Pubwication of software papers, and authorship on them

Pubwication. Pubwication is what bwings us togethew today. Pubwication, that bwessed awwangement, that dweam within a dweam. And authorship, twue authorship, wiww fowwow you fowevah and evah. So tweasuwe youw authorship.

Tue 29 September 2015
By *C. Titus Brown*
In science.
tags: publication software citation sustainability lior yet again

Last week, our software paper on khmer 2.0 was published on F1000Research. We intend this paper to be a citation marker, but it also represents and recognizes some significant software engineering work done between khmer 1.x and khmer 2.0.

As part of the paper process, we offered authorship to everyone who has contributed to the khmer git repository - anyone who contributed to the repo was invited to sign on to the paper.

Addendum: I would like to credit Michael Crusoe with the initial suggestion to offer authorship to all git committers. This is in no way backing away from my own support for this decision, but I only realized a few days after writing the post that I had failed to properly credit Michael. So, kudos, Michael!

This policy has caused some consternation amongst the Twitterati, some (well, ok, one) of whom recoiled in horror at our author list, pointing at the recommendations of (e.g.) the International Committee of Medical Journal Editors. These recommendations are that authorship be based on (1) design and/or analysis, AND (2) writing, AND (3) final approval, AND (4) accountability. While the third and fourth criteria were met by all of the authors of this paper, the first and second were probably not met by all authors.

A few points are in order:

- our condition for authorship is explicit, verifiable, and transparent. You can look up contributors in our release notes and GitHub repo to find out exactly what they did.

  We note that this clarity and verifiability is in contrast to most authorship.

- authorship has been getting more complicated, and traditional authorship roles are both ill defined and clearly inadequate for modern research. There is an ongoing effort to define authorship roles more clearly and explicitly, and, coincidentally, GigaScience just announced they're signing on to this.

  For those who are curious, the majority of authors on our F1000Research paper fall under #3 on the CRediT taxonomy.

  (We did not talk with F1000Research about their support for this taxonomy in advance.)

- our project is an open source project that is developed by a community, with contribution requirements and an

extensive process for contributing. Our author list is an explicit acknowledgement of the role that the community has played in developing khmer, and the work that each and every contributor invested in our process.

- assuming this paper passes into the peer reviewed literature, there may be some interesting consequences of our authorship criteria. For example, since most formal definitions of Conflict of Interest include shared authorship, Jared Simpson and Lex Nederbragt and I would now be in conflict and I would be unable to review their grants or papers. This seems silly to me!

Speaking with my senior author hat on, I hope it's clear that we are not trying to mock authorship in any way, and this is a serious publication on a serious project.

That having been said, we are trying out something new - in particular, we would like to figure out how to acknowledge software authorship within the scientific literature, both because it's the right thing to do AND because we'd like to incentivize community development of software. This is part of an ongoing discussion about the changing roles of contributorship in research (see first point, above).

On names

Some note has been made of the presence of what people presume to be pseudonyms in our author list. There is a long history in science of choosing a specific name or pseudonym to publish under; see Student, for one example (ht @rgcjk) of very, very many. We support this tradition.

Respecting peoples' chosen names is also important for many other reasons. I suggest people read through the Nymwars Wikipedia page, and pay special attention to the "criticism" section, which raises ethical, moral, and legal reasons why a "real names" policy is problematic.

Please note that there is a special place in hell reserved for people who attempt to deanonymize someone's pseudonym on a whim; this is both unprofessional and potentially harmful to the individual in question. Yes, I'm talking to you, Lior.

--titus

p.s. Please comment responsibly! On this post and all future posts, I am going to follow the Captain Awkward comments policy - specifically, "...sometimes comments don't show up because I delete them. This is a dictatorship, and I can delete any comment at any time for any reason."

Comments!

(Please check out the comments policy before commenting.)

Isn't part of why this way of 'granting' authorship is causing so much debate caused by the fact we put so incredibly much value on authorship in the first place? It currently is (one of) the biggest currencies in science: who is on what paper (in what journal with - oh my - what impact factor). If we could let go of all that, and find other ways to determine who is a good scientists (whatever that means), the issues some people have with the author list of this paper would probably disappear...

7 ▲ | ▼ • Reply • Share ›

David Dotson ➚ Lex Nederbragt • 3 years ago

I find the hand wringing over authorship and the overfocus on publications (perhaps at the cost of the basic quality of the science the publications are only description of) tiresome. Perhaps that's why I'd rather work on software, anyway; at least with version control systems credit is easier to decipher.

Most of the time, however, I don't think the question of "Who did X" really matters; what matters is that the overall product (say, the software package) actually does something useful.

4 ▲ | ▼ • Reply • Share ›

Nicolas Bray ➚ Lex Nederbragt • 3 years ago

I know they would for me, Lex, but that's if (or rather when), such new systems are created. For now this seems like just taking the biggest currency in science and completely devaluing it.

Well, that's not exactly true, I guess. The first and last authors get even more credit for heading such a large project, it's just the middle authors whose contribution gets basically nullified.

▲ | ▼ • Reply • Share ›

Richard Smith-Unna ➚ Nicolas Bray • 3 years ago

Nicolas, I'm struggling to see this as anything other than snide and disingenuous. If the system is broken, pressure to change it is valuable. Waiting until a perfect alternative is in place is both unrealistic, and a way to avoid having to do anything.

The first and last authors in this project really do not need the Science Points that might come with authorship - their careers are built not on the silly metrics, but the community appreciation of their (and their work's) value. What is achieved here is giving recognition to people who, in most projects, are left out of the current system.

I for one welcome the devaluation of authorship as a credit system. It's silly, it hurts us (by which I mean early career researchers), and we can do better. Exploring ways to do better *with it* is a crucial part of figuring out how we can do better (though I really believe we'd do better without it altogether).

▲ | ▼ • Reply • Share ›

Titus Brown Mod ➚ Richard Smith-Unna • 3 years ago

Richard,

I thought Nicolas made an interesting point, albeit one I didn't agree with. My point would be that this is really an experiment, and, alone or even in small aggregate, nothing we do is going to significantly affect the system. But we can see how it shakes out, hopefully.

▲ | ▼ • Reply • Share ›

mem_somerville • 3 years ago

I do a lot of software testing that never gets credited in papers. I mean, I don't really need the publications for career issues. But stuff I find leads to code changes and improvements in the final product. I wish that was valued a bit more.

7 • Reply • Share ›

Nicolas Bray mem_somerville • 3 years ago

If I were on a project with someone who did extensive testing of software leading to improvements, I'd definitely consider that a significant contribution and think that deserved an authorship.

I'd also be upset if that significant contribution were equated with something as insignificant as correcting a typo.

• Reply • Share ›

Titus Brown Mod mem_somerville • 3 years ago

Thanks, @mem_somerville.

• Reply • Share ›

ZomBeaver • 3 years ago

Titus, I appreciate you doing this. I remember as a young scientist long ago, I contributed to much of the software that led to the analysis of a Nature Genetics paper, and I wasn't even in the acknowledgements. I remember the first author coming to me every time she needed to do anything computational (permutation tests, etc). Thus, I find value in what you are doing. It was never going to be easy including all those authors. But the heart of what you're doing is important.

4 • Reply • Share ›

Titus Brown Mod ZomBeaver • 3 years ago

Thanks, ZomBeaver!

• Reply • Share ›

Claus Wilke • 3 years ago

I think the recommendations of the International Committee of Medical Journal Editors have to be taken with a grain of salt. For example, if I followed them strictly to the letter I don't think I could ever justify undergraduate authors on any paper. These guidelines were written to prevent senior scientists from just putting their names on papers without having made a meaningful intellectual contribution. The guidelines don't always work for more junior scientists. They also don't work for very large collaborations, e.g. in high-energy physics with thousands of authors on one paper.

Ultimately, each community has to find its own rules that it thinks make sense. I'm fine with a software paper where every person who has made any contribution to the code is a co-author, even if that contribution was just fixing a typo or cleaning up indentation.

4 • Reply • Share ›

Michael Hoffman Claus Wilke • 3 years ago

+1

• Reply • Share ›

Titus Brown Mod → Claus Wilke • 3 years ago

Thanks, Claus!

• Reply • Share ›

Istvan Albert • 3 years ago

It surprises me that you and all the co-authors fail to see the moral hazard and ethics of it.

In a typical publication process authorship is earned. in the majority of cases its quite clear cut who deserves to be on it and who does not.

In this case the majority of the authors have been "gifted" the authorship by you. They are now in your debt because you decided to include them. They would have no case to make to be included otherwise. This leads to most authors actually having a really a strong conflict of interest as now they all "owe you one".

I see this as social experiment. Normally one would refuse to be on a paper where they only had some tiny contributions. But then when it seems that everyone else does the same, all of a sudden it is ok. Why not me? Look "everyone else does it".

It is like that Oprah show: you win a car, and you and you and you. But we all know everyone won't win a car. Only those win that are in the right place at the right time. The same is with this paper, all authors know (or should know) that it would not be possible to scale this up so that all science looks like this: every minor thing would net you an authorship.

But since this is the first paper it might get through and they get a "free" publication out of it hence they all sign up for it.

4 • Reply • Share ›

Andreas Leimbach → Istvan Albert • 3 years ago

I do appreciate Titus for doing this and especially making it public for an apparently much needed discussion.
I think there're quite some cases where we're kidding ourselves with the traditional ways of authorship and its order. Some commenters have already touched on this, but I'll give a couple more examples (somewhat exaggerated):

- the technicians that never get an authorship in some labs
- the student that couldn't finish the big paper and doesn't get credit because he left the lab
- the new student in the lab that "just" finishes up the last experiment for the big paper and gets a first authorship
- the student that gets bumped from first authorship to make room for a postdoc, collaborator etc.
- the lab that includes authorships for sharing a strain, plasmid etc., while others do not
- the pet bioinformatician that doesn't get credit because he doesn't do the "important" wet-lab experiments
- the PI that is on every paper without "significant" contribution (maybe to push the publication with a well-known name)
- reciprocal authorships

see more

1 • Reply • Share ›

Titus Brown Mod → Andreas Leimbach • 3 years ago

Thanks, Andreas -- this is a great list!

Two quick responses -

• Reply • Share ›

Titus Brown Mod → Titus Brown • 3 years ago

Whoops, truncated...

And frankly, I don't think there are a lot of students out there who will confront senior researchers about authorship disputes.

I challenged my graduate advisor on this once: I told him I disagreed with author placement (and gave him reasons why), but also told him that I accepted that someone needed to make the decision and would abide by his. He seemed a bit taken aback ;).

At last, the authorship debate somewhat reminds me of peer review. A lot of people pretend it's working, although there're many examples where it does not.

+1 excellent point!

But, in response - do we *need* something that is cut and dry and works 100% of the time? What's the problem with some ambiguity and complexity? I haven't thought too much about this, but my immediate response would be that this is tied to evaluation more than anything else: we want to know how much to "respect" someone's contribution to a paper, for whatever reason.

• Reply • Share ›

Titus Brown Mod → Istvan Albert • 3 years ago

Thanks, Istvan, for your perspective. I don't really know how to address the moral hazard issue except to say that I think most people have more interesting things to do than to contribute to khmer on the hope of getting an authorship on the next software paper. Also note that it has not yet passed peer review, so no moral hazard has yet entered into the equation. Maybe it'll pass peer review and then we'll get a flood of contributors! But, you know, this is potentially falsifiable...

1 • Reply • Share ›

Andreas Leimbach • 3 years ago

Because this post has attracted such an interesting discussion, I posted a question also on the Open Science Q&A website.
"What are the current practices of granting authorships/deciding authorship order and how can it be improved?"
https://openscience.uni-bie...

This is not meant to "steal" the discussion, but maybe attract an audience from a more diverse background. You might like to follow up there. Thanks.

2 • Reply • Share ›

Titus Brown Mod → Andreas Leimbach • 3 years ago

Thanks, Andreas!

• Reply • Share ›

en zyme • 3 years ago

one broken or missing line of code can easily down a plane, or stop a heart, empty your bank account, or have no effect whatsoever. (caveat: your experience may vary)

the Heartbleed security bug, resulted from a missing bounds check, essentially a one liner, "Cyberoam Security Advisory - Heartbleed Vulnerability in OpenSSL". April 11, 2014.

for some tales from the bitbucket check out :
http://www.computerworld.co...

2 • Reply • Share ›

Pat Schloss • 3 years ago

I like this a lot, Titus. One problem that we face in developing open source software is how to incentivize people to help out - offering authorship is one tool we have. I think what you've done is go a long ways towards recognizing this problem and providing a solution. Many of the early commenters seem to not appreciate this difficulty. Back in 2009, I sent out an email to my mailing list of early mothur adopters telling them that if they provided an example on our wiki of how they've used mothur, I'd make them a co-author. Many of these people I have never met IRL. The hope was that this would spurn people on to contribute to the wiki and the overall project. So what have been the rewards?

Benefit to me: For whatever reason, getting people to help with documentation hasn't worked out (people still email me with typos), but I can at least say I tried to engage the community. So, that's a bonus. It also helped launch a forum that I use to show people how to use the software. Periodically, I do get people contributing an analysis example, which is pretty sweet.

Benefit to co-authors: Those early adopters now have a paper on their CV that has more than 3000 citations. I see this as a major boon to their professional development. They have to live with the responsibility of telling people what they actually did to contribute to mothur.

Going forward. I doubt there will be a second incarnation of the paper (I'm too greedy with the citation count). But I do appreciate that there's a big problem in getting people to contribute to open source projects. One thing I can do is offer to write letters of recommendation to junior scientists that I notice carry some weight on our user forum and in making contributions to the wiki.

2 • Reply • Share ›

Titus Brown Mod → Pat Schloss • 3 years ago

Excellent perspective, Pat - and that 2009 anecdote is great :)

• Reply • Share ›

Carl Friedrich Bolz • 3 years ago

I love this policy, and have done similar things in the past: we had a one week coding sprint that eventually resulted in a paper, and everybody who was at the sprint became an author:
http://link.springer.com/ch...

Judging from that and other similar experience I'm actually not that worried about the 'typo fixing author problem'. Authors that didn't really contribute much tend to just decline authorship anyway.

The slight worry I have is more about really huge OS projects that would lead to pages of authors, but I guess I'll cross that bridge when I get to it (high energy physics seems to have no problem with pages of authors either, after all).

2 • Reply • Share ›

Titus Brown Mod → Carl Friedrich Bolz • 3 years ago

Thanks, CFB!

• Reply • Share ›

Michael R. Crusoe → Carl Friedrich Bolz • 3 years ago

We did have some authors decline when I sent out a follow up survey: https://docs.google.com/for...

For example, I sent this reply to one contributor and now co-author who had initially declined:

"Hello [redacted],

We have several papers, both published and in preparation, that describe novel features within khmer. This paper is for the software project as a whole and we made the decision early on to offer co-authorship to every contributor so we wouldn't have to adjudicate on what makes a 'significant' contribution. We feel that this is fair and we explain this in the Author Contributions (draft) section:

'CTB is the primary investigator for the khmer software package. MRC is the lead software developer from July 2013 onwards. Many significant components of khmer have their own paper describing them CITE CITE CITE. The remaining authors each have one or more Git commits in their name.'

I am happy to leave you on the co-author list and fine with taking you off. The choice is yours.

You will be getting another email when we post the pre-print and when the paper is published."

• Reply • Share ›

Tal Yarkoni → Michael R. Crusoe • 3 years ago

The fact that most people will probably do the right thing in this type of situation and decline authorship is great, but it isn't really a valid reason to leave the choice up to people who have a large incentive to say "Yes Please". It's a policy that would spell disaster if universalized, as it would become very hard to determine whether someone made a meaningful contribution or not without spending a lot of extra time and effort. (I mean, it's really no different from the senior author who puts everyone in his lab on the paper because they answered a question about a reagent somebody used for one of the experiments.)

• Reply • Share ›

Michael Barton • 3 years ago

I agree with your point about deanonymisation, Titus. I feel like there are better ways to have a

discussion about authorship, while still disagreeing, that add to a constructive debate.

My feeling is that authorship serves the people who measure academic performance such as funding bodies and review panels. It makes it easier to create an academic score such as number of first and last author papers. Modern collaborative science has people performing different roles that often cannot be easily represented by authorship position. For example soft skills are often undervalued, such as project coordination, yet are often critical to researchers working in teams. Another way to describe it would be that measuring scientific contribution is complex and has multiple orthogonal axes, yet we are forced to reduce it down to three essential categories for publication: first, last and in the middle.

I think that authorship position leads to divisive discussions based on contributions that often lead to feelings of resentment. This is overdue for a more fairer and inclusive approach that represent how different people contribute in different ways.

2 ^ | v • Reply • Share ›

Titus Brown Mod ➚ Michael Barton • 3 years ago

thanks, Michael!

^ | v • Reply • Share ›

Andreas Leimbach • 3 years ago

Hi Titus,

I think this is great and how it should be done. Although probably not many life scientists outside the twitter-/bioinformatics-sphere will agree, I'm afraid.

Life scientists still can learn a lot from physicists. I recently talked to a theoretical physicist, he said they publish exclusively on arXiv and the author order is always just alphabetically (as you did besides first and last).

One thing to your earlier comment, isn't the F1000 publication also a research paper?

2 ^ | v • Reply • Share ›

Titus Brown Mod ➚ Andreas Leimbach • 3 years ago

Thanks, Andreas - we regard this as a software paper, not a research paper; the article discusses research that is either already done or about to be submitted, and doesn't make any arguments beyond referencing things that have been done. Not sure how clear that distinction will be over time but we will see!

1 ^ | v • Reply • Share ›

KDM • 3 years ago

Disclaimer: I'm a contributor and co-author.

One thing I haven't seen mentioned, largely due to the demographics involved in the usual twitter shitstorm, is what more junior co-authors think.

My opinion is that everyone has contributed roughly according to their abilities or academic "rank" (blerg, someone suggest a better term please). A few non-academic or Honours/Masters/Early PhD students committed one or two typo/minor bug fix changes, and got authorship. Which is *awesome* as they really need any paper they can get right now. Other PhD students like myself committed several to hundreds of commits, and got credited in the same manner, which is *awesome* as showing that we

can collaborate (i.e. middle authorship) is important career-wise, and I'll get first author papers using khmer (in my case, the contributions to khmer were entirely such that I could use khmer in my papers). Mr-c and Titus were (at least in my experience) the gatekeepers to any contributions of any middle author and deserve their first/final authorship.

/me braces for chagrin, but that's my 0.000002BTC

Cheers,
K

1 • Reply • Share ›

Titus Brown Mod → KDM • 3 years ago

Thanks, K. I really like the point that you are contributing to khmer for your own "selfish" purposes and expect to get your own research papers out of it (and I do not expect to be a co-author on those papers :). This is also true of most of the major contributors, note.

• Reply • Share ›

Bob Carpenter • 3 years ago

We've tried to be very open about authorship with our multi-authored software package, Stan (general-purpose Bayesian inference). But it's a tricky business when people's careers hang in the balance. Especially because not everyone contributes equally. Of course, that's one of the benefits of having different people --- they contribute different things, ranging from key algorithms, low-level library code, text editing modes, theory, methodology, examples, and interfaces to external languages. All extremely valuable contributions, no matter the scale. Especially that testing (props to **@mem_somerville**); I was at Julia Con (for the language) and one presenter said she wrote unit tests when she got bored. Biggest ovation of the three day event, spontaenously, right in the middle of the talk; now that's a crowd of serious developers.

For better or worse, contributing to software packages isn't valued very highly by most academic departments, so there isn't much "value" to divvy up.

I was once on a paper (http://www.genomebiology.co... that involved 30+ authors, each of whom had contributed a system for a bakeoff evaluation on gene mention detection in biomedical research papers run by the first two authors at NIH. Mainly it just caused headaches formatting my conflict of interest statements in grants downstream.

As Harry Truman said, "It is amazing what you can accomplish if you do not care who gets the credit." I know the quote because it's the .sig of a friend of mine who's a very successful (and well loved) junior high principal. Like many other successful leaders, he operates on the principle that if something goes well, there's plenty of credit to go around. As Truman's contemporaries working on game theory might have put it, credit is not a zero-sum game!

1 • Reply • Share ›

Titus Brown Mod → Bob Carpenter • 3 years ago

Thanks, Bob - great perspective!

• Reply • Share ›

Tal Yarkoni • 3 years ago

I agree with Istvan Albert's comment and some of the other commenters. I think the policy of rewarding

contributors to the repository with authorship is admirable and should be the norm for software papers, and in that respect, I'm glad to see you pushing the envelope (as usual!). But I don't think the desire to give authorship to contributors absolves you of having to decide what counts as a meaningful contribution and what doesn't--and it's clear that some of the contributions can't possibly qualify on even the most liberal definition of "meaningful intellectual contribution".

Your point about clarity and transparency is well taken, except that (a) nobody really has time to go look up who contributed what in a long list of authors; (b) the cultural norm is that authorship reflects a meaningful contribution, so most people will quite reasonably assume authors on the paper didn't just correct a typo; and (c) the fact that there's a place somewhere that lists all contributions explicitly surely doesn't justify granting authorship to just about anyone. I mean, if my friend babysits my kids to buy me some time to work on my paper, I don't think it should be okay for me to list him in the author list just so long as I explicitly detail his contribution in a publicly accessible README file that's linked from the manuscript. That devalues the contributions of everyone else on the paper and, if widely adopted, would set up some very perverse incentives (e.g., everyone who reads this should immediately open a PR on all your repos adding a comment or fixing a typo).

1 • Reply • Share ›

John Longinotto ➚ Tal Yarkoni • 3 years ago

If the most perverse outcome you could think of is people come together to improve scientific tools, then I'd say that a pretty clean indication that acknowledgment of contribution - no matter how small - is worth it.

• Reply • Share ›

Tal Yarkoni ➚ John Longinotto • 3 years ago

That's just the short-term effect. The long-term effect is what we're already seeing in many fields of biomedical science where papers commonly have hundreds of authors: it becomes much more difficult to assign credit appropriately. Giving co-authorship to everyone who contributed a single line of code ultimately devalues the contribution of the people who did the actual work, because you don't have any way to tell who really played a role in shaping the project, and who just happened to catch a typo. In the long term, it actually achieves exactly the opposite of what the stated goal is, in that we end up in a kind of arms race where in order to appear productive, one has to be listed as an author on an ever-increasing number of articles with ever-growing author lists. Alternatively, past a certain point of authorship fatigue, readers just start completely ignoring everything but the first, second, and last authors, in which case the gesture becomes essentially meaningless anyway.

To be clear, I'm 100% in favor of giving developers credit for the work they've done. I'm entirely on board with making the N most important contributors to a repository authors (where N is some arbitrary and necessarily subjective number). But if the authorship credit doesn't denote a meaningful intellectual effort, then it actually becomes deleterious to scientific evaluation.

1 • Reply • Share ›

Titus Brown Mod ➚ Tal Yarkoni • 3 years ago

Hi Tal,

I'm somewhat sympathetic to the concern, but I don't see it happening with khmer

so far - the paper's been out for over a month, and there's been no flood of new contributors. At what point would you acknowledge that this is an unfounded concern?

Moreover, I think experimentation and trying new things out is important. I'm puzzled by the view that this stuff is too important to experiment on. Doesn't it make it more important to do experiments?? If we had received a flood of minor contributors out to get coauthorship on the next paper, then that would have been interesting to see. But we haven't, in the past month, and I don't think we will. And that's pretty informative - it suggests that your mental model for how stuff works is incorrect.

--titus

1 ^ | v • Reply • Share ›

Tal Yarkoni ➚ Titus Brown • 3 years ago

Oh, Twitter snark aside (sorry!), I didn't seriously expect you to get a flood of new contributors--as you can probably tell from the fact that I didn't drop everything myself to go edit your README. ;)

I think the short-term "perverse incentive" part of this is less of a concern than the long-term impact on publishing norms in science. Again, I think there's a direct parallel with the trend in many other fields of science. There's nothing intrinsically wrong with long author lists, and in some cases they're unavoidable (i.e., some large-scale projects really do depend on substantial intellectual contributions from hundreds of people). The concern to my mind is the devaluation of authorship. I didn't seriously think that hundreds of people would flock to contribute to your repo any more than I think people would flock to the lab of a PI who announced that everyone who walks through the door gets authorship on every paper. The issue is what effect that kind of culture has on science in the long run.

So, to answer your question, I'm not against experimentation at all. My concern is that the approach you've adopted to authorship isn't scalable: if everyone did the same thing, it would become very difficult to determine who contributed what, which would actually end up hurting the people who played the biggest roles on any given project.

That said, my concern is largely conditioned on the assumption that the current model of scientific evaluation (based largely on publication lists and CVs) sticks around for the long term. And I would much prefer that it didn't. Ideally, I think we should be evaluating scientists more like developers. :)

^ | v • Reply • Share ›

John Longinotto ➚ Tal Yarkoni • 3 years ago

I think our disagreement exists because we are looking at the purpose of authorship from two different angles. I agree with you that we need a metric from which to evaluate an individual's contribution to science, I just don't think that the author list should be the place to do it. I believe that if we go down that path, we end up in strange places like "I'm sorry person X, you may have contributed to the project in a meaningful way, but if I simply acknowledge your existence it will

project in a meaningful way, but if I simply acknowledge your existence it will *devalue the work of all other scientists*."

And I think it may come down to scarcity vs abundance thinking; I don't believe that there is a cap on the amount of good science that can be done worldwide in a single year, or from a single journal, or in a single paper - so why do we feel guilty about having too many people listed as authors on a paper? If, in the future, huge teams of people all working together via GitHub and are all getting authorship - who really cares? If the bottom line is better science, i'm all for it! If it's roughly the same level of science but people's contributions are being more fairly represented, i'm still all for it :)

If another group of people who are using authorship as a proxy for an individual's contribution/value no longer have a metric that makes sense... well, sucks to be those guys. They should get a better metric.

Reply • Share ›

Tal Yarkoni ➚ John Longinotto • 3 years ago

I have mixed feelings about this. On the one hand, I agree with you that in an ideal world (and hopefully in this one too, in a few years), we would use metrics other than authorship to evaluate contributions. And I'm certainly not going to argue that we should preserve the sanctity of traditional pre-publication peer reviewed scientific publications (I've argued to the contrary on many occasions).

On the other hand, much of our ability to evaluate one another's work and give credit appropriately depends on people abiding by certain standards of behavior--and one of the most basic is that you don't take credit for work you didn't do. In science, the cultural expectation is that authorship implies a meaningful intellectual contribution to a paper. This isn't just a suggestion; it's an explicit part of the author guidelines at most journals. If I list 20 papers on my CV, and my contribution to 15 of them consists primarily of correcting typos, I'm effectively misleading other people, and I'm also potentially screwing over someone else who has 8 papers on their CV, but made major contributions to all of them.

Now, I agree with you that if we could find metrics that aren't susceptible to these kinds of issues, we'd be in a much better place. And I'm all for evaluating scientists more like developers. But the reality is that much of science currently isn't a matter of machine-readable record, and won't be for the foreseeable future. So while I'm sympathetic to the argument that a metric that can be easily gamed is a shitty metric, I'm not sure that's a justification for trampling all over prevailing codified norms until we have a better approach.

1 Reply • Share ›

John Longinotto ➚ Tal Yarkoni • 3 years ago

I agree with everything :) I just wish we had a better metric.

Reply • Share ›

Keith Bradnam • 3 years ago

The author contributions section makes it clear that you are including everyone who made at least one Git commit. I'm not sure how someone would feel if they missed the cut by one day though, i.e. if

Git commit. I'm not sure how someone would feel if they missed the cut by one day though, i.e. if someone added one line of code the day after this manuscript is published they would get no recognition as 'author' even thought they have contributed as much (or more so) than other 'authors' on the paper (of course they would still be tracked through GitHub).

The other issue here is that all coding contributors who are on the paper gain equal billing whereas it looks like maybe 4–6 people (apart from you and Michael) made significant, and sustained contributions to the code base.

My suggestion would therefore be to divide people into two sections in the 'Author contributions' section: those that made 'substantive' contributions to the code base (add initials for their names), and everyone else.

1 ^ | v • Reply • Share ›

Titus Brown Mod ➚ Keith Bradnam • 3 years ago

Hi Keith, thanks for the comment!

For your first comment, we did authorship through the release of khmer 2.0. People who contribute after that will probably have billing on the next release. Whether or not this is fair, at least it's cut and dry ;)

People who have made significant and sustained contributions to the code base will be on research papers, and/or received salary in order to do so, and/or got other benefits. I'm not too worried about it and we don't want to impose an arbitrary cutoff here, either.

^ | v • Reply • Share ›

Keith Bradnam ➚ Titus Brown • 3 years ago

I appreciate than any cutoff would be arbitrary, but no cut-off raises the problem that Lior mentioned and which Mick alludes to above.

As for the salary and benefits of people who made substantive coding contributions, this information may be known to you but it is not transparent to people looking at the paper.

In any case, I am glad you are doing this. It is good to experiment with models of authorship and publication.

^ | v • Reply • Share ›

Titus Brown Mod ➚ Keith Bradnam • 3 years ago

Keith - is no cut-off better than a subjective and arbitrary one? I think it's worth a shot.

I don't think salary/benefits are generally known wrt papers...?

And thanks, I agree that we need some experimentation :)

^ | v • Reply • Share ›

Titus Brown Mod • 2 years ago

Here's a nice link about anonymity and pseudonymity in online spaces that I'm posting for later reference: https://blog.coralproject.n...

^ | v • Reply • Share ›

Titus Brown Mod • 3 years ago

Nicely relevant link: http://people.uleth.ca/~dan...

• Reply • Share ›

en zyme • 3 years ago

fascinating thread. i'll reserve more detailed comments for F1000

• Reply • Share ›

Load more comments

ALSO ON **LIVING IN AN IVORY BASEMENT - BLOG**

Can bits be the basis for a digital commons? (No.)

1 comment • 5 months ago

Titus Brown — Hi Roman, I think data meets the definition of non-rivalrous - there is no prevention of consumption or utility to others of the data …

Four steps in five minutes to deploy a Carpentry lesson for a class of 30

5 comments • a year ago

Marianne Corvellec — Wonderful! Thanks for sharing, Titus.I'm curious, did you get some feedback from the students specifically on this …

Classifying genome bins using a custom reference database - maybe this time it'll …

1 comment • a year ago

Titus Brown — I just noticed how confusing the command is -- here, the 'delmont/*.sig' refers to the signatures of the genomes you want to …

A quick read of _The genomic and proteomic landscape of the rumen microbiome_

7 comments • a month ago

Titus Brown — https://twitter.com/JeffsMi...

Proudly powered by pelican, which uses python.

The theme is subtlely modified from one by Smashing Magazine, thar

For more about this blog's author, see the main site or the lab

While the author is employed by the University of California, Davis, his opinions are his own and almost certainly bear resemblance to what UC Davis's official opinion would be, had they a