



## Authorship & Contribution Conflicts

### Background

Data publications in repositories have their own author list, independent of other works. Data publications do not routinely include competing interest declarations and there is not a consistent credit system implementation for contributions. It is expected that issues may arise around authorship and author order.

The recommendations here are intended for cases that arise during or after the publication of datasets. Some possible mitigations that are recommended for data repositories pre-publication are: consider including declarations of competing interests in submission forms, consider implementing authorship taxonomies, capture changes (to metadata) in provenance logs, and consider easy features to gracefully add authors for dynamic datasets.

It is recommended that data repositories send a notification to all co-authors when a dataset has been published, if the repository has access to co-author emails.

### Examples

Authorship and contribution concerns may arise over: author order, author(s) missing from list, data having been posted by someone who did not have the authority to publish the data, data theft. More specifically:

- Authorship is different from manuscript order (e.g. graduate student expects first authorship on dataset) even though data publications and articles do not need to have the same authorship order

- Author omitted from the author list completely (e.g. reader writes in and says they were involved but they have not been included e.g., [in retraction watch](#), or a postdoc leaves lab and is not included on dataset)
- Secondary study based on an existing dataset, where the secondary study involves a separate publication and different authorship to that of the initial dataset
- Deceased author or incarcerated author (out of all potential contact) - remaining an author versus footnote
- Institutional investigation on an author, where the institution rules misconduct
- Dataset is composed in part or full of publicly available data (from the web, other repositories) and the original authors ask for authorship
- Retrospective requests for author name changes e.g., by a transgender individual

## How cases may arise

Concerns could be raised by collaborators, authors, readers, or institutions. This will likely not arise from the journal peer review process but rather post-publication of the dataset.

It is important to bear in mind that a data publication (data published in a repository with a PID) is a stand alone published work and its authorship or author list may be completely different from other related research objects such as a journal publication based on the dataset.

In a first instance, it is recommended to raise any concerns directly with the author and the host of the dataset (e.g. data repository, journal), rather than via public commentary for example on social media or blogging sites.

## Recommendations

In all the scenarios below, the data publisher that first received the concern (e.g., data repository, journal) should take reasonable steps to establish whether another party (e.g., related journal) should be notified and where necessary, communicate to the other party that an issue has arisen. It may not always be possible for a data repository to establish whether associated objects exist for the dataset, and thus, the author is also responsible for notifying the hosts of objects associated with the dataset. Once a resolution is reached the data publisher that first received the concern should notify this to the person raising the concerns.

Given that authorship may differ between a data publication and a journal article, the outcome may involve differences in the author list for the dataset, and for other objects (the exception to this would be for article publications that are focused on dataset descriptions where the authorship should match between dataset and article).

### **What actions should be taken if the dataset has not yet been published? Who needs to be involved in this decision?**

For most instances under this category, the follow up will involve the authors and the repository. The repository should first ask for further detail on circumstances (e.g., if an author is omitted, or if an author is requesting a change in order) from the individual who raised the concern. If the person raising the concerns is not an author, the repository should contact the corresponding author and suggest they contact the co-authors as well as the person who raised the concern to rectify the situation. If an agreement is reached, the repository should modify the authorship metadata.

If a resolution cannot be reached among the authors and the person who raised the issue, and the requested change is more invasive than author order, consider involving the institution of the corresponding or affected authors.

### **What actions should be taken for a published dataset? Who needs to be involved in this decision?**

Post-publication metadata changes are common on data publications. Authorship, affiliation, and author order can typically be changed with a simple metadata update without the need of a notification to readers. Authors are expected to maintain current contact information associated with the data set in order to approve or reject changes.

If the request is for a change in author order, and the data repository has contacted the authors about the request, the data repository may proceed with the change if the corresponding author is on board and a response is not received from co-authors within a suitable time frame.

In situations where the request involves a retrospective change in name for one of the authors, but where no other authorship changes are required, the data repository can implement the requested change without approval by the co-authors. In this scenario it is particularly important that no public notification is posted as this could have ramifications for the author's privacy.

A metadata update is the likely solution for all cases under this category.

### **To whom and when does it need to be reported?**

Data authorship issues would not typically merit reporting to the institution unless there are verified or unresolved concerns about misrepresentation of contributions or data ownership, for example, if a researcher posts as their own output data that were generated by others and for which the posting researcher was not involved.

### **How should the public be notified?**

The public does not need to be notified except for making sure the public versions of record are correct (post-decision), with a metadata update.

**How do we handle inaction or silence from stakeholders (e.g, the publisher, the authors, the institution)?**

If the corresponding author does not respond, or cannot be reached, the repository should go to the institution and ask for assistance in reaching the author for fairness and respect to the researchers involved. If there is no response: the issue should be raised to the submitting author's institution (e.g. the Research Integrity Officer, if there is one), with a notice on the dataset that there is an unresolved issue around contribution.

If the concern is raised with a journal and the journal or publisher takes no action, the matter may be raised to COPE if the journal is a COPE member. It is important to note that COPE advises journals to refer authorship disagreements to the institution to investigate, and thus the journal/publisher would refer the issue to the authors' institution(s) for adjudication.

If there is no response from the institution, the data repository or journal would defer the issue back to the complainant and note that they need to raise it with the relevant institution as the issue is beyond the purview of the repository or journal to adjudicate.