Conferences, Journals, Preprints, and Reviewer Expectations

Alistair Moffat
The University of Melbourne
ammoffat@unimelb.edu.au

Joel Mackenzie
The University of Melbourne
joel.mackenzie@unimelb.edu.au

Abstract

We consider the precedence and priority issues that arise from the increasingly common trend of distributing unrefereed preprints via services such as arXiv prior to or at the same time as they are submitted for peer review, and the effect that this practice can have on the integrity of the scientific review process. We offer some suggestions which could be incorporated into the peer-review process to provide better safeguards to authors.

1 Introduction

The submission-review-revision-acceptance-publication cycle is a critical component of science. Once a paper has been accepted, we have a degree of confidence that its claims have been scrutinized by one or more experts, lifting our trust in its conclusions. There is no certainty, of course; and the review process is not an infallible guarantee. Nor does "not yet reviewed" or "under review" mean that work is not yet valid. It is a continuum, built on the understanding that the larger the number of independent people that have thought critically about any given set of claims, the greater the confidence we can have that those claims are sound. At the beginning of that continuum, the "peer reviewed" label gives us an assurance that at least someone other than the authors has undertaken an evaluation of the claims made in the paper.

Against that backdrop we have the increasing trend for authors to release preprints of their work, sometimes as soon as they have written their manuscript and submitted it for review. Issuing a preprint via a service such as arXiv asserts a priority claim on the ideas, and allows other researchers to have early access to the results. But preview lodgment can also affect the review process. Our purpose in this note is to describe – via the story in Section 2 – one way in which there can be interaction between preprint availability and peer review that has the net effect of valid work becoming difficult to publish. Section 3 then discusses some possible remedies.

2 Beaten to the Punch

Consider the case of two groups of authors, Group A and Group B. At roughly the same time they independently commence work on some topic X, and develop software and experimental

Group *A*

submit to journal

lodge on arXiv

working on problem *X*   writing paper   under review

working on problem *X*   writing paper   under review

Group *B*

submit to conference

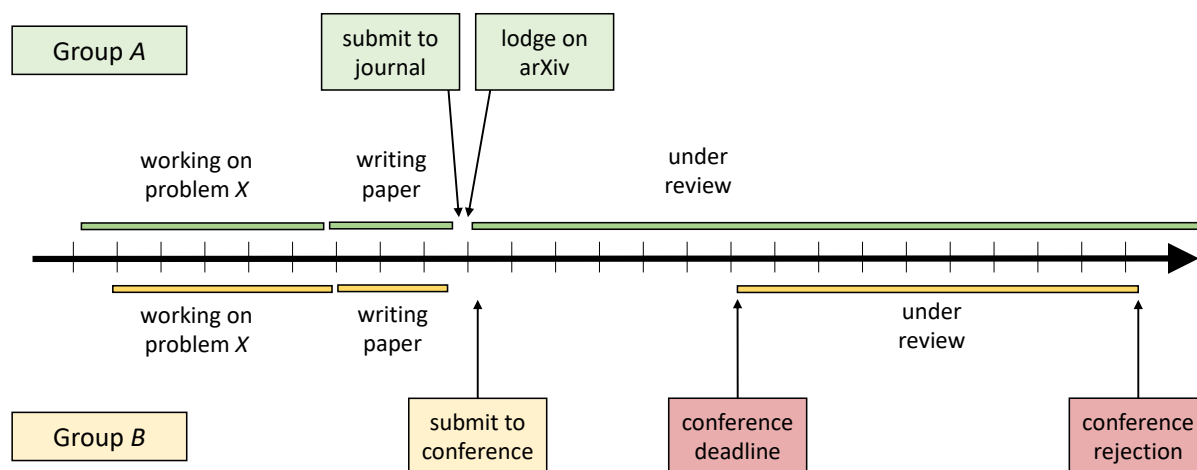conference deadline

conference rejection

Figure 1: The time line considered in Section 2.

protocols that allow them to investigate X. As chance would have it, they (still independently) write up their work through the same time periods, and concurrently consider their publication options. Figure 1 summarizes the time line we are considering.

Group A decides to submit to a journal, and prepares a manuscript accordingly. Because journal review in the field of Information Retrieval is typically single blind (where referees are anonymous to authors, but authors are known to referees), Group A also decide to post their draft paper to arXiv as an unpublished preprint, so that other researchers can see what they have been working on (an act of generosity) and so that their priority on the ideas is clearly established (an act of protection). Making their manuscript and findings public does not compromise the integrity of the review process they seek, as their identities are known to their referees anyway.

Group B, on the other hand, decide to pursue conference publication. They are ahead of schedule in terms of dates, and the next submission deadline they are interested in is still nearly two months away. Nevertheless, they finalize their manuscript, and submit it to the conference just as soon as submissions become available, a week after they had finished their work, six weeks before the submission deadline (yes, it *can* happen!), and just a few days after Group A register their arXiv preprint. Group B read the conference's submission guidelines very carefully, because to not do so is folly, and risks desk rejection. They find that those guidelines confirm that review will be double-blind – that is, reviewers will not know the identity of authors, and vice versa, the normal arrangement for IR conferences. Those guidelines usually include a statement like this one, taken from SIGIR 2021:[1]

> Authors are required to take all reasonable steps to preserve the anonymity of their submission. [. . .] While authors can upload to institutional or other preprint repositories such as arXiv.org before reviewing is complete, we generally discourage this since it places anonymity at risk (which could result in a negative outcome [. . .])

or like this one, taken from ECIR 2022:[2]

> Submission is permitted for papers that have previously been made available as a technical report ([...] like arXiv). However, we discourage this since it places anonymity at risk; in particular, please do not publish your paper at arXiv and submit to ECIR at the same time, some days before, or during the reviewing period of ECIR.

That is, Group B must be very careful to have no content in their paper that might point to their identity, and nor do they register their work at arXiv.

Group B are aware that their work is now entering peer-review "limbo" for a period, but they have confidence in what they achieved, and have their fingers crossed that the assigned referees will also find their results interesting. They are also aware that the conference submission system itself is a kind of "registration" system, should it ever become necessary to demonstrate the provenance of their submission, and so, like Group A, they believe that their work has a degree of protection.

When Group B receive their referee reports, another nine weeks after the submission deadline, and more than three months after they finished their investigation and wrote the first draft of their paper, they find that their work has been rejected. One of the referees compared their approach to the Group A arXiv report (which has still not emerged from the peer-review process, one of the perils of making journal submissions), and voiced an opinion that the Group B work is neither timely nor original. That negative view sways the other two referees and the meta-reviewer, who had been inclined to accept the paper based on their own reading of it. During the discussion phase that knowledgeable referee perhaps pointed at the long interval between the arXiv registration date and the conference submission *deadline* to help explain their recommendation – that deadline being, of course, the *last* moment at which Group B might have completed their paper.

An implicit rule of conference reviewing processes is that no matter what happens, authors are expected to accept the outcome ("suck it up"), even if they think that mistakes have been made. There is no recourse or independent appeals board, and in the IR area at least, decisions are regarded as being immutable and final, no matter how egregious they may appear to be.

And hence, Group B must re-think their plans for publication. For example, they could make a fresh submission to a different venue, and re-start the review process. They may get a completely new set of referees, or might end up with a set that overlaps and hence carries a degree of corporate memory. In this second submission they must note the existence of Group A's work, and either explain how their project adds value to the results that Group A has already made visible, or seek to claim that their work was independent of Group A. Either way, they need to cross their fingers and trust that the new referees will understand the subtlety of what is being argued, and will heed the implicit request that the paper not be assumed to be developed and written as a consequence of Group B having read Group A's preprint. That places Group B in limbo a second time, probably with decreased odds of success compared to their first submission.

Moreover, if Group B do plan to argue that their work was undertaken concurrently via the conference submission time-stamp, they must reveal that they were rejected from the conference, which establishes a prior that may influence the new review process. On the other hand, if they re-submit to another venue (journal or conference) without that disclosure, their risk of rejection rises with time, with each added week of delay increasing the likelihood of their work being perceived as simply extending Group A's arXiv preprint. Eventually, of course, Group A's paper will emerge

---

[2]https://ecir2022.org/calls/full-papers/, accessed 9 December 2021.

from the review process (as a result of the initial submission, or as a result of further submissions) and reach "to appear" status. Once that happens Group B has very few options still possible.

Our question is this: Was Group B treated fairly, and if not, where did the system let them down? Or, were they obligated to take note of the Group A arXiv work and respond to it in some way between when they finalized their submission to the conference, and the actual final deadline for conference submissions?

3   Some Suggestions

The reader will have inferred from our language that we believe that there is room for improvement in the processes that led to the situation Group B found itself in. We now assemble some thoughts in regard to those issues.

Citations to Preprints   In early 2018, 174 members of the SIGIR community participated in the *SIGIR Community Survey on Preprint Services* (Kelly [2018]). The majority of respondents indicated that they "seldom" or "never" cite preprints. Some respondents shared concerns that preprints are not trustworthy since they have not been peer-reviewed. Interestingly, while some respondents said that their review scores would be negatively affected if a paper they were reviewing did not compare with the empirical results in a very relevant preprint, the majority stated that they would make a comment, but that their scores would not be affected, or that they wouldn't penalize the paper whatsoever. While this survey influenced some change in the way preprints are handled by conferences and journals, there are still no clear guidelines in the IR area on how referees should proceed in situations such as the one illustrated in Figure 1.

Timeliness and Concurrency   If Group B were responsible for monitoring arXiv and all other forms of early dissemination, at what point does that responsibility end? In the tale above, Group B finished their work nearly two months before the conference deadline. They believed they were done, and moved on to their next project. On the other hand, the referee believed that they should have been aware of work that, by the time of the submission deadline, had been public for six weeks (see Figure 1). Would the referee have lodged the same opinion if Group A's arXiv paper had been registered six days before the deadline? Or six hours? Clearly there needs to be line drawn in the sand in terms of what things can be expected to be sequential, and what things can be assumed to be concurrent.

Refereeing Guidelines   Cognate fields, such as computational linguistics, already address this issue. For example, the *ACL Policies for Submission, Review and Citation*[3] state:

> Papers submitted to ACL conferences and TACL should in principle cite and compare to all relevant prior work, regardless of when and how that work was presented to the community, and must credit work that influenced them. To this general rule there are two important qualifications:

---

[3] https://www.aclweb.org/adminwiki/index.php?title=ACL_Policies_for_Submission,_Review_and_Citation&oldid=74128, accessed 30 November 2021.

- For citation, refereed publications take priority over preprints. Specifically:
    - Authors are expected to cite all refereed publications relevant to their submission, but may be excused for not knowing about all unpublished work (especially work that has been recently posted and/or is not widely cited).
    - If a preprint has been superseded by a refereed publication, the refereed publication should be cited in addition to or instead of the preprint version.

- For comparison, papers (whether refereed or not) appearing less than 3 months before the submission deadline should be considered contemporaneous to the submission. This relieves authors from the obligation to make detailed comparisons that require additional experimentation and/or in-depth analysis, but they are still expected to cite and discuss contemporaneous work to the degree feasible.

It follows from our recommendations about citation and comparison that failure to cite relevant prior work or failure to compare to (non-contemporaneous) empirical results may affect the assessment of a submission regardless of how the prior work was published. However, reviewers should be instructed to give authors the benefit of the doubt in cases where the work appears in preprints with no corresponding refereed publication, especially preprints that are recent and/or not widely cited. In such cases, reviewers should point authors to the non-cited work (so that they can discuss it in the camera-ready version) but not penalize the authors for missing the citation.

Under these guidelines the referee would not (or, at least, should not) have argued for rejection of Group B's work on timeliness and originality grounds. The three-month definition of "contemporaneous" would have made a critical difference.

Referee Oversight　It is not wrong for referees with strong opinions to seek to sway the overall decision – often they have noted some aspect of the paper that the other members of the review team have missed. But where a single referee argues for rejection on procedural grounds (in the case here, based on publication precedence and prior statement of ideas), the other reviewers have a duty to test that recommendation, and ensure that it stands up. That is, the referee teams assigned to each paper should be willing to directly question each other in regard to what gets written in the collected reviews, and the conclusions that get drawn; and ensure that any strong statements being made are factually correct. This is a second way in which referee guidelines might be drafted to provide more explicit advice.

Appropriate Use of Dates　There is also the question of access to chronologies. Should the referee (and other reviewers of the paper) have been more careful, and checked the actual submission date rather than assuming that the work was completed immediately prior to the submission deadline? For example, in the case of EasyChair, referees and meta-reviewers can see the time-stamp attached to each submission. And if dates are not visible, one imagines that the PC Chairs could be asked to check the time line, even if via an email to the authors.

The scenario in Figure 1 might have played out differently had the referee (or any other member of the review team assigned to the paper) checked a time-stamp.

Disclosure Statements If Group B had become aware of the Group A preprint prior to the final submission date, they might have decided to squeeze a one-line footnote into their paper and then immediately re-upload it, declaring that their work had been developed independently of Group A. Would a declaration to that effect have altered the way the paper was judged by the conference referees? Perhaps; but we have never seen such a note on a paper, and so cannot judge how it might be interpreted. And it might also have gone the other way too: a referee might have even more determinedly taken an "it's a tough world, you lost the race" attitude.

Adding commentary to the referee guidelines that anticipates such disclosures and normalizes them is another possible adjustment that might be worth considering.

Rebuttal, or Appeals Processes As already noted, conference review in IR is (by and large) a binary outcome: acceptance or rejection, elation or dejection. There is sometimes a rebuttal period embedded in the review process, but never an appeals process. Yet appeals can provide critical oversight of what are sometimes quite inscrutable outcomes. Appeals should normally be restricted to procedural and process errors, and not be permitted as a way of arguing difference of opinion over relative importance or significance; but even with that restriction, can sometimes be a way to resolve mistakes. For example, the Australian Research Council recently reversed a total of 32 decisions in regard to project eligibility. In the ARC's original decision, they argued that grant proposals should *not* be citing preprints, with the 32 projects deemed ineligible for funding as a result. That decision was overturned on appeal (and after a wave of academic protest), with all 32 projects deemed eligible for funding, and six of them then awarded funding as a result of their position in the ranking developed during the review process.[4]

It might be appropriate for the SIGIR Executive Committee to establish a "Review Appeals Committee" to consider cases (in SIGIR sponsored conferences) where authors have evidence to suggest that the stated process was not followed. Appeals would need to be resolved quickly, and the committee would require non-trivial involvement from experienced academics, but nevertheless might be an investment worth considering; review forms would need to partition the issues of academic merit and procedural compliance into separate questions.

Carry-Forward Reviewing The conference paper review system starts each time with zero official knowledge, with each paper considered to be a completely fresh submission; and with rejected papers "logjamming" their way downstream through the flow of conference deadlines. So rejected SIGIR papers get submitted next to CIKM and are fully re-reviewed; and perhaps rejected CIKM papers are re-submitted to ECIR and fully re-reviewed. In practice, there is often overlap of assigned referees, and the downstream CIKM or ECIR referee might also have reviewed the same work for SIGIR a few months earlier. If that happens, a typical referee would re-read their previous comments, and then look to see whether the authors have responded to the feedback that was provided. If the paper is unchanged, it likely leads directly to a "reject" recommendation from that referee, based on "I already told them what was needed, and they chose not to do it". On the other hand, there is no option in current IR conferences for authors to indicate that a paper has already been reviewed, and for them to provide the previous reviews – and a response/rebuttal to them – as part of their downstream "roll the dice again" submission.

---

[4]See <https://www.arc.gov.au/news-publications/media/communiques> and the item "Outcomes of ARC Appeals committee on preprints", posted on 2 December 2021, accessed 10 December 2021.

Suppose now that every review bundle issued from EasyChair had an `md5` checksum associated with it, and EasyChair maintained a on-going service whereby supply of a valid `md5` checksum allowed that set of prior reviews to be viewed, plus validation details, such as paper title, conference name, date submitted, and so on; but still completely anonymized in terms of both paper authorship and reviewer details. Introduction of this service would then allow a CIKM submission to be accompanied by an `md5` checksum that linked to and hence made visible the SIGIR reviews for the paper. Suppose further that in cases where authors opted to supply that checksum, they were also permitted to upload a two-page response document, giving them the opportunity to address those previous referee comments, and to explain what changes had been made to the paper since that earlier rejection – or perhaps (more courageously) to try and explain why the first set of referees had misunderstood the paper.

Such a system might be employed by the Group B authors in their second submission, giving them a chance to rebut the criticisms of the referee, and to make a clear statement of independent originality, without the reviewing "cold start" risk that was noted above. Such a system might also help reduce the refereeing load that we are all feeling. A paper that had already been reviewed by three SIGIR referees plus an SPC member might be allocated to just one or two CIKM referees plus an SPC member if it arrives via this "resubmission including reviews" pathway. It might also reduce the volume of work associated with an Appeals Committee.

Note that a number of conferences now require mandatory declaration of resubmissions, and that any resubmission be augmented by the previous reviews, plus further explanatory text that reports on how the reviews have been integrated. For example, IJCAI requires disclosure of any rejections of the work in the previous six months.[5] Our suggestion here is for an opt-in process, supported by access to the previous review feedback package in a verifiable manner.

Double-Blind Review for Journals   Another possible change could be to switch IR journals to double-blind review, and ask *all* authors to not expose their work via preprints. It it a little puzzling that authors must bend over backwards to hide their identity in one type of submission, but can put their names on the front page of the other type. For example, Tomkins et al. [2017] carried out an experiment in which half of the WSDM 2017 PC members were aware of author names and the other half were not; they found that knowledge of author identity affected both the bidding processes and also the recommendations entered by reviewers: "single-blind reviewers are significantly more likely than their double-blind counterparts to recommend for acceptance papers from famous authors, top universities, and top companies" [Tomkins et al., 2017].

If the main pool of IR journals could be persuaded to shift to double-blind review, the issues raised in our Group A/Group B tale might not have arisen, albeit at the price of neither group being able to provide an early dissemination of their findings.

Other experiments have explored the consistency of review outcomes. In 2014, Corinna Cortes and Neil Lawrence managed the NIPS paper review process in a manner that put 10% of the papers into both sides of two otherwise disjoint review processes.[6] They found a relatively high level of inconsistency between the two review pools, confirming that conference paper review protocols are rather imprecise in their outcomes. In a similar study, Hannah Bast set up two completely

---

[5]See https://ijcai-21.org/cfp/, accessed 6 December 2021.
[6]See https://inverseprobability.com/2014/12/16/the-nips-experiment and http://blog.mrtz.org/2014/12/15/the-nips-experiment.html, both accessed 5 December 2021.

parallel evaluation committees for papers submitted to Track B of the 2018 European Symposium of Algorithms (ESA 2018), and also found that repeatability of the paper selection process was relatively low.[7] These studies provide further impetus to establishing a facility for "carry-forward reviews", discussed above.

4   Conclusion

We recognize the importance of open research, and have no issues with the use of preprint services to provide early dissemination – and, as we have noted, early protection – of work. We also recognize that the peer-review process, while very important as a "checking" process that helps validate academic research and is a critical part of scientific dissemination, is far from perfect; it is inevitable that it sometimes accepts work that is later discredited, and sometimes rejects work that later becomes seminal. And between those two extremes, valid and interesting work sometimes gets accepted and sometimes (when the acceptance rate is low, or when a referee just misunderstands it) gets rejected.

Nevertheless, peer review is the "democracy"[8] of academic life, and as such, is worth protecting with precise guidance to referees, plus the provision of carefully chosen checks and balances to ensure that work that does not receive the protection afforded by early preprint dissemination is treated equitably.

References

D. Kelly. SIGIR community survey on preprint services. *SIGIR Forum*, 52(1):11–33, 2018. doi: 10.1145/3274784.3274787.

A. Tomkins, M. Zhang, and W. D. Heavlin. Reviewer bias in single- versus double-blind peer review. *Proc. Natl. Acad. Sci. USA*, 114(48):12708–12713, 2017. doi: 10.1073/pnas.1707323114.

---

[7]See https://cacm.acm.org/blogs/blog-cacm/248824-how-objective-is-peer-review/fulltext,   accessed 5 December 2021.

[8]"Many forms of Government have been tried, and will be tried in this world of sin and woe. No one pretends that democracy is perfect or all-wise. Indeed it has been said that democracy is the worst form of Government except for all those other forms that have been tried from time to time...", Winston Churchill, 1947, see https://winstonchurchill.org/resources/quotes/the-worst-form-of-government/, accessed 4 December 2021.