

Open Research

EXPLORING RESEARCHER PREFERENCE FOR THE VERSION OF RECORD

White paper

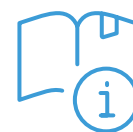
ADVANCING
DISCOVERY



Open Research

Journals	Data
Books	Tools

Contents



Foreword	1
Executive Summary.....	3
Main findings.	3
Introduction	5
Usage analysis	6
Reader and author attitudes towards article formats	7
Findings	8
Part one: usage of article versions	8
Part two: user preference (reader survey)	10
Discussion and Conclusions	16
Appendix	17
References	18
Acknowledgements	21

Authors

Mithu Lucraft; Imogen Batt; Katie Allin

February 2021

This white paper and associated raw data from the survey have been made openly available in the figshare repository.

■ **Access whitepaper:**
<https://www.doi.org/10.6084/m9.figshare.13834532>

■ **Access survey data:**
<https://doi.org/10.6084/m9.figshare.13834712>

Foreword

Springer Nature is committed to achieving the benefits of open science for all. We are constantly looking at ways to support and enable a fully open research future, where all outputs of scholarly research – from the research article to its underlying data and code – are interconnected and accessible to all. Open science increases the speed, integrity, reproducibility and overall efficiency of the whole research system, in turn helping to address global societal challenges, such as those identified by the UN's Sustainable Development Goals. In the past year we have seen the real time value of open science in advancing our understanding and treatment of COVID-19, with the development of vaccines in record time.

It is why we have been very clear in our support for full and immediate open access (OA), as a fundamental building block of open science. As our Chief Executive Frank Vrancken Peeters explains, we are adamant that for it to be 'full' OA, it must be 'gold' not 'green'.¹ This is because only gold OA gives immediate access to the final published version of record (VOR), is sustainable and is dynamically updated in perpetuity. In addition, gold OA articles act as an integrated hub for all the elements necessary for open science such as data and code.

This is why we feel that gold OA is the only way to achieve real and meaningful open science and it is why we are concerned that Plan S' support for green OA via their Rights Retention Strategy is misplaced. First, it removes the ability of authors to control how their work is used by prescribing the type of open licence they must seek to apply to their accepted manuscript (AM). Second, it provides access to an unfinished AM, which requires the continuation of subscription payments to fund the creation of these articles, preventing the gold OA transition that will enable the benefits of open science.

Those are our concerns, but what about the views of those who author and use OA papers? The voices of researchers and authors have not yet been heard. How valuable is the VOR to them? Would such increased access to the AM meet their needs? This is what we have sought to explore in this white paper.

Our findings are that researchers, both as readers and as authors, are most comfortable with the VOR on every count. They prefer to use the VOR themselves, they prefer to cite the VOR, and they prefer others to use the VOR when it comes to their own work.



Steven Inchcoombe, Chief
Publishing and Solutions Officer,
Springer Nature

**Gold OA is the only
way to achieve real and
meaningful open science.**

1. Vrancken Peeters (2021)

So if gold OA is what researchers say that they want then could, as Plan S funders say, green OA be looked upon as a transitory model to get us there? We are cautious of this without a clear plan of how we manage this period of transition and call on funders to work with us to start to address this now. Our concern is that without such action, we could unwittingly find ourselves in a situation where green OA is deemed by some to be “good enough” and we fail to move quickly enough to gold. Providing only access to the unfinished AM, which does not have the benefits of post-accept improvements, is not linked up with data or code, does not show corrections, or retractions, and ultimately relies on the continuation of library subscriptions, risks embedding an inferior version of OA, with the result being we fall short in delivering on the promise of an open science future that is so crucial to the whole research enterprise.

Executive Summary

Publishers have pointed to the final published VOR as a critical component in achieving an open research future, where all scholarly outputs are credible, searchable, linked and available in perpetuity. They have highlighted that while gold OA provides immediate access to the authoritative VOR, enhanced through the publication process, green OA provides often delayed access to earlier, incomplete and potentially inferior versions of the content, with limited permissions for re-use.²

Yet, to what extent does article version matter to researchers? Does the VOR offer significantly more value to them, to the extent that it would impact the way a researcher might discover, read or share a research output? In this white paper we explore researcher preference for the VOR, compared to other article versions: AMs or preprints.

In collaboration with ResearchGate, Springer Nature undertook an assessment of the usage of Springer Nature content syndicated to the ResearchGate platform, followed by a survey of users to determine:

- The extent to which usage for the VOR differs from author uploaded content, such as a preprint or an AM
- The extent to which users prefer one version of an article to another depending on whether it is being used for reading or research
- Author attitudes towards the sharing and reading of their own work in different formats.

We found overwhelming evidence that researchers prefer the VOR both for general reading and for research, seeing this version as the most authoritative and credible source. As such, this white paper provides clear support for the need to protect and sustain funding for immediate gold OA as a means of increasing access to the VOR. It further emphasises why efforts to increase the sharing of AMs via green OA should be discouraged.

Main findings

Where multiple versions of OA articles are available, the majority of researchers choose the VOR

For Springer Nature content syndicated to ResearchGate, when users have access to both an earlier version (either a preprint or an AM) and the final published VOR, they are more likely to use the VOR. Where the article is OA, and therefore accessible to all readers, use of the VOR was highest, with fewer than 5% accessing earlier versions of the article.

Researchers prefer to read and cite the VOR

In a survey of nearly 1,400 researchers using the ResearchGate platform, researchers showed strong preference for the final published VOR, compared to the AM or preprint. This is true for 'general reading', reading with the intention to inform their own work, and to cite, with the preference strongest for citing articles.

- 83% of respondents preferred working with the VOR for citing content in their own work, compared with 9% preferring AMs, and 2% preferring preprints.

We found overwhelming evidence that researchers prefer the VOR.

Researchers believe the VOR is easier to read and is more reliable

In open text answers, respondents commented on the reassurance that peer review and proof of publication give to the VOR, pointing to the lack of time researchers have to read a large volume of content, and the desire to quickly assess and cite an article. Compared to earlier versions, researchers highlight value added to the VOR through the publication process, including copy editing and typesetting, the inclusion of figures and links to data, and the continued maintenance of the VOR after publication (e.g. corrections).

Researchers are more likely to look for ways to find the VOR, rather than an AM or preprint

Where authors did not have access to the VOR (i.e. they did not have access via a subscription or as a result of it being published OA), the majority – nearly 9 in 10 researchers – will take direct action to gain access to the VOR (e.g. contact the author). There is little variation across discipline, seniority or region, however respondents in the field of Computer Science were most likely to use any article version for both reading and citing.

Alternative versions can offer value, but with caveats on use

Even though the VOR is preferred, many researchers also feel comfortable using a preprint or an AM for reading and, in some instances, for citing. Speed of availability, in particular, is noted as a benefit from preprints. Researchers feel more comfortable using the earlier versions of a manuscript when it is in their research field so that they can make their own judgement on the quality.

The VOR is considered the most authoritative and credible source by the majority of researchers

Researcher preference for the VOR highlights the value added by publishers, in particular with reference to the 'stamp of credibility' that publication in a recognised journal brings.

These findings strongly support publishers' previous assertions that AMs are an inferior, incomplete version of research that do not meet the needs of the research community. Attempts to enable the further growth of green OA and to make AMs more widely available may only add confusion to the scientific record and do not reflect researcher preference as shown here. Instead, these findings support the need to widen the availability of VORs via the gold OA route. Gold OA provides immediate, unrestricted access to the authoritative VOR, and requires continued and sustainable funding as a critical step towards a fully open research future.

Most researchers – nearly 9 in 10 – will take action to gain access to the VOR.

Introduction

In scholarly publishing, multiple versions of an author's publication are often available online, prior to and post-publication. NISO/ALPSP guidelines are regularly cited as providing a set of recommended definitions of scholarly article versions for the publishing community.³ In this white paper, we will focus on the preprint, AM and VOR version. Author original manuscripts, also known as 'preprints', refer to the version of a scholarly paper that precedes formal peer review. AMs, sometimes referred to as 'postprints', are the versions of manuscripts which have been through any applicable journal, editorial and peer review processes and have been accepted by the publisher, but have not yet undergone copy editing to ensure clarity of meaning, and formatting and typesetting to enhance usability. The final, edited version of the article made available by the publisher in the journal (whether physical or digital) is known as the VOR. This is the version maintained for the scientific record by the publisher, including links to any post-publication corrections or retractions.

The availability of different article versions is linked to ongoing debates around OA to scholarly publications, with the OA policy landscape fragmented by regions and disciplines.⁴ AMs can be deposited in institutional and disciplinary repositories, made available online via personal websites and/or shared via social networking sites such as ResearchGate, Mendeley and Academia.edu. Where they are deposited by authors (or their institution or publisher on an author's behalf) in institutional or disciplinary repositories, generally after an embargo period set by the publisher, this is commonly referred to as the 'green' route to OA. In contrast, the final published VOR is made available online by the publisher on the journal website immediately upon publication. Where these are published OA, thereby free-to-read, share and re-use, it is referred to as the 'gold' route to OA. While some promote the green route as the "easy" pathway to OA, others, including Springer Nature, actively champion the gold route to OA as the simplest, most open and most sustainable route to OA.⁵ A recent joint blog from members of OASPA, including both mixed model and fully OA publishers, stated, "All of us recognise the value of full open access to the...VOR. We want to ensure that for those who are committed to OA, this route is fully enabled and supported."⁶

Support for the gold OA route has pointed to increased availability of VORs as a key benefit over green OA. Gold OA delivers immediate access to the authoritative VOR for researchers and readers. The VOR drives open science and open research, by connecting the final published article to a range of further open outputs from preprints to open research data, open protocols, open code, and transparent peer review reports. Sharing open data, code and protocols in particular is supported by gold OA and continued efforts to remove barriers in sharing open data have been enabled by key publisher initiatives.^{7,8} Connecting the VOR to open data is a key driver in speeding up research, facilitating greater re-use, and increasing reproducibility.⁹

In 2020, cOAlition S announced plans to enable AMs to be made OA without an embargo period under a CC BY licence.¹⁰ This zero-embargo green OA approach was met with concern from many publishers, including Springer Nature and Cambridge University Press (CUP).¹¹ CUP's response describes embargo-free green OA as "financially self-defeating", pointing to the potential of undermining both the subscriptions on which it relies, as well as newer models for OA such as Read and

3. NISO (2008)

4. Hook et al., (2019); Eger and Scheufen, (2018)

5. Lucraft et al., (2018)

6. Ferguson et al., (2020)

7. Byrne (2017)

8. Lucraft (2019)

9. Burgelman et al., (2019)

10. cOAlition S

11. Day (2020)

Publish agreements. Importantly, publishers also question the quality of AMs, and the challenges that increased access to AMs may cause, since they provide both a poorer reading experience and an increased risk of errors. The VOR provides researchers with easily navigable full text, embedded figures, supplementary information and extended data. Downloadable citation details, article-level metrics, enhanced reference lists with links to articles or full-text, and links to corrections and additional commentaries are also available as part of the VOR. These dynamic and enhanced features provide researchers with the full and complete reading experience to use and build upon in their own work, which is not matched or comparable when accessing the AM. In their joint blog, OASPA members stated, “[green OA] confuses the scholarly record with multiple inferior versions of manuscripts. Do we want researchers to have to search through repositories for an earlier version of a manuscript, and then spend further valuable time seeking out accompanying data, or checking whether there have been post-publication corrections? Or would we rather that they have immediate access to the trusted and enhanced VOR on the publisher platform, with links to relevant data and other outputs?”¹²

This white paper looks to provide further evidence of the value of the VOR and immediate gold OA, bringing together both analysis of VOR usage, and feedback from readers and authors via an online questionnaire.

Usage analysis

Where previous studies have examined the impact of publishing research gold OA, examining use of the VOR on the publisher platform (Jubb et al., found that OA articles are downloaded from publishers' platforms on average between two and four times as much as subscription-based articles, while a Springer Nature white paper on hybrid journals yielded similar results)¹³, there has been limited analysis comparing usage of VORs to AMs where both versions of the article are available. Numerous difficulties arise in doing so, with the analysis of AM usage particularly challenging, notably because of the scale and diversity of repositories that vary both by regions and subject communities (OpenDOAR indicates that in 2020 there were over 5,300 repositories worldwide).¹⁴ In addition, AMs are often shared by individual authors via their own websites or institutional sites, and social sharing sites such as ResearchGate.

Since March 2019, Springer Nature and ResearchGate have partnered to make full-text articles published in certain Springer Nature journals available to researchers on ResearchGate.¹⁵ The syndication agreement sees Springer Nature VORs automatically uploaded to an author's profile on ResearchGate at the time of publication. Doing so has been shown to increase reach (in terms of overall usage and discoverability) and author satisfaction, ensuring readers gain access to the most up-to-date version of an author's work and making the process as easy as possible for authors by managing the upload on their behalf.¹⁶ Initially a pilot, the partnership has now been extended indefinitely, and aims to include all Nature journal content from 2017 onwards and content from other Springer Nature journals from 2015 onwards, with earlier articles to be included when feasible. Access to the VOR is the default option shown where a user has either an institutional subscription to Springer Nature content ('entitled user') or where the article is published OA. Entitled users can also access alternative versions of an article if these have been uploaded by the author. Other users ('non-entitled'; those without institutional subscriptions) can also access these other versions of the article, along with an enhanced abstract, but not the VOR. Although we cannot assess

12. Ferguson et al., (2020)

13. Jubb et al., (2017), Draux et al., (2018)

14. OpenDOAR

15. Springer Nature (2018)

16. Hawkins et al., (2020)

a true like-for-like comparison of version usage, given the preferential presentation of the VOR on the ResearchGate platform, we are able to dig deeper into usage of articles with multiple formats when both VOR and author uploaded content (AM or preprint) is available, in particular to compare differences in behaviour where the article is published OA. We can then also explore the reasons behind this behaviour via a user survey.

Reader and author attitudes towards article formats

Several studies have sought to understand authors' perception and usage of scholarly papers. The impact and speed of publication is a chief concern among researchers. Subject repositories and social networking sites are often used as discovery tools bringing clear visibility benefits within their epistemic community, whereas authors are more reluctant to use institutional repositories.¹⁸ Although slightly outdated, a questionnaire of 3,000 respondents, and four focus groups with authors in Europe, found that Physicists and Mathematicians had a preference for speed of dissemination over peer review.¹⁹ Similar trends have been observed in other disciplines with long peer review timeframes, such as Economics, and appear to have played a part in the emergence of several longstanding preprint servers.²⁰ There is also a longstanding debate over the relationship between embargo periods and journal 'half-lives' – the median age of articles downloaded from a publisher's website. A 2013 study showed that half-lives were typically shorter for journals in the Health Sciences (median half-life: 25-36 months), and were considerably longer for journals in the Humanities, Physics and Mathematics (median half-life: 49-60 months).²¹

Very little is known about researchers' attitudes towards different article formats, in particular whether and to what extent the article version has, in itself, a particular value. A limited-scope qualitative study investigating authors' motivations and decision factors in OA publishing found that for most authors, green OA is not a fully viable alternative to gold OA and is seen as a second-best alternative.²² This is due, in part, to the poorer layout of the AM compared to the VOR, and in larger part to perceived lack of visibility of AMs, with many authors being unaware that their manuscript had been made available in the institutional repository. By contrast, VORs published OA on the journal site were seen more favourably as authors believed they drove downloads and reached a wider audience.²³ Similarly, Jamali et al., found that academics viewed the usage of repositories negatively and were sceptical of their ability to gain a wider reach.²⁴ Another reason cited in favour of the VOR is the fact that AMs can still be found online even if the VOR article is retracted. There has been concern in recent years at the growing number of retractions, which have increased from about 30 a year in the early 2000s²⁵ to more than 1,800 papers in 2020.²⁶

Geographical, disciplinary and generational factors would all likely present different attitudes and behaviours with regards to article versions. To explore this, we again collaborated with ResearchGate to deliver a survey to understand attitudes towards different article versions. The survey was hosted on ResearchGate and was served to active members who have had some interaction with any Springer Nature publications in the 60 days prior to hosting the survey in October 2020. Nearly 1,400 responses were received representing a range of regions, disciplines and career stages. Further details on methodology and demographics can be found in the appendix, along with limitations.

17. Nicholas et al., (2017); Rowley et al., (2017)
18. Poynder, (2016)
19. Creaser et al., (2010)
20. Chiarelli et al., (2019)
21. Davis, (2013)
22. Van der Graaf et al., (2017)
23. Ibid.
24. Jamali et al., (2014)
25. Van Noorden (2011)
26. Retraction Watch (2020)

Findings

Part one: usage of article versions

ResearchGate enables authors to upload preprints, AMs, and VORs to a single “publication page”. ResearchGate users who have access to an article included in the Springer Nature-ResearchGate partnership through their institution (entitled users who are logged in to ResearchGate) can read and download the full-text PDF of the VOR. These entitled users may be able to access various versions of the full-text of the publication, including early author submitted drafts alongside the VOR provided by the publisher, if such versions have been uploaded by an author. For OA articles, all users can read and download the full-text VOR, as well as accessing other versions, if they are present. As noted in the introduction, the VOR is presented as the default option to both entitled users and where articles are published OA.

ResearchGate provided aggregated usage data on Springer Nature articles included in the pilot during January 2020. There were approximately 38k total accesses across a total of around 44k Springer Nature articles during the month. Of these, approximately 35k consumes occurred on a VOR (Figure 1). Here we use the ResearchGate term ‘consumes’ of a publication to refer to either a download of the full-text PDF, or the user clicking to view the full-text PDF in line on the publication page.

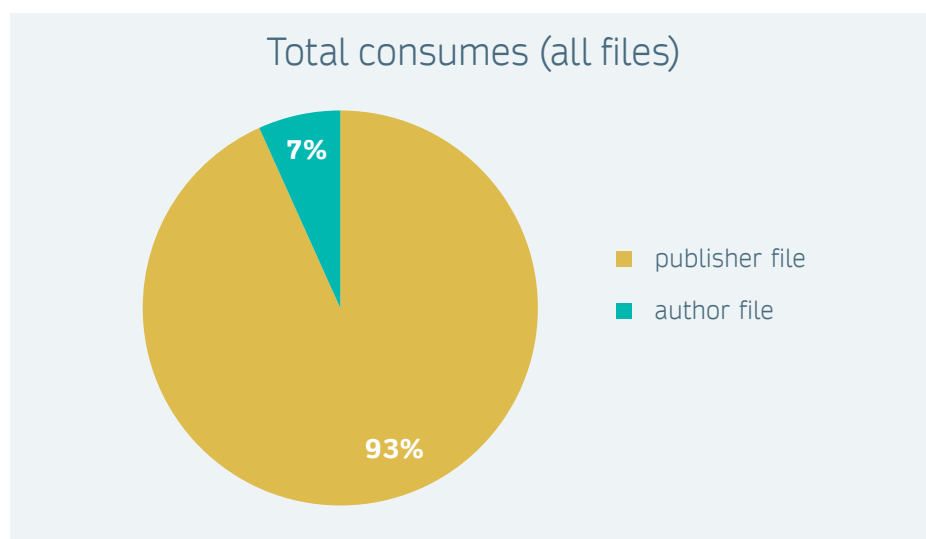


Figure 1: Total consumes of Springer Nature articles syndicated to ResearchGate, January 2020, by type. Shows combined total article PDF downloads or full PDF views for all publisher-uploaded VORs added during the pilot phase ('publisher file') and for 'author files'; author-uploaded files (AMs or preprints) in January 2020, with the majority of consumes occurring on the VOR.

Approximately 1.4k of total consumes came from logged-in entitled users (with institutional access) on articles that had multiple article versions available, with the default option of the VOR gaining most consumes (over four times as many as the number of AM consumes, as shown in Figure 2).

Approximately 2.7k consumes occurred on OA articles. Here, we see noticeably higher use of the VOR. Again in this instance, the VOR is presented as the default option to users, but we see less than 5% accessing author-uploaded versions of the article (Figure 3). Overall, for all publications with multiple versions available, the majority of article

January 2020 logged-in consumes (pilot publications with multiple files) - entitled users only

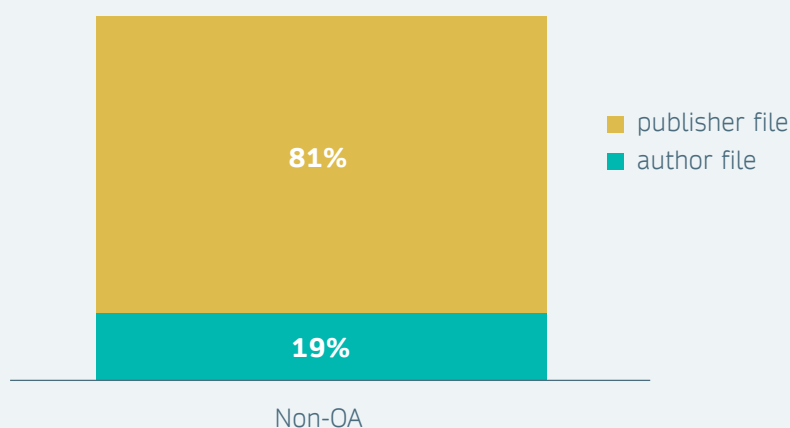


Figure 2: Pilot article consumes by type (publisher file or author file), for logged-in entitled users. *Shows subset of all pilot articles that had more than one version available in January 2020, with usage by logged-in entitled users (with default access to the VoR).*

consumption occurs on the VOR. Given the presentation of this version as the default, this is perhaps not surprising, but the higher usage of the VOR where articles are published OA requires further exploration. Although it is not possible to distinguish whether consumes of the VOR are instead of, or in addition to, author files, it is likely that for non-OA content, users download multiple versions of the same file, as the AM is visibly different to the VOR. In the case of OA articles, it is more likely that the author-uploaded version and publisher-uploaded version is the same version (VOR), reducing the need to download multiple files. In the second part of our white paper, we will continue to analyse user behavior and try to validate those assumptions, turning our attention to researcher attitudes and perceptions when it comes to article versions.

January 2020 OA consumes (pilot publications with multiple files) - all users

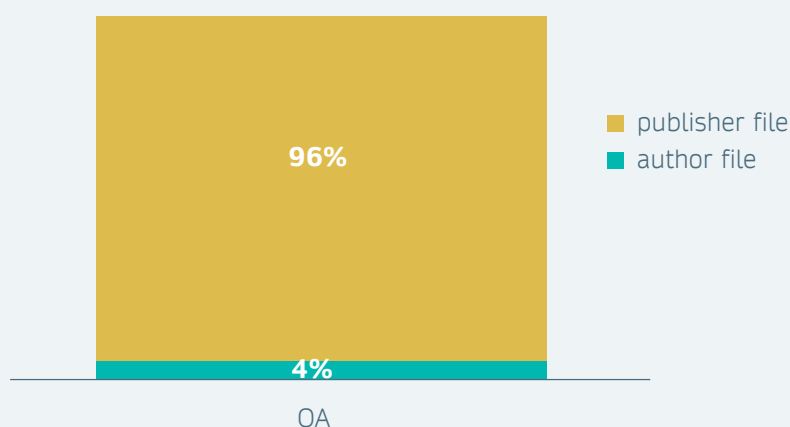


Figure 3: OA article consumes by type (publisher file or author file), for all users. *Shows subset of all pilot articles that have more than one version available, where the final published article is OA.*

Part two: user preference (reader survey)

To be able to examine preference for different article versions, participants in the survey were firstly asked to familiarise themselves with the terms as defined below and to then check a box to confirm they had read and understood them:

Preprint: The author's version of a research manuscript, before formal peer review, deposited on a public server. Preprints can be posted at any time during the peer review process.

Accepted manuscript (AM): The version of a research paper accepted by a journal after peer review. This version does not include any changes made after acceptance such as copy editing, or typesetting. Usually AMs are shared on a repository or the author's website, and not the journal website.

Final published version: The version that has been published in a journal, in print and/or online. This article will include any editorial improvements such as copy editing, or typesetting, made after the peer review process is complete. It is usually available on a publisher website, in PDF and/or HTML format.

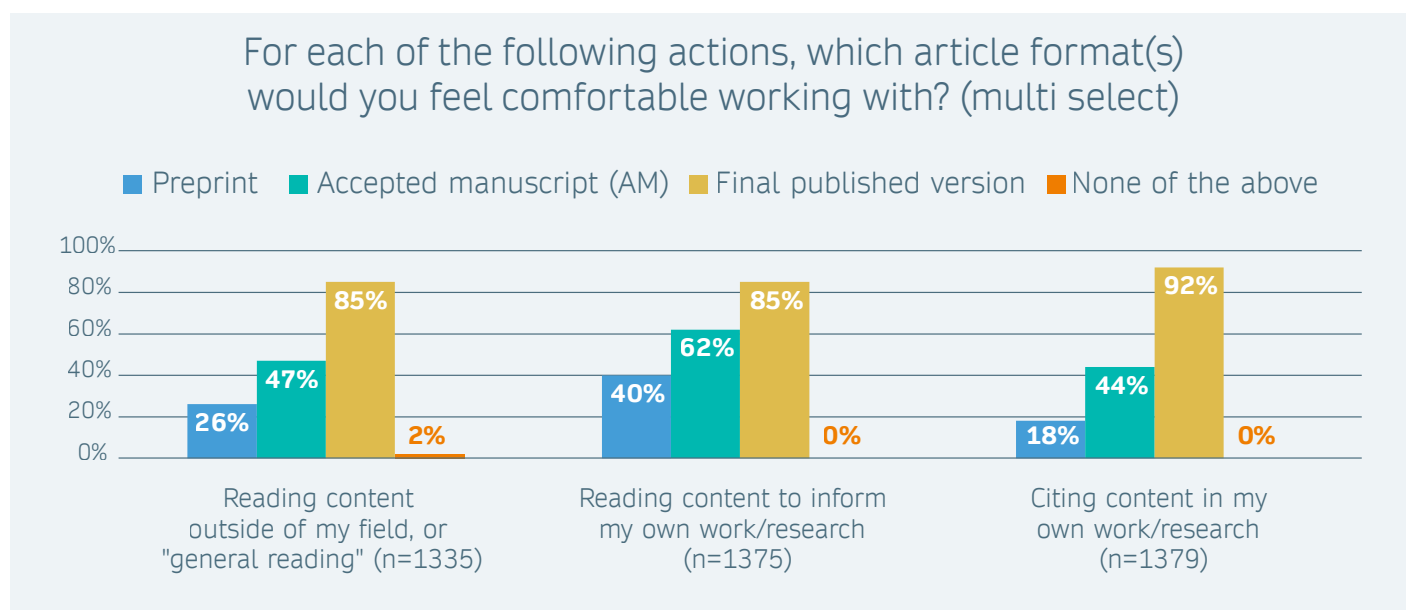
For simplicity, in the white paper we will continue to refer to the final published version as the VOR. For more details on the methodology and demographics see the appendix.

Researchers feel more comfortable working with the VOR for reading and research

Readers were asked for each article format which they would feel comfortable using, considering reading for general interest, reading to inform their own research, or citing for their research.

For all three activities, respondents feel more comfortable with the VOR than other article versions, with limited variation by discipline, region or career level (Figure 4).

Figure 4: Which versions do researchers feel comfortable working with for different activities?²⁷



²⁷ Those who responded 'Unsure' have been removed. Significance testing $p=0.05$ (5% probability finding is by chance)

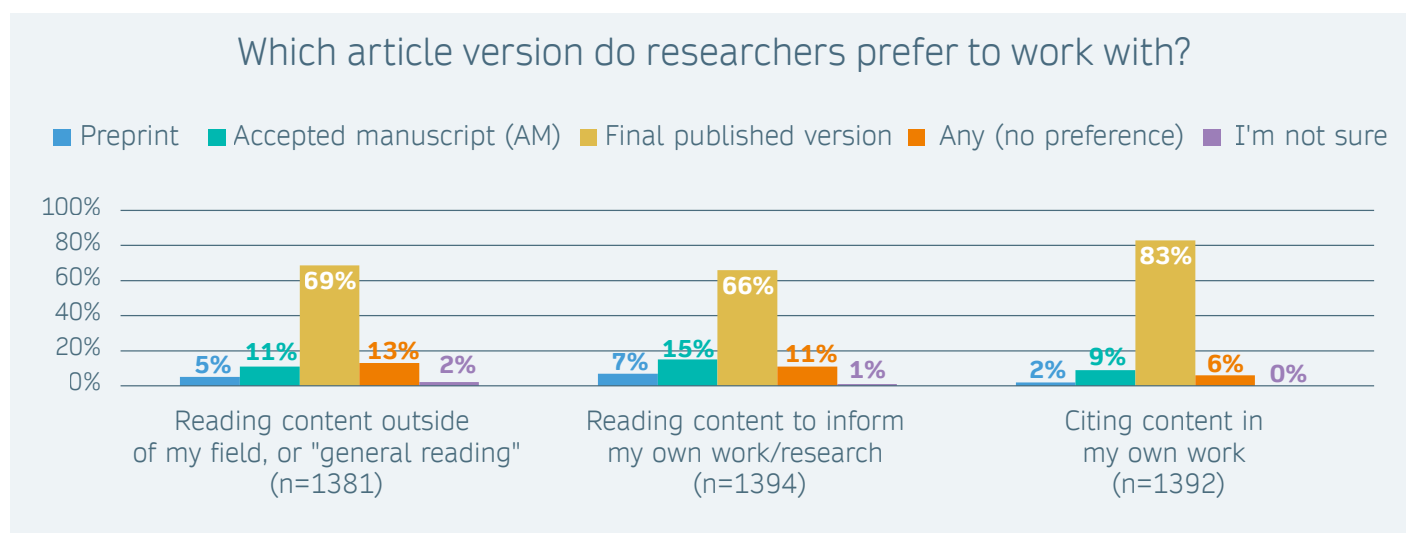
The preference is strongest when researchers consider what version of an article they would be comfortable citing in their own work, with 92% feeling comfortable using the VOR, compared with less than half who would feel comfortable citing an AM (44%), and less than a fifth (18%) who would feel comfortable citing a preprint.

For reading, more researchers were comfortable using alternative article versions, although again most were comfortable with the VOR (85% for both general reading and to inform research). Although minimal demographic variation was seen in general, more respondents from Western Europe were comfortable working with preprints (54%) or AMs (74%) for reading to inform their work, compared with other regions. A possible reason for this is the number of funder policies in this region that include AM deposition as a requirement of their funding mandate. As OA policies have been longer established within Europe, it is also more likely that researchers in this region are more familiar with funding grants and their requirements. As such, researchers are more familiar with the AM, or more comfortable working with it.

Researchers express preference for the VOR over other formats

Respondents were then asked which version of articles they *preferred* to work with. Again, here we see similarly high preference for the VOR compared with other article versions, especially for citing articles. 83% preferred to use the VOR when citing, compared with just 9% who prefer the AM and 2% who prefer to cite preprints (Figure 5).

Figure 5: Which article version do researchers prefer to work with?



There is slightly more variation in answers when it comes to reading, with more researchers showing no preference on what version of article they use, or showing preference towards using preprints or AMs, compared to when they are asked about citing.

In coded open text responses, the majority of respondents point to the reliability of the VOR. Published articles were seen to be the most credible as they were peer-reviewed (Figure 6):

"Final published versions are endorsed by peers, editors and publishers, thus representing material with the highest value in academia."

Experienced researcher²⁸, Western Europe, Biology

Many referred to VORs being better for citations, with stable DOIs, and in many cases were understood to be the only citations accepted by major journals:²⁹

28. See summary of the categorisation of researchers in the appendix

29. In reality, this is a mixed picture and policies vary by publisher. Springer Nature's preprint policy states that preprints may be cited in the reference list of articles under consideration. For further information see <https://www.springernature.com/gp/policies/editorial-policies>. See also Chiarelli et al., (2019)

Please elaborate on the reasons for your preferred article type - coded open text answers



"I want my material to be stable, have a DOI, be easy to refer to. Especially the DOI is important to me, that [sic] why it should be properly published. Also the time between acceptance and online publication tends to be short, so I don't see much benefit to using something else."

Early career researcher, Western Europe, Computer Science & Mathematics

Figure 6: Coded open text comments that explain preference for different article formats (868 comments)

VORs were also acknowledged as being easier to read or having better formatting, and therefore quicker to use:

"I prefer to read and study a final formatted paper with well structured text, figures and tables."

Career stage unknown, Western Europe, Engineering

Overall, the preference for the VOR references authority and speed, both of which complement the previous research on researcher behaviours:

"If I'm going to take the time to skimread through something – and there's both tons to skimread and very limited time for me to do it – I need the text to be final so that I can credibly cite it. I don't feel Preprints, having not yet gone through review, meet this standard. Similarly, for citing pieces (not just reading), I'm somewhat hesitant about AMs, because final precise wording may change, which matters with quotations."

Experienced researcher, US & Canada, Anthropology

10% of respondents highlighted the potential for errors in earlier versions. Comments mentioned that these manuscript types were not preferable to use in research due to the possibility of changes in its final form:

"A preprint may lack the complete data set, so I hesitate to cite it unless the field is new and rapidly developing (e.g. SARS-CoV-2)."

Senior researcher, US & Canada, Medicine

Although, as previously noted, we see minimal differences across demographics, the field of Computer Science ($n=99$) is one area where there are significant differences in answers provided, with 24% saying they would use 'any' article version for general reading, the same proportion answered 'any' about reading to support their own work, and 14% said they would use any article version for citing. These are all significantly higher than other disciplines and the averages ($p=0.05$). For the small group of respondents (12%) who preferred to use preprints and AMs, this was particularly for general reading in their own field of research. In fact, 33% of those who gave a reason for using preprints or AMs also said it depended on if they were reading within their own area of expertise. In other words, if an article was within their research field, they were more confident in evaluating its validity:

"It is most convenient to cite the published version, which may be more easily accessible (including DOI etc.), and it may be more acceptable by some journals than citing preprints. When I am familiar with the research field, reading preprints give [sic] me some months advantage in terms of being aware of the "leading edge" in the research field."

Experienced researcher, Western Europe, Engineering

In a number of comments, the ability to get up to date with research quickly (notably in fast moving fields) was cited as a reason for preferring preprints:

"I'm a virologist by training and one of my current projects involves SARS-CoV-2. Therefore, it's just impossible to read/cite only those articles with a final published version available."

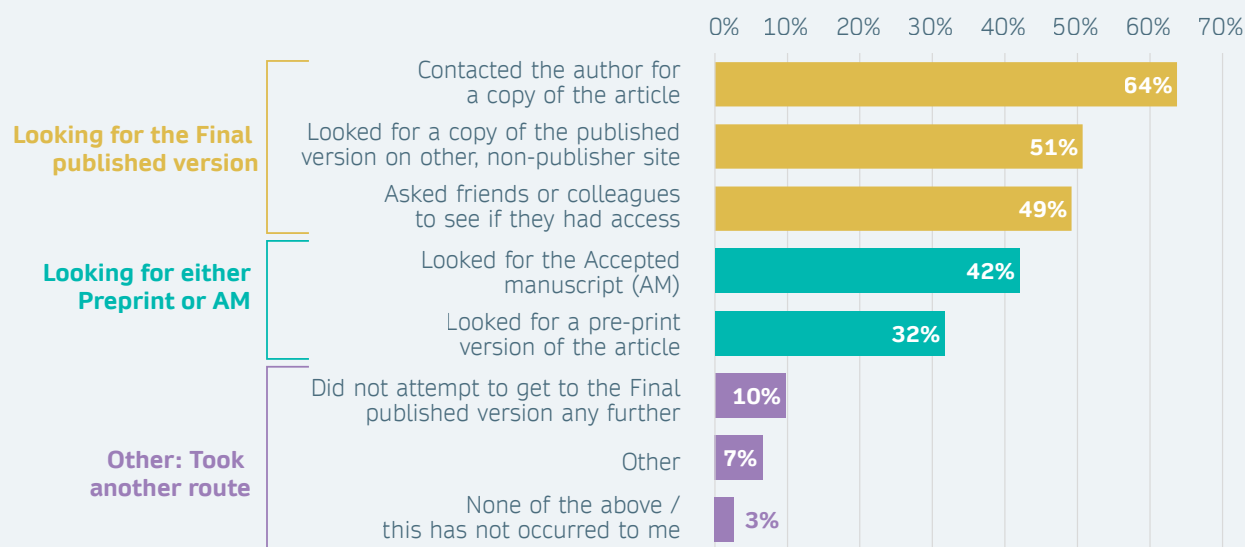
Experienced researcher, US & Canada, Engineering

Most researchers will look for ways to find the VOR rather than AMs or preprints

We asked respondents whether they took any action if they did not have access to the VOR, for example contacting authors directly, or looking for a copy on non-publisher sites. The majority of respondents were contacting the authors to request a copy (64%).

Approximately half of respondents would look for the VOR on other sites (51%), or ask colleagues if they had access (49%). This shows the desirability of the VOR to researchers, compared with AMs or preprints, where only 42% of respondents said they would look for the AM, and 32% would look for the preprint in the absence of the VOR (Figure 7). Although this was not explicitly asked in the survey, it is likely that researchers are seeking out the VOR when citing research, based on the evidence of version preference we noted above.

Actions taken when previously unable to get hold of final published version of a paper (n=1398) (multi select)



What is particularly interesting to note is that of these responses, 89% of respondents – nearly 9 out of 10 researchers – took at least one action to gain access to the VOR. Where just over half of the respondents would look for the VOR or preprint/AM in combination, preference for the VOR is again seen with 35% of respondents who would look only for the VOR and no other version, compared with just 6% who reported looking for either the AM or preprint and taking no other action. This again shows a clear preference, and determination from researchers, to have access to the authoritative final publisher VOR (Figure 8).

Figure 7: What action will authors take in the absence of a VOR?

Actions taken when previously unable to get hold of Final published version of a paper (grouped answers)

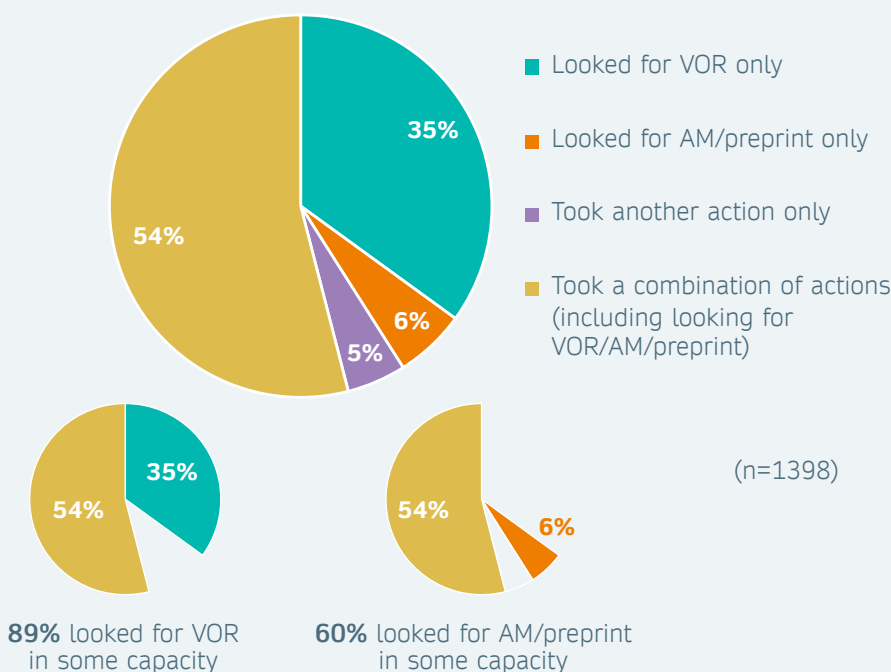


Figure 8: Actions taken when unable to get hold of Final published version, by article version³⁰

30. Multiselect question as shown in Figure 6. Other actions people could select in question were i) Other; ii) None of the above / this has not occurred to me; iii) Did not attempt to get to the Final published version any further

Authors of research articles are most comfortable with readers having access to the VOR

Researchers overwhelmingly prefer the VOR when thinking about how others make use of their own research (91%), with nearly half (47%) suggesting that this was the only version they would like readers to access. In open text comments (579 responses), authors call out reliability and ease of reading again, as seen in the earlier comments:

"As a researcher it has always been important for me to know the credibility of my research by sending them for peer-review to make it scientifically accepted which may increase its impact. Also, being published in its final format increases its value, showing that these paper/s they are reading or citing is scientifically accepted."

Senior researcher, Asia, Biology & Chemistry

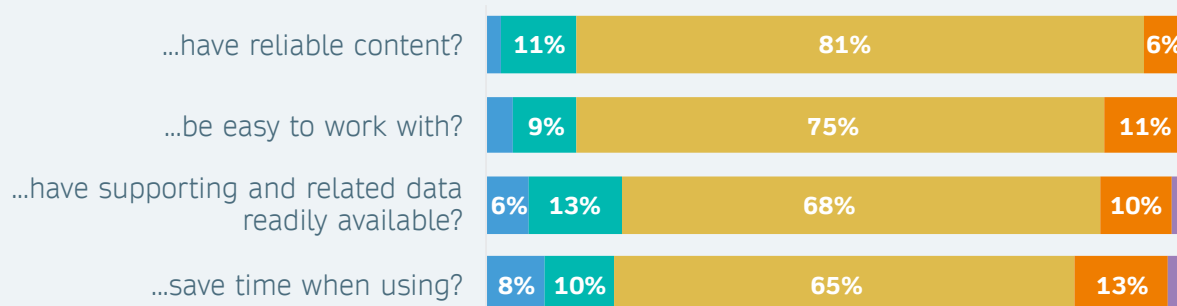
Overall the VOR is perceived to be more reliable, easier to use, and saves time for researchers

Respondents were asked to rate all three formats against the same criteria: reliability; ease of use; availability of data; and whether the format could save them time. As with previous questions, we again see the VOR consistently scoring highest across all factors, especially reliability, as might be expected from the previous comments on why the VOR is preferred for citations, and the impact of peer review, copy editing, typesetting, as well as links to data and corrections. The preference for the VOR further strengthens the position asserted by publishers on the value of the VOR as the most credible, reliable source that researchers are looking for to support their work.

Figure 9: Comparison of article formats against expected value factors

Which of the three article formats would you most expect to... (n=1,378-1,383)

■ Preprint ■ Accepted manuscript (AM) ■ Final published version ■ Any (no preference) ■ I'm not sure



NB. % under 5% not shown

Discussion and conclusions

By setting out to better understand researcher behaviours regarding article formats, we have demonstrated here that the VOR is the most valued version. Researchers show strong preference for the VOR, due to the need to access an easily navigable, reliable and trusted source of the scientific record and to ensure they base their research on the authoritative, most up-to-date, and complete version of the work. These findings strongly support publishers' previous assertions that AMs are an inferior, incomplete version and may add confusion to the scientific record.

As we continue towards a fully open research future, where articles are connected to the complete research output, and the many benefits this brings for speed, integrity and efficiency in science, gold OA with immediate access to the VOR is required. Although open science is not explicitly referenced in the feedback from researchers in this white paper, the clear preference for the VOR as the complete, authoritative version shows why attempts to make the AM more widely available should be discouraged.

This does not diminish the importance of preprints and early sharing as part of open research. As noted by many respondents, preprints in particular provide early insights, particularly valid for fast moving disciplines. However researchers comment that they feel most comfortable using these within their own field of research, and the majority of researchers would not choose to cite either the preprint or the AM version of work in their final research. Similarly, these findings do not suggest there is no relevance for repositories in supporting an open research future. As noted in a recent cOAlition S blog, repositories have successfully increased discoverability of articles.³¹ Rather, the strong author preference indicated in this white paper shows that wider access to the VOR, over AMs and preprints, is critical if we are to further open scholarship. OA versions of articles are also being deposited into repositories and we recommend this practice continues.

Access to VORs through gold OA continues to grow at an unprecedented rate, with DOAJ recording more than 15,000 fully-OA journals and more than 5.5m articles in those journals at the time of writing.³² Publishers continue to support immediate gold OA via Transformative Agreements (ESAC references 125 current transformative agreements in place across 20 countries in Europe, Middle East, Asia and North America, committing nearly 100,000 gold OA articles per year³³). The future growth of this gold OA model relies heavily on the commitment and investment of funders and institutions, ensuring that the authoritative, credible VOR is made immediately available to researchers without restrictions and limitations. Further investment in the AM and green OA, either financially (for example investment in AM infrastructure or discoverability services) or by policies, is not, in the long run, a sustainable or desirable option. It may in fact be detrimental in the drive to grow access to the VOR via the gold OA route, thereby delaying or preventing a full and complete transition to open research. We call for a continued and coordinated effort to grow gold OA, and therefore access to VORs globally, making not only the research article, but every facet of research openly and immediately available, accessible and reusable by all.

The clear preference for the VOR as the complete, authoritative version shows why attempts to make the AM more widely available should be discouraged.

31. Rumsey (2021)

32. Directory of Open Access Journals (DOAJ)

33. Ferguson et al., (2020)

Appendix

Research Methodology

We surveyed 1,398 ResearchGate active members who had interacted with at least one Springer Nature publication in the previous 60 days prior to the survey being live in October 2020. An invitation to take part in the survey was displayed to active members whilst they were on the ResearchGate site. Respondents were offered the chance to enter a prize draw to win one of three \$150 digital gift cards once they had answered all questions.

Demographic overview

Geography: 37% of responses came from people based in the continent of Asia, 33% from people based in Europe, 20% responses came from those based in the Americas, and the remaining 10% came from other regions around the world.

Subject discipline: Respondents were categorised in up to three different disciplines, the most common being: Biology (33%), Engineering (28%), Chemistry (18%) and Physics (15%).

Career Stage: Respondents were categorised into one of three of the following career stages: "Early career researcher" who have published on average 14 papers (18%), "Experienced researcher" who have published on average 27 papers (42%) and "Senior researcher" (35%) who have published on average 49 papers. For the remaining 5% their career stage was unknown.

Limitations

As noted in the introduction, the usage shown in this report reflects the available data for Springer Nature articles on ResearchGate as part of a syndication pilot in January 2020. The presentation of the VOR as the default creates a bias towards the VOR, and it is therefore difficult to compare like-for-like usage across formats. A more accurate method to compare VOR usage to AMs or preprints was not identified for this study.

For the survey, there is a chance of selection bias as the survey only represents the views of ResearchGate members who interacted with Springer Nature articles. As such, these findings should be taken as a sample of researchers, it may not necessarily be representative of all researchers.

References

Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., Schoupe, M., 2019. Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data*. 2: 43. URL <https://www.frontiersin.org/article/10.3389/fdata.2019.00043>. DOI 10.3389/fdata.2019.00043

Byrne, M., 2017. Making Progress Toward Open Data: Reflections on Data Sharing at PLOS ONE. URL <https://everyone.plos.org/2017/05/08/making-progress-toward-open-data/> (accessed 1.28.2021)

Chiarelli, A., Johnson, R., Pinfield, S., Richens, E., 2019. Accelerating scholarly communication: The transformative role of preprints. Zenodo. <https://doi.org/10.5281/zenodo.3357727>

cOAlition S. Plan S Rights Retention Strategy. URL <https://www.coalition-s.org/rights-retention-strategy/> (accessed 1.21.21)

Creaser, C., Fry, J., Greenwood, H., Oppenheim, C., Proberts, S., Spezi, V., White, S., 2010. Full article: Authors' Awareness and Attitudes Toward Open Access Repositories [WWW Document]. URL <https://www.tandfonline.com/doi/full/10.1080/13614533.2010.518851> (accessed 3.31.20)

Davis, P., 2013. What is the Lifespan of a Research Article? Sch. Kitchen. URL <https://scholarlykitchen.sspnet.org/2013/12/18/what-is-the-lifespan-of-a-research-article/> (accessed 4.4.20)

Day, M., 2020. CUP's response to the Rights Retention Strategy from cOAlition S. URL <https://www.cambridge.org/core/blog/2020/10/19/openresearch-cups-response-to-the-rights-retention-strategy-from-coalition-s/>

Directory of Open Access Journals (DOAJ) URL <https://doaj.org/> (accessed 1.21.21)

Draux, H., Lucraft, M., Walker, J., 2018. Assessing the open access effect for hybrid journals | Open research | Springer Nature. <https://doi.org/10.6084/m9.figshare.6396290>

Eger, T., Scheufen, M., 2018. The Economics of open access: On the future of academic publishing. Edward Elgar, Cheltenham. URL https://www.jipitec.eu/issues/jipitec-9-3-2018/4812/JIPITEC_9_3_2018_345_Steinhauer_Book_Review (accessed 1.28.2021)

Ferguson, L., Sutton, C., Taylor, S., Webster, C., Wilson, E., Hall, S., Peters, P., Day, M., Legge, M., Clark, D., Ross, D., OASPA. 2020. Open post: the rise of immediate green OA undermines progress. URL <https://oaspa.org/open-post-the-rise-of-immediate-green-oa-undermines-progress/> (accessed 1.21.21)

Hawkins, E., Hofmayer, S., Noyes, D., Schoenenberger, H., Winter, S., 2020. Researchers at the centre: content discoverability, visibility, and access. <https://resource-cms.springernature.com/springer-cms/rest/v1/content/18300962/data/v4>

Hook, D., Hahnel, M., Calvert, I., 2019. The Ascent of Open Access (report). Digital Science. <https://doi.org/10.6084/m9.figshare.7618751.v2>

Jamali, H.R., Nicholas, D., Watkinson, A., Herman, E., Tenopir, C., Levine, K., Allard, S., Christian, L., Volentine, R., Boehm, R., Nichols, F., 2014. How scholars implement trust in their reading, citing and publishing activities: Geographical differences. *Libr. Inf. Sci. Res.* 36: 192–202. <https://doi.org/10.1016/j.lisr.2014.08.002>

Jubb, M., Plume, A., Oeben, S., Brammer, L., Johnson, R., Bütün, C., Pinfield, S., 2017. Monitoring the transition to open access: December 2017 [WWW Document]. URL <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/monitoring-transition-open-access-2017.aspx> (accessed 3.31.20)

Lucraft, M., Baynes, G., Allin, K., Hrynaszkiewicz, I., Khodiyar, V., 2019: Five Essential Factors for Data Sharing. figshare. Journal contribution. <https://doi.org/10.6084/m9.figshare.7807949.v2>

Lucraft, M., Calder, C., Pyne, R., Monaghan, J., Spinka, V., 2018. Gold Open Access in the UK: Springer Nature's Transition. figshare. Journal contribution. <https://doi.org/10.6084/m9.figshare.6230813.v1>

Nicholas, D., Watkinson, A., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Xu, J., Abrizah, A., Świgoń, M., Herman, E., 2017. Early career researchers: Scholarly behaviour and the prospect of change. *Learned Publishing*, 30: 157–166. <https://doi.org/10.1002/leap.1098> (accessed 3.31.20)

National Information Standards Organization (NISO)., 2008. Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group. URL <https://www.niso.org/sites/default/files/2017-08/RP-8-2008.pdf> (accessed 1.29.2021)

OpenDOAR. URL https://v2.sherpa.ac.uk/view/repository_visualisations/1.html (accessed 1.21.21)

Poynder, R., 2016. Open and Shut?: Q&A with CNI's Clifford Lynch: Time to re-think the institutional repository? Open Shut. URL https://poynder.blogspot.com/2016/09/q-with-cn-is-clifford-lynch-time-to-re_22.html (accessed 8.17.18)

Retraction Watch. 2020. A look back at retraction news in 2020 — and ahead to 2021. URL <https://retractionwatch.com/2020/12/31/a-look-back-at-retraction-news-in-2020-and-ahead-to-2021/> (accessed 01.21.21)

Rowley, J., Johnson, F., Sbaffi, L., Frass, W., Devine, E., 2017. Academics' behaviors and attitudes towards open access publishing in scholarly journals. *J. Assoc. Inf. Sci. Technol.* 68, 1201–1211. <https://doi.org/10.1002/asi.23710>

Rumsey, S., 2021. cOAlition S and repositories (part II) The many benefits of repositories. URL <https://www.coalition-s.org/blog/coalition-s-and-repositories-part-ii/> (accessed 1.21.21)

Springer Nature. 2018. Springer Nature and ResearchGate announce new cooperation to make it easier to navigate the sharing of academic journal articles URL <https://group.springernature.com/gp/group/media/press-releases/springer-nature-and-researchgate-announce-new-cooperation/15705990> (accessed 1.21.21)

Van der Graaf, M., Johnson, R., Chiarelli, A., 2017. The Role of Hybrid Open Access in Extending Author Choice [WWW Document]. URL <https://zenodo.org/record/3958621#.YAluolX7TIU> (accessed 1.22.21)

Van Noorden, R., 2011. Science publishing: The trouble with retractions. *Nature* 478, 26-28. doi:10.1038/478026a URL <https://www.nature.com/news/2011/111005/full/478026a.html> (accessed 1.28.2021)

Vrancken Peeters, F., 2021. Partnerships needed to advance open science. URL <https://www.springernature.com/gp/advancing-discovery/blog/blogposts/partnerships-to-gain-trust-and-advance-open-science/18767374> (accessed 1.2.2021)

Acknowledgements

We would like to thank the following for their contributions to this white paper:

Springer Nature

Carrie Webster

Jessica Monaghan

Susie Winter

Dan Penny

Sebastian Bock

Henning Schoenenberger

Harald Wirsching

Emma Goldsmith

Nick Campbell

Steven Inchcoombe

Lucy Frisch

Christina Emery

Isabel Roth

Chloe Davis

ResearchGate

Thomas Boyle

Michael Hauesler

Sören Hofmayer

Dave Hora

Dan Noyes

Sebastian Reichau

Research Consulting

Rob Johnson

Victoria Ficarra

Mattia Fosci

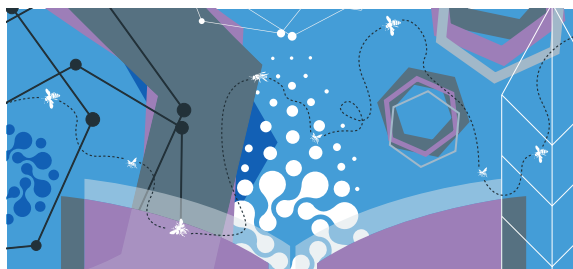
Shift Insight

Elsie Lauchlan

Gee St. Design

Rob Dicks

Around our complex and interconnected world, the research community is advancing discovery for all of us. These illustrations celebrate some of the great minds who have helped advance discovery through history.



Jean-Claude Bradley (1969–2014)

Jean-Claude Bradley was a chemist and passionate proponent of Open Science. Following an early career in patent driven nanotechnology, Bradley came to believe that the work he was doing wasn't having the impact or benefitting mankind in the way he had hoped. At Drexel University, working on antimalarials, he coined the term Open Notebook Science for an approach which aimed to make the details and raw scientific data of every experiment done in the lab freely available within hours of production. Bradley was founding Editor-in-Chief of *Chemistry Central Journal* and a founding Editor of the *Journal of Cheminformatics*. In 2007 he was awarded a Blue Obelisk award for achievements in promoting Open Data, Open Source and Open Standards.

The Open Research portfolio:

BMC

Journals including:

The BMC Series
Genome Biology
Genome Medicine
BMC Biology
BMC Medicine

Nature Research

Journals including:

Nature Communications
Communications Biology
Communications Chemistry
Communications Physics
Scientific Data
Scientific Reports
 Nature Partner Journals

Springer books and journals including:

Springer Open Choice

Palgrave Macmillan

Books and journals including:

Humanities and Social
Sciences Communications