, the people who actually collected the data might be mentioned in acknowledgments, or the names of individuals responsible for particular observations might be embedded in the data themselves.

In any case, in the interest of promoting transparent and reproducible science, it is important to formally document contributions whenever possible [using unique persistent identifiers](https://eos.org/opinions/why-seismic-networks-need-digital-object-identifiers) so that the research community can trace the provenance and impact of contributions. This documentation is also important so that credit for work, still the prevailing currency of science, follows those who performed it.

**Looking Beyond Author Lists**

The culture of how we value different activities and contributions in science—and thus how we evaluate individuals for promotions, awards, funding, and more—is evolving. These evaluations must go beyond just assessments of the articles a researcher has written. Moreover, in evaluations of research artifacts other than articles, the contributions of people beyond just those listed as authors or creators should be considered. Author lists do not tell the whole story of how valuable scientific products came to be.

A [narrative description of contributions](https://www.gla.ac.uk/media/Media_804252_smxx.pdf) could be more useful than a traditional curriculum vitae with a list of publications in facilitating fairer evaluations. Better yet, a network graph can show how a scientist contributed to the production of various data sets and software that, in turn, fed into subsequent articles and other data sets. We believe more holistic and interconnected approaches like these help make science more inclusive and transparent.

We hope the work of the ESIP Research Artifact Citation Cluster to date and the observations presented here prompt further conversation in the scientific community. Structured discussions will continue in both the [cluster](https://wiki.esipfed.org/Research_Artifact_Citation) and a [new AGU-led community of practice](https://data.agu.org/DataCitationCoP/) that’s working with the Research Data Alliance to consider how to assign credit transparently for large, complex data collections. Both groups welcome public participation. Ultimately, outcomes of these ongoing discussions and efforts will help give credit where it’s due and help the practice of science evolve to become more robust, accessible, and trusted.

**Acknowledgments**

This short article resulted from the work of scores, if not hundreds, of people. The ESIP Research Artifact Citation Cluster worked collectively to determine the questions we pursued and the methodology we used. We entrained scores of people to contribute to our workshops both online and in person. The work was truly a community effort.

Ironically, *Eos* guidelines limit us to only five authors in the byline, but more than a dozen people contributed directly to writing this article. We capture that in Figure 1 and the list below based on the Contributor Roles Taxonomy (CRediT). Names under each role are listed alphabetically, with organizational contributors at the end.

**Conceptualization:** Robert R. Downs, Ruth Duerr, Nancy Hoebelheinrich, Daniel S. Katz, Madison Langseth, Mark A. Parsons, Hampapuram Ramapriyan, Sarah Ramdeen, Lesley Wyborn, ESIP Research Artifact Citation Cluster

**Data curation:** not applicable (n/a)

**Formal analysis:** n/a

**Funding acquisition:** n/a

**Investigation:** Robert R. Downs, Ruth Duerr, Nancy Hoebelheinrich, Daniel S. Katz, Madison Langseth, Mark A. Parsons, Hampapuram Ramapriyan, Sarah Ramdeen, Lesley Wyborn, ESIP Research Artifact Citation Cluster

**Methodology:** ESIP Research Artifact Citation Cluster

**Project administration:** Megan Carter, Madison Langseth, Mark A. Parsons, ESIP

**Resources:** ESIP

**Software:** n/a

**Supervision:** n/a

**Validation:** n/a

**Visualization:** Ruth Duerr, Hampapuram Ramapriyan, Sarah Ramdeen

**Writing – original draft:** Mark A. Parsons

**Writing – review & editing:** Ruth Duerr, Rob Casey, Robert R. Downs, Chris Erdmann, Nancy Hoebelheinrich, Daniel S. Katz, Madison Langseth, Matthew Mayernik, Mark A. Parsons, Hampapuram Ramapriyan, Sarah Ramdeen, Lesley Wyborn

We are also grateful for the anonymous reviews from the U.S. Geological Survey (USGS). We are especially grateful for the support, logistics, and collaborative spirit of the ESIP Community with support from NASA, NOAA, and USGS.

**References**

ESIP Data Preservation and Stewardship Committee (2019), Data citation guidelines for Earth science data, version 2, Earth Science Information Partners, Figshare, <https://doi.org/10.6084/m9.figshare.8441816.v1>.

Habermann, T. (2021), Contributor roles crosswalk (version 0), Zenodo, <https://doi.org/10.5281/zenodo.4767798>.

Hanson, B., and S. Webb (2018), Recognizing contributions and giving credit, *Eos*, *99*, <https://doi.org/10.1029/2018EO104827>.

McNutt, M. K., M. Bradford, J. M. Drazen, B. Hanson, B. Howard, K. H. Jamieson, V. Kiermer, E. Marcus, B. K. Pope, R. Schekman, S. Swaminathan, P. J. Stang, and I. M. Verma. (2018), Transparency in authors’ contributions and responsibilities to promote integrity in scientific publication. *Proc Natl Acad Sci U S A* 115 (11): 2557–60. https://doi.org/10.1073/pnas.1715374115.

**Author Information**

Mark A. Parsons (map0046@uah.edu), University of Alabama in Huntsville; Daniel S. Katz, University of Illinois at Urbana-Champaign; Madison Langseth, U.S. Geological Survey, Denver, Colo.; Hampapuram Ramapriyan, Science Systems and Applications, Inc., Lanham, Md.; and Sarah Ramdeen, Lamont-Doherty Earth Observatory, Columbia University, Palisades, N.Y.

**Image Information**

Main image options:

1) <https://www.dreamstime.com/stock-illustration-teamwork-network-three-dimensional-illustration-concept-made-pictogram-people-image59629322>

2) <https://www.shutterstock.com/image-vector/vector-illustration-social-network-scheme-which-738913864>

3) <https://www.shutterstock.com/image-photo/network-interconnected-people-interactions-between-employees-1985087213>

4) <https://www.shutterstock.com/image-vector/people-searching-creative-solutions-teamwork-business-1935960514>

5) <https://www.dreamstime.com/stock-photo-teamwork-business-team-meeting-unity-jigsaw-puzzle-concept-image47350521>

6) <https://www.dreamstime.com/stock-photo-startup-business-people-teamwork-cooperation-hands-together-image92936339>



Figure 1 file name: EOSArticleGraphV7.png

Caption/credit: Fig. 1. Defining contributor roles and attributing credit for various aspects of a project reveal a complex network of interacting participation. Here we see participants and their roles in the project leading to this article. Credit: Ruth Duerr