The version of record as a central organizing concept in scholarly publishing

Lisa Janicke Hinchliffe*
University Library, University of Illinois at Urbana-Champaign, 1408 West Gregory Drive, Urbana, IL, USA

Abstract. The "version of record" is an organizing concept in scholarly publishing. It is by reference to that version that others are understood and it is the object of financial models, policies, and recognition and reward systems. At the same time, many of the core functions of academic publishing – in particular, registration and dissemination – are decoupling from the version of record. Scholarly publishers are also expanding their remit to encompass other article versions, as well as other research outputs, and efforts to systematically link together and track these into a "record of versions" are growing.

Keywords: Scholarly publishing, journal articles, version of record

1. Introduction

The "version of record" is an organizing concept in scholarly publishing. It is by reference to that version that others are understood and it is the object of financial models, policies, and recognition and reward systems. At the same time, many of the core functions of academic publishing – in particular, registration and dissemination – are decoupling from the version of record. Scholarly publishers are also expanding their remit to encompass other article versions, as well as other research outputs, and efforts to systematically link together and track these into a "record of versions" are growing. This article presents a landscape scan of the state of the "version of record".

2. Version of record defined

The version of record is defined both formally by industry organizations as well as colloquially. In everyday conversation with scholars, librarians, and publishers, we often hear the version of record identified as "the publisher PDF" or even just "the PDF". While a researcher may be willing to accept a different version, such as a preprint, when they ask for the PDF, no librarian asks which version they mean. We immediately look to locate the publisher PDF. We can see this conceptualization operating nascently as well when someone tweets about an article and then tags it "[paywalled]" even though there are other versions of the article that are open.

Crossref's document on "Version Control, Corrections, and Retractions" defines the version of record as the "typeset, copyedited, and published version" and observes that version control is important for traceability, identifiability, clarity, reduced duplication, and reduced errors [1]. While there are many potential versions prior to version of record, there is only one version that follows sequentially.

---

*E-mail: ljanicke@illinois.edu.

NISO's definition of the version of record further underscores the role of the publisher in the creation of the version of record. The NISO Recommended Practice on Journal Article Versions states that the version of record is "a fixed version of a journal article that has been made available by any organization that acts as a publisher by formally and exclusively declaring the article 'published [2]'". Most notably, then, publishers have the agency in identifying the version of record. We know which is the version of record through a kind of "speech act" by the publisher – by declaring it the version of record, it becomes the version of record.

We should also take a moment to address some potential confusions about the version of record. In the early days of electronic journals, when there were discrepancies between the printed version and the electronic version, the print version was typically considered the version of record in various style guides and the like. Today, the printed version of a journal issue – when one even exists at all – is no longer given this primacy.

Additionally, the version of record is also not always the most recent version of an article. NISO identifies two follow-on possibilities (corrected and enhanced). Crossref presents a single category for all follow-on possibilities (updated), including retraction.

Finally, in the case of publicly available or open access articles, there is some additional nuance. For open access, the version of record may be gold and/or green – as the "color" of open access is not determined by version but by whether the copy is served by the publisher platform or elsewhere.

3. The centrality of the version of record

What we can see then is what I would term the teleological and central nature of the version of record in organizing the scholarly record. As the version of record comes into being, it has the effect of fixing the earlier versions as predecessors to it. In addition, anything that comes after the version of record is defined relative to it. This centrality is reflected throughout the scholarly communications system in a number of ways.

4. Editorially privileged

The version of record is editorially privileged in many citation and style guides. For example, the APA Style Guide to Electronic References stated that authors should "Update your references frequently prior to publication of your work; refer to the final published version of sources when possible [3]". More recently, the APA Style Blog directed that an author should "Ideally, use and cite the final, published version of a work. However, if you used the preprint version of a work, cite that version [4]".

In addition, the instructions to authors for manuscript preparation may include admonitions to not cite other versions instead of the version of record. For example, even in physics, a discipline that led on the development of preprints, the American Physical Society states in its Information for Authors that "citations to e-print archives should not be used in place of primary references [5]".

5. Prioritized by funders

Funders have focused on the version of record in their open access policies. Even when an author accepted manuscript is accepted for compliance instead of the version of record, the requirement to provide open copy is tied to the existence of the version of record and, when embargoes are allowed, to the timing of the publication of the version of record.

For example, the NSF Public Access Policy "requires that either the version of record or the final accepted manuscript in peer-reviewed scholarly journals and papers in juried conference proceedings or transactions (also known as "juried conference papers") be deposited in a public access compliant repository designated by NSF; be available for download, reading and analysis free of charge no later than 12 months after initial publication [6]".

Some funders go further and express a preference for the open copy to be the version of record. For example, cOAlition S states that: "we are mindful that the AAM [author accepted manuscript] version differs from the VoR [version of record]. Not only does the latter contain all the changes from the copyediting process, journal formatting/branding etc., but it is also the version maintained and curated by the publisher, who has the responsibility to ensure that any corrections or retractions are applied in a timely and consistent way. For this reason, our preferred option is to ensure that the VoR is made open access [7]".

In addition to open access compliance, the version of record is also the locus of compliance for mandates related to ethics declarations, data availability, and funder recognition as well as documentation of contributor roles. Prior to the version of record, these elements are often either fluid and/or not required.

6. Focus of recognition and reward structures

As a faculty member myself, I am very aware of the focus on publications in promotion and tenure processes. For example, at my institution, our departmental annual review directions state "do not include works submitted but not yet accepted [8]" and the dossier outline for campus says to list "in print or accepted" works [9]. Other research institutions have more expansive parameters for listing preprints or in progress works; however, I know of none that would judge positively a tenure case that included only such works. And, when I am asked by other universities to write external evaluator letters for candidates being considered for tenure or promotion, the focus requested by the university is on published work. In my more than two decades of writing such letters, sometimes as many as seven in a given year, I have never been sent a preprint to evaluate.

It is also notable that the research assessment reform effort DORA, while recognizing the expansion of research outputs, organizes its efforts around the version of record. Specifically, the "San Francisco Declaration on Research Assessment" states that "Outputs other than research articles will grow in importance in assessing research effectiveness in the future, but the peer-reviewed research paper will remain a central research output that informs research assessment. Our recommendations therefore focus primarily on practices relating to research articles published in peer-reviewed journals but can and should be extended by recognizing additional products, such as datasets, as important research outputs [10]".

And, of course, institutional rankings and related processes are also heavily dependent on publications and citations to those publications. For example, the "*THE World University Rankings*" take significant account of research productivity and citations as assessed through Elsevier's indexing and bibliometric tools, which take as their object a corpus of published articles, books, and conference papers.

7. Researcher perspectives

Unfortunately, there is relatively little systematic investigation into researcher perspectives on the version of record; nonetheless, the two significant reports we do have find a similar pattern – researchers value the version of record but also find value in other versions for particular use cases. There are also instances where they are unwilling, or at least quite hesitant, to substitute other versions. This may be,

at least in part, due to the editorial, funder, etc. privileging discussed above; researcher preferences and actions are shaped by environmental factors not only personal beliefs.

"Exploring Researcher Preference for the Version of Record", a white paper from Springer Nature by Mithu Lucraft, Katie Allin, and Imogen Batt, presents a study conducted in the context of the Springer Nature syndication partnership with ResearchGate [12]. The study reported that researchers prefer to read and cite the article version of record, which they believe is easier to read and more reliable as well as the most authoritative and credible source. Specifically, the survey found that 83% of respondents preferred working with the version of record for citing content in their own work. In contrast, 9% preferred author manuscripts and 2% preferred preprints. Open text comments evidence that researchers value the "stamp of credibility" that they believe is signaled by the version of record. Nonetheless, researchers do see value in other versions, which are particularly valuable for skimming, reading, and keeping up on the literature. Speed and ease of access are seen as a significant benefit of versions that are publicly available, particularly of preprints.

The study also reported that researchers are more likely to look for ways to find the published article than an author manuscript or preprint, including by attempting to contact the author directly. I have noted elsewhere that the open data set that accompanies the report also includes multiple mentions of Sci-Hub as a mechanism for getting copies of published articles, which underscores the drive to access the published copy, though this was not analyzed in the formal report [13].

"What's the Big Deal?" – an inquiry into the emergence, growth, and decline of library subscriptions to large scale journal collections by Danielle Cooper and Oya Y. Rieger of Ithaka S+R – uncovered some important details of researcher perspectives on the version of record and other versions through examination of researchers' experiences in the context of significant journal subscription cancellations [14].

Most significantly, respondents were clear that they prefer not to use preprints for most teaching purposes and discourage students from using preprints in their papers. These instructors comment on the difficulty students have with respect to understanding the context of a preprint. Specifically, the report shares that researchers find that "preprints add an extra level of complexity to the already challenging process of introducing students to the scholarly communications ecosystem. With the need to assign a limited number of articles that correspond to specific pedagogical aims, instructors therefore generally prefer to opt for those that they perceive as the easiest to navigate as a novice reader, which are peer reviewed and published in respectable journals". The report continues on, "Given that students are at an earlier point of experience with evaluating scholarly literature, the version of record was seen by our interviewees as the easier format for them to navigate".

Researchers themselves also sometimes struggled to conceptualize the role of preprints, though this appeared at least somewhat influenced by disciplinary context. Cooper and Rieger observe that "Some scholars who did not come from fields with an established preprint culture also expressed confusion as to how to engage with preprints meaningfully … In fields where preprint practices are less established, interviewees frequently questioned their quality and expressed their ongoing commitment to the peer review process. There was also confusion among this group as to how a preprint relates to or can be linked to the version of record".

In fields where preprints are more firmly established, researchers seem to have developed more nuanced perspectives, reflecting different use contexts even within the single use case of citation. This quote from a researcher included in the Ithaka report is illustrative: "I use preprints ... primarily to stay up to date with the latest and greatest in my field. And that's primarily directed towards when I'm working on my grants. Less so when I'm writing papers".

Additional studies on this topic of researcher perceptions and practices are greatly needed. These two offer just a glimpse into scholars' mindsets and approaches, which one suspects are evolving relatively rapidly, at least in some fields. Understanding how scholars conceptualize the role of the version of record and its relative value would be of great use to all players in the scholarly communications industry.

8. Financial sustainability

This review of the state of the version of record would be greatly remiss if it overlooked the centrality of the version of record to the financial sustainability of the scholarly publishing industry. Subscriptions are payments for access to the version of record. Subscribe-to-open is a payment for assured access to the version of record. APCs and other open access funding frameworks (e.g., transformative agreements such as read-and-publish or pure publish agreements) operate by paying for the production of the version of record.

Of course, this financial model is threatened as components of the value proposition of the version of record decouple from it. Preprints are carving out the value of the role of the version of record in registration; public posting of preprints and author accepted manuscripts are carving out the value of the role of the version of record in dissemination. Certification and preservation remain tied relatively strongly with the version of record; however, in some fields being posted on a specific preprint platform shows glimmers of threatening the role of certification. Perhaps these value challenges are why we see an increasing turn to payment-to-publish rather than payment-to-read revenue models as well as publishers pivoting to workflow support and hosting preprints, data, etc. directly. Publishers are innovating in order to attempt to shore up their centrality.

9. The future of the version of record

The version of record's role as an organizing principle in scholarly communications is so deeply inherent in the system that dislodging it from that centrality would be a significant challenge and very destabilizing to the system as a whole. One can imagine a future in which research projects generate a document (or a set of documents) that are continuously revised and updated [15], perhaps over the course of multiple researchers' careers, with evolving authorship and contribution markers. In such a future, perhaps there is with no identified exclusive version of record, only a set of the most recent versions, and perhaps they are selected into overlay journals [16], with a given text possibly selected into more than one journal [17]. It seems more likely that the version of record will be the locus of future innovations as the fixed nature of a version of record has proven utility – beyond its financial value – as an infrastructure of knowledge building, information, and recognition and reward systems.

Note: This article is an edited version of "The State of the Version of Record" (https://scholarlykitchen.sspnet.org/2022/02/14/the-state-of-the-version-of-record/), which based on my plenary talk at the APE 2022 Conference: The Future of the Permanent Record (https://www.ape2022.eu/full-program/program).

References

[1] Crossref. Version control, corrections, and retractions, 2020. [cited 2022 Feb 14]. Available from: https://www.crossref.org/documentation/crossmark/version-control-corrections-and-retractions.

[2] NISO. Recommended practice on journal article versions, 2008. [cited 2022 Feb 14]. Available from: https://www.niso.org/publications/niso-rp-8-2008-jav.

[3] American Psychological Association. *APA Style Guide to Electronic References*, 6th edn, American Psychological Association, Washington, DC, 2012, p. 46.

[4] American Psychological Association. Preprint article references, 2020. [cited 2022 Feb 14]. Available from: https://apastyle.apa.org/style-grammar-guidelines/references/examples/preprint-article-references.

[5] American Physical Society. Physical Review A: Information for authors. [cited 2022 Feb 14]. Available from: https://journals.aps.org/pra/authors.

[6] National Science Foundation. What is NSF's public access policy? [cited 2022 Feb 14]. Available from: https://www.nsf.gov/pubs/2016/nsf16009/nsf16009.jsp#q1.

[7] cOAlition S. Plan S rights retention strategy. [cited 2022 Feb 14]. Available from: https://www.coalition-s.org/rights-retention-strategy/.

[8] University of Illinois Library at Urbana-Champaign. 2021 annual review report, guidelines and outline, 2021. Available online: https://www.library.illinois.edu/staff/wp-content/uploads/sites/24/2017/11/2021_AnnualReviewReport_Guidelines_Outline.docx.

[9] Office of the Provost, University of Illinois at Urbana-Champaign. Communication #9: Promotion and tenure, 2022. Available online: https://provost.illinois.edu/policies/provosts-communications/communication-9-promotion-and-tenure/.

[10] San Francisco Declaration on Research Assessment, 2013. [cited 2022 Feb 14]. Available from: https://sfdora.org/read/.

[11] Times Higher Education. THE world university rankings 2021: Methodology, 2020. [cited 2022 Feb 14]. Available from: https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2021-methodology.

[12] M. Lucraft, K. Allin and I. Batt, Exploring researcher preference for the version of record, 2021. [cited 2022 Feb 14]. Available from: https://www.springernature.com/gp/open-research/version-of-record.

[13] L.J. Hinchliffe, Publishers care about the version of record, do researchers? 2021. [cited 2022 Feb 14]. Available from: https://scholarlykitchen.sspnet.org/2021/04/05/publishers-care-about-the-version-of-record-do-researchers/.

[14] D. Cooper and O.Y. Rieger, What's the big deal?: how researchers are navigating changes to journal access, 2021. Available from: 10.18665/sr.315570.

[15] D.R. Shanahan, A living document: Reincarnating the research article, *Trials* **16** (2015), 151. doi:10.1186/s13063-015-0666-5.

[16] T. Alves, Overlay journals, overlay reviews: Has their time finally come? *Science Editor* **44** (2021). doi:10.36591/SE-D-4402-55.

[17] T. Vines, Two new initiatives at eLife to start the Eisen era, 2019. [cited 2022 Feb 14]. Available from: https://scholarlykitchen.sspnet.org/2019/08/15/two-new-initiatives-at-elife-to-start-the-eisen-era/.