



This is a repository copy of *The elusive gold mine? The finer details of creative commons licences – and why they really matter.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/155227/>

Version: Published Version

Article:

Mallalieu, R. orcid.org/0000-0002-0409-6199 (2019) The elusive gold mine? The finer details of creative commons licences – and why they really matter. *Insights*, 32 (1).

<https://doi.org/10.1629/uksg.448>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The elusive gold mine? The finer details of Creative Commons licences – and why they really matter

This article explores the tensions inherent in the ownership and reuse of scholarly works, with a focus on achieving open access (OA) aims via Creative Commons licences. This US initiative has had an enormous impact on the access to, dissemination and reuse of UK-authored scholarly literature since the Finch report of 2012. However, confusion abounds within the funding, publishing and academic communities about the correct uses and long-term implications of using such licences. This has legal consequences, as well as consequences for the author, readers and institutions who have to report compliance regarding OA in order to secure future research funding.

Ownership is an important part of this picture. Creative Commons licences are legally binding on licensor and licensee, and only the copyright owner may release their work under such a licence. The complex research funding and sharing ecosystem has resulted in a 'policy stack' challenge¹ with authors given little choice about their options. This paper examines some of these challenges through an exploration of current UKRI policy and the copyright licences of one publisher, with a focus on text and data mining.

Keywords

Licences; Creative Commons; Elsevier; RCUK; UKRI; copyright; ownership; reuse; text and data mining; TDM

Introduction: Creative Commons licences – what's all the fuss about?

In 2001 the Budapest Open Access Initiative held that 'the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited'.² Two years later the Berlin Declaration stated that open access (OA) contributions must 'grant to all users a free, irrevocable, worldwide, right of access to, and a license to copy ... and to make and distribute derivative works ... subject to proper attribution of authorship'.³

In this context, the word 'access' is something of a misnomer. Interpreted narrowly, it means free to view and read, and potentially to download for personal use.⁴ These ground-breaking statements of intent went much further than that – and with good reason. True access under these concepts means freedom to access, download, interrogate, rework, republish. Academic norms of attribution and moral integrity must remain. Without going over all of the compelling arguments for open access here,⁵ suffice it to say that – at a minimum – the outputs from publicly funded research should be available to all to read and scrutinize. In order for the benefits of research to flow back into society, certain rights that are usually the preserve of the copyright holder must be shared – although to what extent is a matter of ongoing debate. In UK law, these consist broadly of the rights to copy the work, issue copies to the public, lend the work,



RUTH MALLALIEU

Scholarly
Communications
Licensing Manager
Copyright & Licensing
Team
University of
Sheffield

'for the benefits of research to flow back into society, certain rights ... must be shared'

- 2 perform the work, communicate the work to the public, make derivative versions of the work – and to perform any of these acts with any derivative versions.⁶

Creative Commons licences have been in existence, in various forms, since 2002. These founding principles of the OA movement were written before their invention and subsequent widespread use. Their definitions of open access – requirement of attribution, commercial and derivative use allowed with no further restrictions save the moral rights of the author – align with what is now the Creative Commons Attribution (CC BY) licence. Open licences provide a neat solution for a copyright holder to retain copyright whilst making their work available under clear terms.

The starting point, then, is the copyright holder. First ownership of academic works is not straightforward to unpick – and in an environment in which authors are expected to behave in certain ways with regards their intellectual outputs, it may become something of a moot point, as is described below.

Copyright and first ownership in the UK

Under the Copyright, Designs and Patents Act 1988 (CDPA), the author of a work is the first owner of the copyright in that work. However, in law certain categories of works, including literary works that are created by employees in the course of their employment, are treated differently. In such cases the employer is the first owner of copyright, subject to any agreement to the contrary.

Copyright ownership, especially in the higher education context, raises ideological and practical considerations. Studies show a very mixed picture of ownership within universities and actions regarding assignment to publishers, with one clear conclusion: many authors are not aware of, firstly, who owns the rights in their scholarly works, and, secondly, what they and others are allowed to do with them. The latest, a 2017 study into university intellectual property policies, showed that a majority of respondents stated that either staff own the copyright outright in scholarly works (44/81), or the higher education institution (HEI) owns but waives their first ownership claim (16/81) or assigns the rights to staff (7/81) with a mixture of other approaches taken.⁷

'HEIs do not exercise their ownership rights on scholarly outputs created by their staff'

Sometimes these matters are explicitly laid out in contracts of employment, sometimes by custom and practice. Either way, in the majority of cases HEIs do not exercise their ownership rights on scholarly outputs created by their staff.

On the face it, therefore, many academic authors are free to choose how they license their works. This is pertinent when it comes to Creative Commons licences. Only the copyright owner may make their work available under licence. Once they do so, the terms are irrevocable insofar as a person who has accessed a work under the terms may continue to rely on those terms in future – even if the licence is later withdrawn – and apply for the full term of copyright. In the UK, US and many European countries, this is 70 years following the end of the year in which the author dies. The duration of copyright remains tied to the lifetime of the creator even if the creator subsequently assigns their copyright – for example, to a publisher as a condition of the publishing contract.

Policy, publishing and the illusion of choice

Many universities have traditionally waived ownership claims for practical purposes – to allow authors to sign publishing contracts, assigning copyright if necessary – as well as for cultural reasons. Historically, many publishers have asked for copyright assignation as standard, with authors not necessarily realizing that this results in a transfer of all rights: a 2003 study found that 79% of academics claimed that they owned copyright despite 90% of publishers asking for an assignation.⁸

3 In the current UK policy landscape, gold OA⁹ is the preferred model, with the weight of UK research funding behind it. The 2017 *Monitoring the transition to open access* report showed that, in 2016, 37% of UK-authored articles were accessible immediately on publication, with 30% of those under gold OA. The Research Councils (RCUK) – now part of UK Research and Innovation (UKRI) – have provided funding via block grants, year on year, with a 2016 sample of 38 universities showing that [then] RCUK funding accounted for 80% of article processing charge (APC) payments.¹⁰

The RCUK/UKRI policy language is the same today as it has been since the first post-Finch policy in 2013, espousing a preference for ‘immediate Open Access with the maximum opportunity for re-use’.¹¹ It goes on to say that, where an APC has been paid by one of the Research Councils, the published version must be made available immediately on publication under the most permissive Creative Commons licence that passes the litmus test of ‘true open access’: the Attribution (CC BY) licence.

Conversely, the track of green OA compliance can be extremely convoluted. No publishing fee is paid and publishing costs are typically supported by subscriptions, with the resulting ‘access’ to research sometimes nebulous. Before the merger of RCUK and the Higher Education Funding Council for England (HEFCE), each had separate policies. HEFCE permitted the use of a Creative Commons Attribution Non-Commercial No-Derivatives (CC BY-NC-ND) licence.¹² RCUK, however, required that the accepted manuscript (AM) version be made available ‘without restriction on non-commercial re-use’, naming the Creative Commons Attribution Non-Commercial (CC BY-NC) licence as the most restrictive that would apply.¹³ The FAQs accompanying the policy have been updated over time. In their most recent incarnation, they specifically exclude the use of a CC BY-NC-ND licence. There is also a requirement that no restriction is placed on ‘non-commercial text- and data-mining’.¹⁴

This is for very good reason. CC BY-NC-ND is the most restrictive of the CC licences. Anything beyond CC BY does not, strictly speaking, meet the Budapest/Berlin standards. Nevertheless, the non-commercial element makes more sense when the research was unfunded. Restrictions on derivative works make less sense. Many national copyright laws allow fair reuse of all or parts of a work for a particular purpose, but to truly allow research to be rebuilt, interrogated and reshared, a licence that permits the creation and use of derivatives of the original must be in place. Authors are protected from derogatory use and false attribution by copyright norms (and by their moral rights, enshrined in part of the CDPA), and the attribution element of a CC licence means that uses of a work can be tracked and monitored. It remains the case that some academic authors prefer the use of more restrictive licences, as found in a 2014 study,¹⁵ particularly in the arts and humanities. Misunderstandings about the use and applications of CC licences will contribute to this perception. Additionally, CC licences are legal contracts and have been used successfully in courts to defend the rights of creators.¹⁶

‘Authors are protected from derogatory use and false attribution’

Elsevier, UKRI and Creative Commons – the policy stack in action

To highlight the tension between copyright ownership, funder policy and publisher policy, I have chosen to focus on the set of licences offered by Elsevier, the largest academic publisher in the world.¹⁷ Elsevier’s policy and interpretation of Creative Commons licences perfectly encapsulate the confusion and tensions of ‘openness’ in the current scholarly publishing ecosystem.

‘the confusion and tensions of ‘openness’ in the current scholarly publishing ecosystem’

Elsevier offer four licences: CC BY, CC BY-NC-SA (Share-Alike: any new work must be shared under the same terms), CC BY-NC-ND, and its own ‘Elsevier user licence’. For gold OA, a choice of a commercial or non-commercial CC licence is offered, depending on the journal – although the non-commercial element would not satisfy the requirements of many

4 funders, including UKRI. The green route, however, is particularly problematic: there is no way to publish with Elsevier and remain compliant with current UKRI requirements.

Its policy states, 'We support green open access ... accepted manuscripts can be self-archived following our sharing guidelines and are required to attach a CC-BY-NC-ND license'.¹⁸ The current UKRI FAQs explicitly state that any licence with an ND element is not acceptable for the green route. The official policy since 2013 is more ambiguous. This requires that the manuscript is made available without restriction on non-commercial reuse, and does not specify a particular licence save that the requirement can be met by use of the CC BY-NC licence, and 'publisher-specific licences are acceptable providing they support the aims of the policy, and allow re-use including non-commercial text and data mining'.¹⁹ This appears to preclude more restrictive licences, such as CC BY-NC-ND.²⁰ Additionally, the RCUK/UKRI policy has always called for the allowance of text and data mining (TDM).

Copyright, TDM and derivative use

A simple definition of text and data mining is: 'electronic analysis of large amounts of copyright works [allowing] researchers to discover patterns, trends and other useful information that cannot be detected through usual 'human' reading'.²¹ It can involve statistical, numerical or linguistic analysis. Therefore, its use potentially applies to a wide range of disciplines from humanities through to hard science. Essentially, it involves machine reading, and a maxim used by many proponents of TDM is that 'the right to read is the right to mine'.

'TDM is our future, but we have not accommodated it in our laws in a perfect way, yet'

In the context of OA, however, this maxim is less helpful. TDM does not just involve machine reading. It involves analysing, annotating and producing new works that encapsulate and develop the knowledge contained within the mined works. Future research is contingent on successful applications of AI (artificial intelligence) technologies: TDM is our future, but we have not accommodated it in our laws in a perfect way, yet. What we have is a series of imperfect solutions that do not always match up with actual needs.

TDM activity can involve some or all of the following rights: accessing text (right to make public); downloading text (right to copy); reformatting text (right to translate/make derivatives); storing text (copy); producing outputs containing text (make derivatives). UK law permits copying of a work for non-commercial computational analysis provided that the person performing the analysis has lawful access to the work. Transferring the copy to another person may not be done save to fulfil the same purpose of performing a non-commercial computational analysis.²² Licences cannot restrict the right to do this – but, crucially, they can enable. Article 3 of the EU's proposed Directive on Copyright in the Digital Single Market contains an exception for TDM, which is similar to the UK exception, but more limiting because it proposes to restrict TDM activity to research institutions, and for the purposes of scientific research.²³

The sharing of derivative works is explicitly disallowed under the ND flavour of licence.²⁴ Therefore, even where works are shared under CC BY-NC-ND, sharing derivative works is not allowed by law or by licence; the CC licence is patently unable to fill the gap in the law. Current UKRI policy is clear: the green OA route must use a licence which places no restriction on non-commercial use, and 'allow[s] for the sharing of adaptations of the material'. This includes remixed and transformed works, and the products of text and data mining. UK TDM law is problematic in other ways; it does not permit its use for commercial purposes, but the definition of 'commercial' can pose problems. For example, would publication of the results of TDM in a charged-for journal or conference proceedings be considered commercial? Furthermore, current UK TDM law does not prevent a publisher putting a stop to what would be otherwise lawful TDM activities if it believes such activities will lead to degradation of its services. As a result, a number of higher education libraries have taken an unnecessarily cautious approach by stating that patrons should not use materials that the library has access to for large-scale TDM purposes.

5 Returning to author choice and rights, it is worth noting that the restrictions imposed by a licence will often affect the author. The green route to OA is far more likely to involve a copyright assignment than the gold route. Publishing contracts will sometimes give rights 'back' to the author. Aside from this variable and complicated solution, they lose the rights to perform acts normally allowed by the copyright holder, including the right to rework and reshare. True OA licences are therefore a method by which authors retain rights to reuse their own work, including their moral rights as author. Ideally, this is balanced with a measure of autonomy for authors to choose their licence. It is incumbent on institutions, publishers and funders alike to provide clear and consistent guidance and advice regarding open licences so that choices are made from a position of knowledge and understanding of the potential impact.

Conclusion

Misunderstandings about Creative Commons licences abound. This is due in part to misinformation, and leads to tensions between publishing and funding policy. It is also due to evolving technologies and legal concepts. UK copyright law underwent significant updates in 2014, yet it has not caught up with many of these technological innovations and applications. Strong laws are a vital bedrock yet they are notoriously slow to update, so licences must necessarily play a critical part in filling the gaps.

The relationship between academics and UK HEIs is usually constructed so that many academic authors have freedom to transfer copyright and/or to license their works. It is, however, increasingly difficult to balance the competing demands of funder and publisher policy, particularly when misinformation about rights and licensing abounds.

Money speaks louder than ownership – for now. When publishers receive research funding in the form of APCs, then there is an incentive to adapt practices accordingly. Paid-for gold OA is oftentimes the most straightforward route to comply with funder mandates and publisher policies. Pay an APC and you can publish your article immediately, with the version of record (VoR) available under a CC BY licence – whilst often retaining copyright ownership in the process, and always retaining the rights to reuse your work. For green OA, it is difficult to ignore the correlation between lack of payment and corresponding variation in access to research, including the imposition of various embargo periods by different publishers.

What is needed now, more than ever, is harmonization between funder and publishing policies which give authors, their institutions and readers choice as well as the opportunity for fair access and reuse. The recently announced cOAlition S initiative, to which UKRI is a signatory, states a preference for the CC BY licence, that hybrid publishing models are incompatible with the initiative, and that 'In all cases, the license applied should fulfil the requirements defined by the Berlin Declaration'. Perhaps most critically, the first point is that authors 'retain copyright of their publication with no restrictions'.²⁵ The mandate for open licences has reignited debates about academic freedom, yet full retention of copyright is a positive move for many authors. Further details, and how UKRI will interpret this in their policy-making, remain to be seen. It seems hopeful, however, that this is the beginning of a strong funder/policy mandate for true OA licences – something which will be of enormous benefit to many stakeholders in the scholarly communication ecosystem.

'it is difficult to ignore the correlation between lack of payment and corresponding variation in access to research'

'this is the beginning of a strong funder/policy mandate for true OA licences'

Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other *Insights* articles can be accessed here – click on the URL below and then select the 'full list of industry A&As' link: <http://www.uksg.org/publications#aa>

Competing interests

The author has declared no competing interests.

References

1. Banks C, Focusing upstream: supporting scholarly communication by academics, *Insights*, 2016, 29(1), 37–44; DOI: <https://doi.org/10.1629/uksg.292> (accessed 5 November 2018).
2. Budapest Open Access Initiative, 2001: <https://www.budapestopenaccessinitiative.org/read> (accessed 5 November 2018).
3. Berlin Declaration, 2003: <https://openaccess.mpg.de/Berlin-Declaration> (accessed 5 November 2018).
4. Downloading a 'fair' amount of a work for personal use is permitted in many national copyright laws, including the UK, but the copyright holder or technical provider may choose to protect their content or platform with technological protection measures (TPMs). Again, in many national laws it is illegal to circumvent TPMs, even if it is to achieve a lawful aim, such as utilizing a copyright exception.
5. An introduction to open access, Jisc, 2015: <https://www.jisc.ac.uk/guides/an-introduction-to-open-access> (accessed 5 November 2018).
6. Copyright, Designs and Patents Act 1988, sections 16(1) and 29A: <https://www.legislation.gov.uk/ukpga/1988/48/contents> (accessed 5 November 2018).
7. Gadd E, UK university policy approaches towards the copyright ownership of scholarly works and the future of open access, *ASLIB journal of information management*, 2017, 69(1), 95–114; DOI: <https://doi.org/10.1108/AJIM-06-2016-0092> (accessed 5 November 2018).
8. Gadd E et al., RoMEO studies 1: the impact of copyright ownership on academic author self-archiving, *Journal of Documentation*, 2003, 59(3), 243–277; DOI: <https://doi.org/10.1108/00220410310698239> (accessed 7 November 2018).
9. The author acknowledges that gold open access is not always paid-for open access. Peter Suber's definitions of gratis (price barrier removed) and libre (price and permissions barriers removed) open access are more equitable and more accurate – yet in the 10 years since he proposed these, they have not become the norm in the UK, nor are they commonly understood – because research funding has shaped the open access debate to a disproportionate extent.
10. Jubb M et al., *Monitoring the transition to open access*, 2017, UUK: <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/monitoring-transition-open-access-2017.pdf> (accessed 5 November 2018).
11. RCUK Policy on Open Access and Supporting Guidelines: <https://www.ukri.org/files/legacy/documents/rcukopenaccesspolicy-pdf/> (accessed 5 November 2018).
12. Policy for open access in the post-2014 Research Excellence Framework: Updated July 2015, HEFCE: <http://www.hefce.ac.uk/pubs/year/2014/201407/> (accessed 5 November 2018).
13. RCUK, ref. 11.
14. RCUK Policy on Open Access Frequently Asked Questions: <https://www.ukri.org/files/funding/oa/oa-faqs-pdf/> (accessed 5 November 2018).
15. Taylor & Francis, Open Access Survey, 2014: <https://www.tandf.co.uk/journals/explore/open-access-survey-june2014.pdf> (accessed 5 November 2018).
16. Creative Commons: https://wiki.creativecommons.org/wiki/Curry_v._Audax (accessed 5 November 2018).
17. Larivière V, Haustein S and Mongeon P, The Oligopoly of Academic Publishers in the Digital Era, *PLOS ONE*, 2015, 10(6); DOI: <https://doi.org/10.1371/journal.pone.0127502> (accessed 5 November 2018).
18. Elsevier – Open access licenses: <https://www.elsevier.com/about/policies/open-access-licenses> (accessed 5 November 2018).
19. RCUK, ref. 11.
20. The Elsevier user licence offers no recourse as it is even more restrictive than the most restrictive CC licence (CC BY-NC-ND) by not permitting self-archiving in a repository, or reusing portions or extracts from the article in other works. This latter requirement attempts to limit rights allowed by many national copyright laws, and as such is legally tenuous. Additionally, the claim that 'Articles published under an Elsevier user license are protected by copyright' is misleading since all articles, including those published under a Creative Commons licence, are protected by copyright.
21. Copyright User – Text & Data Mining: <https://www.copyrightuser.org/understand/exceptions/text-data-mining/> (accessed 4 November 2018).
22. Copyright, Designs and Patents Act 1988, ref. 5.
23. European Commission: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016PC0593> (accessed 5 November 2018).
24. Creative Commons: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode> (accessed 5 November 2018).
25. Science Europe Plan S: https://www.scienceeurope.org/wp-content/uploads/2018/09/Plan_S.pdf (accessed 5 November 2018).

Article copyright: © 2019 Ruth Mallalieu. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use and distribution provided the original author and source are credited.



Ruth Mallalieu
Scholarly Communications Licensing Manager, Copyright & Licensing Team
The University of Sheffield
Western Bank Library, Sheffield S10 2TN, UK
E-mail: r.mallalieu@sheffield.ac.uk | Tel: 0114 222 7347

ORCID ID: <http://orcid.org/0000-0002-0409-6199>

To cite this article:

Mallalieu R, The elusive gold mine? The finer details of Creative Commons licences – and why they really matter, *Insights*, 2019, 32: 1, 1–7; DOI: <https://doi.org/10.1629/uksg.448>

Submitted on 08 October 2018 Accepted on 16 November 2018 Published on 03 January 2019

Published by UKSG in association with Ubiquity Press.