Research England

Sector report
RE-P-2018-05

# Monitoring sector progress towards compliance with funder open access policies

June 2018

Web: http://re.ukri.org
Twitter: @ResEngland
Email: Subscribe to email alerts

## Acknowledgements

This project was jointly commissioned by the former Higher Education Funding Council for England (HEFCE), the former Research Councils UK (RCUK), Jisc, and Wellcome in 2017. This project was carried out between summer and autumn 2017, prior to HEFCE's transition to Research England, and before the formal establishment of UK Research and Innovation in April 2018. References to HEFCE and RCUK in the report are to be understood within this context.

The project partners would like to thank the universities who responded to the open access initial pilot exercise, carried out by Research Consulting in August 2017. We would also like to thank the institutions that participated in the survey circulated to the sector during summer 2017.

We would like to thank members of the project steering group for their feedback and guidance at all stages in this work: Hannah Hope (Wellcome), Angela Hilton, Bill Hubbard, Sarah Fahmy and Helen Blanchett (Jisc), Geraldine Clement-Stoneham (representing the former RCUK), and Steven Hill and Claire Fraser (Research England). Rachel Reynolds (UK Research and Innovation) has also provided some invaluable advice.

We would especially like to extend our thanks to Rob Johnson and Andrea Chiarelli at Research Consulting for developing and piloting the survey. We would also like to thank Kim Hackett (Research England), Richard Niblett and Ashley Moore (RCUK) for their contributions to the project.

Authors: Claire Fraser, Geoff Hill, Helen Snaith and Jade Taffs.

# Contents

# Executive Summary

During August 2017, the former HEFCE, RCUK, Jisc and Wellcome surveyed Higher Education Institutions about their compliance with, and experience of funder open access policies. The aim of the work was to investigate progress towards our shared goal of increasing the extent to which the outputs of publicly-funded research are freely and openly available. This report details the findings from the survey responses, and represents a snap-shot of the progress towards open access. The report should be read in the context of other work that examines the extent to which scholarly content is openly available in the UK and beyond, especially the report commissioned by the Universities UK Open Access Co-ordination Group.[1]

The publication of this report coincides with the open access reviews of major UK funding bodies. Both Wellcome and the newly-established UK Research and Innovation (UKRI) are conducting an internal review of their open access policies. The evidence put forward in this report will be considered as part of both reviews.

The survey and its findings focus on policy and implementation issues. Specifically, the report examines:

- The open access policy for the 2021 Research Excellence Framework (REF), which has applied to journal articles and conference proceedings made first publicly available since 1 April 2016;
- The RCUK open access policy, which applies to outputs resulting from funding from the Research Councils, and has applied in its current form since 1 April 2013, and;
- The Wellcome/Charity's Open Access Fund (COAF) open access policy, which applies to outputs resulting from research charity funding, and has applied in its current form since 1 October 2014.

It is therefore important to read the findings in the context of the maturity of the policy environment, which differs between the policies.

The following is a summary of the key findings.

1. **There has been significant progress towards meeting the REF open access policy requirements.** Although the survey period covered only the first year of the REF

[1] 'Monitoring the transition to open access' (2017) - report to the UUK Open Access Co-ordination Group. http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/monitoring-transition-open-access-2017.aspx

open access policy, over 80 per cent of the outputs covered meet the policy requirements, or an exception is known to apply (Table 9). This represents an enormous achievement, contributing to increased access and use of UK research.

2. **Institutions are making open their most significant and important research outputs, although there are disciplinary differences in the types of output that are generally selected for assessment.** While the focus of REF has been on both journal articles and conference proceedings, it is important to recognise that there are disciplinary differences in the importance of different types of output. There is evidence that institutions are not making conference proceedings open in those disciplines where they are less likely to be submitted for assessment (Paragraph 62).

3. **There are a number of reasons why 19 per cent of outputs are not meeting the REF 2021 policy requirements, or an exception is not known to apply**. Disciplinary differences and types of output returned to the REF is just one reason: we recognise that some tools do not recognize the flexibility of the REF 2021 OA policy and that some institutions are currently not tracking exceptions. It is important to remember that this report captures in-scope outputs in the first year of the policy, and that not all of these outputs will be returned to the next REF exercise. We hope universities will consider the information presented in this report as part of their own benchmarking activities.

4. **The use of immediate open access is a significant aspect of the change, especially in the science disciplines.** Making research outputs available open access immediately on publication, generally with extensive re-use permissions, is being used extensively (Table 17). This approach is encouraged by the RCUK and COAF policies and removes the need to comply with the deposit requirement of the REF policy. Evidence suggests that this route to open access is being more widely adopted in discipline areas covered by the natural science panels of the REF.

5. **Where funding is required for immediate access, a significant proportion comes from the dedicated funds provided by RCUK and COAF.** Immediate access often requires funding in the form of an Article Processing Charge (APC) or equivalent, and the survey confirms the reliance on dedicated open access funding from RCUK, Wellcome and COAF in meeting this funding requirement (Table 19).

6. **There is variation in the extent to which institutions are complying with the REF OA policy, but this does not necessarily relate to levels of research intensity.** While the overall level of compliance with REF policy requirements is over 80 per cent, there is considerable variation between the institutions that responded to the survey, varying from a minimum of 32 per cent to a maximum of 100 per cent (Figure 2,

Paragraph 49). Initial findings also indicate that there is no systematic relationship between this variation that is related to Transparent Approach to Costing (TRAC) grouping or reflecting volume of research outputs.[2]

7. **Dedicated staff time is required to deliver open access, but this varies across the sector.** Institutions reported that a substantial part of the effort required for open access comes from staff in roles dedicated specifically to this objective. Across the sector, some 300 full-time equivalent staff are employed in such roles with a significant majority being found in the larger, more research intensive institutions (TRAC group A) (Figure 7). These staff contribute to open access via all routes, and in compliance with all funder policies.

8. **The increased open access to research is resulting from considerable effort on the part of researchers, libraries, research offices.** The achievements detailed in the report are the result of efforts from all parts of the research system within institutions. Despite progress on software systems and tools, and interoperability between them, a significant amount of manual intervention is still required in order to make content open and ensure fully documented compliance with funder and other requirements.

9. **The Jisc Publications Router service is an increasingly important tool for helping institutions to make content open and meet policy requirements.** Among a range of software systems and tools used by institutions, particular importance was placed on the Jisc Publications Router, which provides automated and timely information on accepted and published articles. While the tool was used (at the time of the survey) by 13% of respondents, more than 50% intend to make use of the service in the future (paragraphs 32-34, Table 8) although this figure will depend on the willingness of the Current Research Information System (CRIS) vendors to collaborate technically and on the willingness of publishers to supply information.

This report, alongside other sources, provides important evidence that will be considered by the reviews of open access policy currently being conducted by UKRI and Wellcome. It is also useful for Jisc in planning the development and deployment of services to support the goal of increased access to the outputs of research.

---

[2] The Transparent Approach to Costing (TRAC) is the methodology developed with the higher education sector to help them cost their activities. It is an activity-based costing system adapted to academic culture in a way which also meets the needs of the main public funders of higher education. As part of TRAC institutions are split into six comparator groups (A-F) with similar sized teaching and research intensive universities grouped together.

# Background

1. This project was jointly commissioned by the former Higher Education Funding Council for England (HEFCE), the former Research Councils UK (RCUK), Jisc, and Wellcome to further understand how far the sector is meeting the funders' open access (OA) policies and the tools which are being used to do so. In August 2017 we invited HEIs to participate in a UK-wide survey on the delivery of funders' OA policies. This report contains the analysis of responses to this survey.

2. A collaborative, multi-stakeholder approach was adopted for the project in order to extend, enhance and share knowledge of OA within the higher education (HE) sector. This approach was also adopted in order to reduce the administrative burden, both on project partner organisations and on institutions participating in the study.

3. This project was carried out between summer and autumn 2017, prior to HEFCE's transition to Research England and before the formal establishment of UK Research and Innovation in April 2018. Operating across the whole of the UK, UK Research and Innovation brings together the seven Research Councils, Research England, and Innovate UK. Any references to HEFCE and RCUK are to be understood in this context.

4. This report is intended for those with an interest in open access within the higher education sector, including (but not limited to) universities, research librarians, service providers and funders. An abbreviations list and a glossary of terms are located at the end of this report.

# Introduction

5. The project was undertaken to further understand some of the challenges that funders' OA policies may place on higher education institutes (HEIs), and to assess how the sector is progressing with and implementing these policies. HEFCE, RCUK, Jisc and Wellcome were not solely interested in the data relating to overall policy compliance, but the methods and tools being employed across the sector. Project partners were keen to understand how ready the sector is for the Research Excellence Framework (REF) 2021 and the software which underpins repositories, research management and monitoring

grants from RCUK and the Charity Open Access Fund (COAF).[3] To note, the project did not seek to fully assess institutional costs of implementing OA, although Questions 37 and 38 capture some information pertaining to the number of staff engaged in supporting OA.

6. The report was developed in the period running up to and after the release of the Universities UK (UUK) report 'Monitoring the transition to open access' (2017). The UUK report highlights the increasing rate at which the UK's research output is available via OA. This publication therefore seeks to understand how the success in increasing OA outputs is being operationalised, noting where the difficulties and burdens lie, in order to inform the sector and stakeholders where improvements can be made.

7. Open access to research outputs is essential in ensuring that new knowledge is shared, both to stimulate the endeavour of research, and to enable its outputs to be put to effective use by all stakeholders. The intention of the REF 2021, RCUK and COAF OA policies is to ensure that publicly funded research is made freely available to anyone who is able to access online resources.

8. UK Research and Innovation has already signalled its strong support for OA as an essential part of the research system. In February 2018 UK Research and Innovation announced that it will be carrying out an internal review of its OA policies. The review will consider if the UK approach to OA is working, identifying areas of best practice and areas where improvements can be made. UK Research and Innovation's priorities for open access policy will include making the best use of public funding and driving significant progress towards OA in the UK.

9. In March 2018, Wellcome also announced that they will be carrying out a review of their OA policy. The review aims to ensure the policy efficiently delivers OA to the growing volume of Wellcome-funded research publications, while also supporting the wider transition to OA publishing.

10. The quantitative and qualitative data presented in this report will be considered by the UK Research and Innovation and Wellcome as part of their OA reviews scheduled to take place over the next year.

---

[3] The Charity Open Access Fund (COAF) [https://wellcome.ac.uk/funding/managing-grant/charity-open-access-fund] is a partnership between six health research charities, including Wellcome, to enable free and unrestricted access to the published outputs the research partners support. The fund distributes the majority of Wellcome's funding for OA and enables Wellcome to monitor policy compliance.

# Methodology

11. In May 2017 HEFCE, RCUK, Jisc and Wellcome commissioned Research Consulting to develop and pilot an OA assessment template for the UK HE sector. Following a review of the evidence base and an initial consultation with five HEIs, Research Consulting developed a draft assessment template in collaboration with the four project partners. The template was piloted by 18 institutions in early summer 2017.

12. Institutions that participated in the pilot stage were selected via a disproportionate stratified random sampling strategy and given two weeks to complete the exercise. The final report provides an insight into the 18 institutions' progress towards the delivery of OA gathered via qualitative interviews conducted as part of the pilot process. The final report from Research Consulting, containing an analysis of their results, can be found at Annex A.

13. Institutions that participated in the pilot exercise highlighted the value of the assessment, with most agreeing that it was a useful exercise. Gathering evidence required by the survey allowed HEIs to paint a picture of their OA activities and workflows, providing an opportunity to stimulate critical thinking around these processes. Pilot HEIs also noted that the draft survey failed to allow sufficient room for more qualitative comments, and valued the opportunity to provide feedback via interviews with Research Consulting.

14. In response to the findings and recommendations put forward by Research Consulting, the four project partners (with Research Consulting) developed a final OA assessment template. A copy of this template can be found at Annex B.

15. In August 2017, 165 institutions across the UK were invited to complete the online survey 'Monitoring sector progress towards compliance with funder open access policies'. The survey was facilitated by SmartSurvey. Project partners also published a guidance document and a Word template of the survey to assist institutions when gathering data and populating the survey.

16. Institutions were asked to provide information on publications that fell within scope of the REF 2021, COAF and RCUK open access policies from 1 April 2016 to 31 March 2017 ("the period of interest"). Only articles and conference proceedings were in scope for this exercise, and institutions were asked to disregard monographs, book chapters, data and eTheses when responding to questions. We advised that it would take no longer than two person-days to complete the exercise.

17. Comprising both quantitative and qualitative questions, the survey was grouped into seven themes:

    a. Approaches to open access software solutions (Q3-7)

    b. The policy for open access in Research Excellence Framework 2021 (Q8-17)

    c. Recording exceptions to the Policy for open access in REF 2021 (Q18-25)

    d. The RCUK and COAF open access policies (Q26-29)

    e. Publication metadata (Q30-36)

    f. Costs of open access – staffing (Q37-38)

18. The online survey attracted responses from 113 institutions (a 68 per cent response rate). The analytical and research policy directorates at HEFCE worked in coordination with RCUK, Jisc and Wellcome on the analysis of responses. To ensure the safeguarding of sensitive commercial data, institutional responses were anonymised and amalgamated into Transparent Approach to Costing (TRAC) peer groups for further analysis (Table 1). A breakdown of TRAC peer groups (2016-17) is provided at Annex C.

**Table 1 Breakdown of institutional responses by TRAC peer group**

| TRAC peer group | Responses | Invited |
|---|---|---|
| A | 33 | 33 |
| B | 20 | 23 |
| C | 19 | 23 |
| D | 13 | 15 |
| E | 19 | 41 |
| F | 9 | 23 |
| None | 0 | 7 |

19. This report presents the findings from the survey in the following structure, and highlights the corresponding questions throughout:

    a. Approaches to open access software solutions

    b. Meeting the requirements of the REF 2021 open access policy

c. The use of exceptions to the REF 2021 open access policy

d. Deposit requirements for the REF 2021 open access policy

e. Compliance with RCUK and COAF open access policies

f. Managing metadata

g. The institutional cost of staffing open access

20. This report provides an evidence base for institutional compliance with OA in the HE sector. Survey responses have been used to inform the decision taken on deposit requirements for the OA policy for REF 2021, published in November 2017. No further policy recommendations are made in this report.

# Approaches to open access software solutions

(Q3/4/5/6/7)

## Institutional systems

21. Institutions were asked to provide details of their research management solutions in terms of software and function. Forty HEIs (35 per cent) responded that they only use an institutional repository, with a further 3 per cent noting that they only use a current research information system (CRIS). The majority of institutions (58 per cent) use both. A breakdown of institutional responses can be found at Table 2.

22. Sixty-two per cent of HEIs highlighted that they used EPrints (an open-source repository platform) making it the most popular software solution among those who responded to the survey. A number of GuildHE institutions cited the Consortium for Research Excellence, Support and Training (CREST) as their preferred repository. A full list of the types of software solution used by HEIs can be found at Table 3.

23. Respondents were given the opportunity to highlight any other software solutions used that were not listed in the survey. A number of institutions indicated that they have created their own in-house solutions with functions similar to CRIS for research management purposes, including databases and/or spreadsheets. A number of institutions were reported to be in the process of implementing CRIS solutions in the run up to REF 2021.

**Table 2 Research management solutions used by institutions**

| Type of software | Number of institutions | Percentage of institutions |
|---|---|---|
| CRIS only | 3 | 3% |
| Institutional repository only | 40 | 35% |
| Both | 66 | 58% |
| Neither | 4 | 4% |

**Table 3 Types of software solutions used by institutions (all that apply)**

| Software | As a CRIS or equivalent | As an institutional repository | As both CRIS and institutional repository | Not used |
|---|---|---|---|---|
| Converis | 6% | 0% | 0% | 94% |
| DSpace | 0% | 20% | 0% | 80% |
| EPrints | 0% | 58% | 4% | 38% |
| Samvera (Hydra) | 0% | 3% | 0% | 97% |
| Pure | 13% | 0% | 14% | 73% |
| Symplectic | 19% | 0% | 0% | 81% |
| Worktribe | 4% | 0% | 1% | 96% |

24. Institutions were asked to report on research management solutions for tracking article processing charges (APCs). Results are shown at Table 4. Twenty-five respondents reported that the question was not applicable to their institution. We expect that this may be due to a negligible volume of APC transactions. Twenty-six institutions (23 per cent of all respondents) reported using one system to track APCs. Forty-one institutions (36 per cent) reported using two software solutions, with 17 HEIs (15 per cent) using three solutions.

25. In-house databases (79 responses) and institutional finance systems (56 responses) emerged as the two most popular solutions to track APCs. As institutions were asked to indicate all systems and software solutions used for this purpose, there will be a certain degree of overlap between these two responses.

26. Twelve institutions use a CRIS in coordination with other platforms: responses from these HEIs (and other respondents) highlighted that their proprietary CRIS was often unable to record APC data satisfactorily.

**Table 4 System(s)/software used for tracking APCs (ticked all that apply)**

| Software | Number of institutions using this software | | | | |
|---|---|---|---|---|---|
| | Only | With one more solution | With two more solutions | With three more solutions | Total |
| CRIS | 0 | 1 | 8 | 3 | 12 |
| In-house database | 21 | 38 | 17 | 3 | 79 |
| Institutional finance system | 3 | 34 | 16 | 3 | 56 |
| Institutional repository | 2 | 4 | 4 | 0 | 10 |
| Jisc Monitor Local | 0 | 5 | 3 | 1 | 9 |
| Other | 0 | 0 | 3 | 2 | 5 |
| No. of institutions | 26 | 41 | 17 | 3 | |

# Tools to track articles covered by funder OA policies

27. Institutions were asked to report on their approaches or tools used to identify publications that fall within the scope of research funder OA policies *on acceptance*. Over three quarters of responses (79 per cent) reported using self-service recording or deposit of accepted manuscripts by authors in institutional systems. Seventy-one per cent of HEIs also reported that author-accepted manuscripts are deposited by a member of staff from professional services. Institutional responses are located at Table 5.

**Table 5 Tools used to track funder policies**

| Approach/tool | Currently use | Plan to use in future | No plans to use |
|---|---|---|---|
| Notification of accepted manuscripts by authors to central administration or departmental staff | 71% | 4% | 26% |
| Self-service recording/deposit of accepted manuscripts by authors in institutional systems | 79% | 10% | 12% |
| Jisc Publications Router | 13% | 57% | 30% |
| Direct notification from publisher to institution | 22% | 23% | 55% |
| Review of pre-print archives | 6% | 8% | 86% |
| Other | 9% | 2% | 89% |

28. Institutions were also asked to report on their approach to tools used to identify publications that fell within the scope of research funder OA policies on publication. Scopus (67 per cent), Web of Science (59 per cent) and EuropePMC (39 per cent)

emerged as the top three tools used for this purpose. Alongside the options provided in the survey, respondents also highlighted a number of other widely utilised sources including; PubMed, DBLP, RePEc, SSRN, Inspirs-HEP, EBSCO, DOAJ, Embase, Espacenet, Microsoft Academic Search, and publisher database email alerts. A breakdown of institutional responses is presented in Table 6.

**Table 6 Tools used to identify publications that fall within the scope of research funder OA policies on publication delineated as either manual or automatic ingestion process, or both (all that apply)**

| Tool/data source | Number of institutions | | | Percentage using tool |
|---|---|---|---|---|
| | Via native interface only | Via CRIS or repository only | Via both | |
| ArXiv | 5 | 26 | 2 | 29% |
| Crossref | 11 | 22 | 2 | 31% |
| EuropePMC | 17 | 22 | 5 | 39% |
| Jisc Publications Router | 3 | 13 | 0 | 14% |
| Gateway to Research | 15 | 0 | 0 | 13% |
| Google Scholar | 24 | 3 | 0 | 24% |
| ORCID | 7 | 17 | 0 | 21% |
| Researchfish | 25 | 0 | 0 | 22% |
| Scopus | 39 | 28 | 9 | 67% |
| Web of Science | 36 | 25 | 6 | 59% |
| Other | 15 | 8 | 0 | 20% |

29. Scopus and Web of Science are the two most utilised sources used to identify publications that fall within the scope of OA policies. However, the significant overlap between these two sources was noted by respondents (estimated by one institution to be 90 per cent of outputs). More HEIs identify publications using Scopus or the Web of Science using a native interface (including manual process) than automatically via a CRIS or repository. Institutions also noted that the resource, time and technical expertise required to handle and understand application programming interfaces (APIs) and their compatibility with CRIS solutions were, or could be, prohibitive. Free text comments also revealed a number of institutions were not aware that some publishers were able to provide direct notification when an article has been accepted.

30. Most institutions reported using tools via a native interface (manual) more frequently than via a CRIS or repository (automatic). In free text responses, a handful of institutions reported that they also rely on authors to self-archive or notify professional services when an output is published.

31. Institutions were asked which of the listed third party tools and data sources they use to monitor whether publications are made available in accordance with funders' policies (Table 7). Almost all respondents (93 per cent) regularly used the online SHERPA/RoMeo database to check compliance. The other two most frequently utilised tools were Sherpa/FACT and the Repository REF compliance checker. Other tools noted by institutions (but not listed in the question) included Wellcome's compliance checker and Symplectic's OA Monitor.

**Table 7 Tools used to monitor whether or not publications are made available in accordance to funders' policies**

| | We use this tool/data source regularly | We use this tool/data source occasionally | We are aware of this but don't use it | We were not previously aware of this | No response |
|---|---|---|---|---|---|
| CORE | 0% | 10% | 71% | 8% | 12% |
| Crossref | 16% | 13% | 57% | 3% | 12% |
| Directory of Open Access Journals (DOAJ) | 14% | 27% | 45% | 4% | 9% |
| EuropePMC | 28% | 18% | 35% | 10% | 9% |
| Gateway to research | 7% | 21% | 41% | 19% | 12% |
| Jisc Monitor Local | 4% | 1% | 68% | 12% | 15% |
| Lantern | 2% | 12% | 29% | 45% | 12% |
| OADOI/Unpaywall | 4% | 12% | 47% | 23% | 14% |
| OpenAIRE | 2% | 13% | 65% | 8% | 12% |
| Open Access Button | 3% | 13% | 58% | 13% | 13% |
| Researchfish | 11% | 25% | 50% | 4% | 12% |
| SHERPA/REF | 22% | 26% | 40% | 3% | 10% |
| SHERPA/FACT | 44% | 25% | 21% | 4% | 6% |
| SHERPA/RoMEO | 93% | 5% | 2% | 0% | 0% |
| Repository REF compliance checker | 45% | 10% | 29% | 7% | 9% |
| Other | 19% | 6% | 1% | 4% | 70% |

# Jisc Publications Router

32. At the time of the survey, 13 per cent of HEIs used the Jisc Publications Router to identify publications within the scope of research funder OA policies on acceptance. Over half of remaining respondents (57 per cent) indicated that they plan to use the tool in the future. Table 8 provides a breakdown of institutional use of the Jisc Publications Router by TRAC peer group (see Annex C). The category 'No plans to use' also includes the seven institutions that did not provide a response to this question. Please see Table 1 for TRAC peer group response rates and the total number of institutions per group.

**Table 8 Use of Jisc Publications Router by TRAC peer group**

| Is Jisc Publications Router used? | TRAC peer group | | | | | | |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | Total |
| Currently use | 3 | 0 | 4 | 3 | 5 | 0 | 15 |
| Plan to use in future | 23 | 15 | 10 | 8 | 6 | 2 | 64 |
| No plans to use | 7 | 5 | 5 | 2 | 8 | 7 | 34 |

33. Although under 10 per cent of TRAC group A currently use the Jisc Publications Router, almost 70 per cent of institutions in this group indicated that they plan to use this resource in the future. Three quarters of respondents in TRAC group B (65 per cent of the total number of institutions in the group) also responded in this way. The chief obstacle to take-up for this group is almost certainly the fact that the vendors of the main CRISs have not yet implemented interoperability with the Publications Router. Many institutions in these groups use such a system as part of their repository workflow.

34. Almost 85 per cent of respondents in TRAC group D reported that they currently use the Jisc Publications Router, or they have plans to use it in the future. Fifty-eight per cent of respondents (11 institutions) in TRAC group E also reported these responses; however it should be noted that a low number of responses were received from group E to the survey (19 out of a possible 44).

## Summary

35. Alongside their use of multiple software solutions and tools for individual tasks, institutional reliance on manual input was a persistent theme running through survey responses. This was particularly prominent with regards to tracking APCs, the approaches taken to depositing outputs on acceptance **or** publication and checking compliance with funder policies. Survey responses highlighted:

    a. Manually depositing authors' accepted manuscripts (AAMs) is resource-intensive. Many institutions put this down to publishers either not informing institutions (as opposed to authors) when a manuscript has been accepted, or not providing the relevant metadata via APIs for software to automatically ingest.

    b. The methods used to deposit AAMs and check versions of record (VORs) are wide-ranging. Managing the number of online tools and resources (for example SHERPA/FACT, SHERPA/RoMEO) alongside the management of journal feeds and emails is also resource-intensive. The majority of institutions rely on professional services staff to deposit AAMs.

    c. The Jisc Publications Router is seen as a helpful addition for notifying institutions of AAMs. Respondents repeatedly recommended that publishers engage with this service. As one institution observed "The [Jisc] Publications Router would be very beneficial to researchers/workflows if manuscripts as well as metadata were routinely supplied by publishers."

    d. Responses highlighted the significant number of systems and software solutions used by HEIs to track APCs, with most institutions using more than one solution to do this.

    e. Arts and design specialist institutions who responded to the survey highlighted that their outputs are practice based and these tools and sources do not adequately reflect their requirements. There were calls for more engagement on this from software developers, funders and Jisc.

    f. The "complex nuances of funder policy" as well the variety of publishers' embargo periods caused institutions to extensively cross-check the results of REF Compliance Checker (particularly) and other compliance tools manually. This was usually checked by emailing journals/publishers or checking their websites.

# Meeting the requirements of the Research Excellence Framework 2021 open access policy

(Q8/9/10/17/36)

## Compliance with the Research Excellence Framework 2021 open access policy

36. Project partners were keen to understand the extent to which universities are meeting the requirements of the REF 2021 OA policy. It should be noted that research outputs are deemed to meet the access requirements during an embargo period, provided that the length of the embargo period is not greater than the maxima permitted by the policy. It should also be noted that not all outputs reported in this survey will be submitted to the REF 2021. Therefore compliance levels are indicative of sector progress in general and do not reflect institutional submissions to REF 2021.

37. Institutions were asked to supply the total number of known research outputs that fell within the scope of the REF 2021 OA policy during the 'period of interest' against the following categories:

    a. How many outputs are believed to have met, or are expected to meet, the deposit, discovery and access requirements?

    b. How many outputs have a known exception?

    c. How many outputs do not yet meet the deposit, discovery and access requirements?

    d. The total number of in-scope outputs.

38. Wherever possible, institutions were asked to present the data by REF main panels (A, B, C and D) as well as providing total numbers. The total number of known research outputs in scope of the REF 2021 OA policy is shown in Table 9. A breakdown of these outputs by REF main panels is provided at Figure 1.

39. Over 85 per cent of institutions reported total output numbers. Institutions who left an entire category blank (as listed in paragraph 37), or responded 'unknown' were removed from the analysis. Where an institution provided figures by panel, but no total, the sum of the panel figures is used as the total in this analysis. Where numbers were reported by

panel, but the sum of these did not reflect the totals reported, the totals provided were used regardless of the discrepancy.

**Table 9 Total outputs known to be in scope from 1 April 2016 to 31 March 2017 by category of compliance**

| Category of outputs | Total outputs | Percentage of outputs |
|---|---|---|
| Meet or expected to meet deposit, discovery and access requirements | 82,192 | 61% |
| An exception is known to apply | 27,101 | 20% |
| Do not meet deposit, discovery and access requirements and not known to be covered by an exception | 26,160 | 19% |

40. Table 9 shows that over 80 per cent of outputs in the scope of the REF OA policy either met the REF policy requirements in the first year (1 April 2016 to 31 March 2017), or an exception to the policy requirement is known to have applied.

41. Fewer institutions reported the figures by panel (Main Panel A, 68 per cent; Main Panel B, 59 per cent; Main Panel C, 66 per cent; and Main Panel D, 70 per cent). Institutions who did not record a response next to a category have been removed from this analysis. Institutions that gave a partial response (for example, reporting figures in some, but not all of the categories) have also been removed from the final figures used. These final figures have been used to inform the data used in Figure 1.

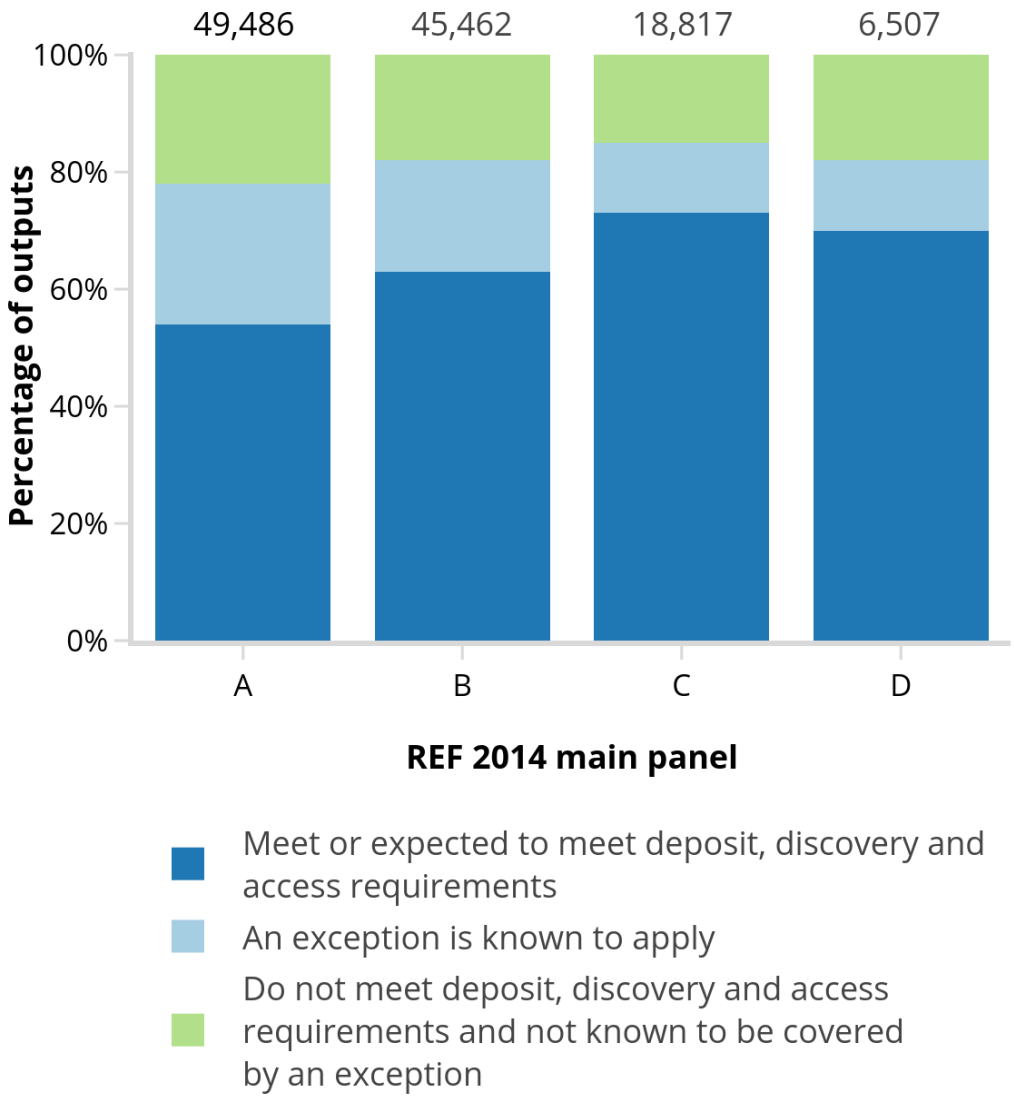
42. The number of total outputs with a known exception which were reported in this section of the survey do not match those reported to the question on exceptions (27,101 compared to 28,984). There are a number of reasons for the discrepancy in the data, which could include:

    a. Analysing a different number or set of institutions in each of the questions;

    b. Institutions may have information by REF main panel, but not by exception type (or vice versa);

    c. Question 8 in the survey asked for known estimates, whereas Question 10 asked for an estimate.

**Figure 1 Percentage of outputs by main panel and category of compliance**




43. Figure 1 shows that Main Panel A has the highest percentage of outputs that do not meet the deposit, discovery and access requirements, and not known to be covered by an exception (22 per cent, or just under 10,800 outputs). It has the lowest percentage of outputs that are known to have met the requirements for REF 2021 (54 per cent, approximately 26,600 outputs) across all Main Panels. It also has the largest number of exceptions at 24 per cent (just over 12,000 outputs). This may be accounted for by the large number of exceptions available through Gold OA. As the section 'The use of exceptions to the REF policy' explores, deposit exceptions where an output has been published under Gold OA was mostly recorded in subject areas aligned with REF Main Panels A and B (see paragraph 76).

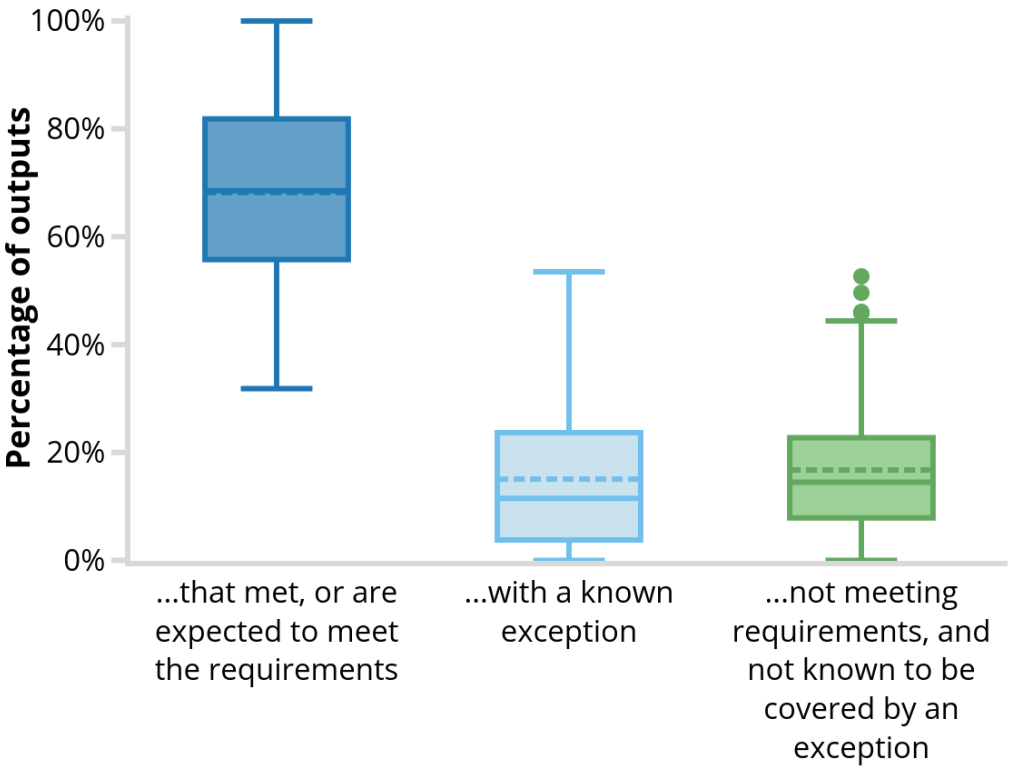
44. Figure 1 also shows that an average of 63 per cent of outputs in Main Panel B are compliant with the REF 2021 OA policy (almost 28,500 from a total of 45,462 outputs). Nineteen per cent (just under 8,800 outputs) are known to have exceptions; as noted in paragraph 76, these exceptions could be accounted for by the number of outputs made available through Gold OA. The remaining 18 per cent of outputs (over 8,100 outputs) do not meet deposit, discovery and access requirements, or an exception was not known to apply.

45. Main Panels C and D reported a higher average of outputs which met or are expected to meet the requirements at 73 (over 13,800 outputs) and 70 per cent (approximately 4,500 outputs) respectively. It should be noted that responses reported a significantly lower number of outputs aligned with Panels C and D in comparison to Main Panels A and B (see Figure 1). Panels C and D also recorded an average of 12 per cent of exceptions to the REF OA policy (over 2,200 outputs for Panel C and just under 800 for Panel D), 10 per cent lower than the average number of known exceptions recorded for Main Panel A outputs. As noted in paragraph 80, lower recorded exceptions of Gold OA across Main Panels C and D may account for this difference. Respondents suggested that there are a higher number of access exceptions due to the inclusion of third party rights in Panels C and D.

46. The average number of outputs not meeting the deposit, access and discovery requirements and not covered by an exception is 15 per cent (around 2,800 outputs) and 18 per cent (just under 2,000 outputs) across Panels C and D respectively.

47. Reporting in paragraphs 42-46 presents the averages across all respondents. For comparison, Figure 2 below shows the spread of responses that reported output compliance with the REF OA policy. The same categories referred to in paragraph 37 have been used for this analysis.

**Figure 2 Percentage of outputs in scope of REF 2021 policy, within the period of interest, and by category of compliance**




48. The top and bottom of each box mark the boundaries of the upper and lower quartiles. Solid and dashed lines that appear across the centre of each of the three boxes show the median and mean values respectively. Vertical lines show the maximum and minimum values, excluding outliers, which are plotted as separate data points. Please note that this figure is based on data from Table 9 rather than panel data shown at Figure 1.

49. For each of the respective categories, survey responses show:

    a. **Met, or expected to meet the requirements.** Half of institutions surveyed indicated that 56 to 82 per cent of in-scope outputs are expected to meet REF 2021 OA requirements. That the highest and lowest values are 100 and 32 per cent respectively demonstrates varying institutional progress. Both measures of central tendency in this category fall around 69 per cent.

    b. **Outputs with a known exception.** Half of institutions surveyed are reporting that five to 22 per cent of in-scope outputs are known to have an exception attributed.

Responses here are skewed towards lower values, reflected by a difference in the mean and median values, with the median number of in-scope outputs with an attributed exception being 10 per cent. While lying within an expected range, one institution reported that 55 per cent of outputs have a known exception.

c. **Do not meet the requirements and not known to be covered by an exception.** Half of institutions surveyed reported that eight to 20 per cent (a narrower range than previously) of in-scope outputs neither meet the OA requirements nor are covered by an exception. Responses are skewed to lower values of non-compliant outputs, reflected in a slightly lower median than mean value. Three outliers are present beyond the uppermost expected range of values (43 per cent).

50. A breakdown of the percentage of outputs that meet the REF 2021 OA policy requirements by TRAC peer group is shown at Table 10. The results show that TRAC group A reported the highest volume of research outputs and the highest percentage of outputs (23 per cent) with a known exception to the REF 2021 OA policy.

51. There is some variation in the number of institutions per peer TRAC group that responded to the survey and the number of institutions included in the response to this question. For example, 33 institutions in TRAC group A responded to the survey and 30 are included in the Table 10 below; nine institutions from peer group F (out of a total 23) responded to the survey, with six responses included in Table 10. Given the variation in the number of institutions from different TRAC peer groups responding to the question, as well as the variation in the volume of outputs per peer group (particularly in reference to TRAC group F) this data may not be robust enough to draw solid conclusions from. However, there appears to be no systematic relationship between the volume of outputs produced by a TRAC peer group and policy compliance. The extent to which individual TRAC groups are meeting the policy requirements of the REF and have a known exception to apply aligns with overall sector compliance.

**Table 10 Percentage of outputs in scope of the REF 2021 OA policy according to TRAC peer group**

| Percentage of outputs | | | | |
|---|---|---|---|---|
| TRAC peer group | Total outputs | Meet or expected to meet requirements | An exception is known to apply | Do not meet requirements and not known to be covered by an exception |
| A | 96,201 | 56% | 23% | 21% |
| B | 21,044 | 71% | 13% | 16% |
| C | 8,872 | 69% | 14% | 18% |
| D | 5,015 | 74% | 10% | 16% |
| E | 4,137 | 75% | 9% | 16% |
| F | 184 | 77% | 7% | 17% |

52. Institutions were able to provide additional qualitative information relevant to their responses to this section. Institutions reported concerns with the data provided, largely reflecting that they expect the number of compliant outputs (whether meeting the requirements, or an exception to the policy) to increase when the data issues are ironed out. Respondents reported issues with tools, availability of metadata, and institutional practices.

53. Respondents noted that systems and tools need to be improved in order to increase compliance with the REF 2021 OA policy. Specifically, institutions noted that:

    a. Some systems are unable to differentiate between conference proceedings that are in scope, and those that are out of scope. This is discussed in further detail at paragraphs 59-62.

    b. Some tools only report compliance when an output meets deposit requirements within three months of acceptance, rather than publication. Some systems do not consider the flexibility in the REF policy from 1 April 2016 to 1 April 2018. A number of institutions reported that they are working with system providers to rectify this.

    c. HEIs must rely on native tools/interfaces. The report's section on 'Approaches to open access software solutions' (paragraphs 21-35) discusses this in further detail.

54. Limited metadata also impacts upon the deposit processes and determining whether they fall in the ‘met, or are expected to meet the requirements’ category. Embargo lengths cannot be set until there is a publication date attached to the output.

55. Some HEIs also noted that institutional processes could be improved to reduce the average number of outputs that are recorded as non-compliant. For example:

    a. Some institutions reported that the use of exceptions will increase, as they are not yet tracking exceptions, and/or are waiting for some clarification from the REF panels on the use of exceptions to the policy.

    b. Some institutions reported that they are tracking the number of outputs by Main Panels A and B combined and Main Panels C and D combined, rather than by each individual panel.

    c. Some institutions reflected that new systems were being implemented and data was difficult to capture at this time.

56. One institution stated that they have been creating reports on the data held in Scopus and comparing this to data held in their CRIS on a regular basis (every three months). Results are then used to encourage departments to increase the number of outputs deposited in their CRIS. This approach has increased outputs in their CRIS from 65 per cent to 90 per cent.

57. Institutions were asked to estimate the proportion of outputs that fell within the scope of the REF 2021 policy which are currently known to the institution. Responses are presented in Table 11 below. The majority of institutions estimated that they know about 80 per cent of outputs.

58. It was reported in the free text comments that it was difficult to estimate how many outputs were unknown. Some institutions used the previous year’s data to estimate the expected outputs, using tools such as Scopus.

**Table 11 The proportion of outputs which are estimated to be known to the institution**

| Proportion of outputs | Number of institutions | Percentage of institutions |
|---|---|---|
| less than 50% | 2 | 2% |
| 50-60% | 4 | 4% |
| 51-60% | 0 | 0% |
| 61-70% | 8 | 7% |
| 71-80% | 9 | 8% |
| 81-90% | 26 | 23% |
| 91-95% | 21 | 19% |
| 96-100% | 13 | 12% |
| Unable to provide an estimate | 27 | 24% |
| No response | 3 | 3% |

## Conference proceedings: challenges

59. Institutions were asked to report on particular challenges faced with making conference proceedings meet the requirements of the REF 2021 OA policy. It should be noted at this stage that there was an error in the smart survey tool, and the categories 'neither agree nor disagree' and 'agree' were presented in the opposing order to a usual survey format (presented in Smartsurvey as left to right: strongly disagree, disagree, agree, neither agree nor disagree, strongly agree). It would appear that institutions recognised this error or 'switch', as survey responses are commensurate with an overall trend that shows most respondents selected either 'Agree' or 'Disagree'. This should be considered in relation to the responses in Table 11 below.

60. Table 12 captures the challenges of ensuring conference proceedings are compliant with the REF OA policy. Notably, the majority of institutions have encountered difficulties ensuring compliance for this output type, with the most challenging aspects listed as:

    a. Interpreting publisher policies (94 per cent either agree or strongly agree);

    b. Determining the acceptance/ publication date (94 per cent either agree of strongly agree);

    c. Determining if the conference proceeding is subsequently published (84 per cent either agree or disagree).

61. The majority of institutions (66 per cent) also agreed or strongly agreed that it was difficult to determine whether conference proceedings are in scope of the REF policy, and that it was difficult to identify and obtain the AAM. However, over 20 per cent of respondents either disagreed or strongly disagreed with these two statements.

**Table 12 Challenges of ensuring conference proceedings are compliant with the REF OA policy**

| Statement | Percentage of institutions | | | | |
|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree |
| It is difficult to find and interpret publisher policies on self-archiving for conference proceedings | 2% | 0% | 4% | 30% | 64% |
| It is difficult to determine acceptance and/or publication dates for conference proceedings | 1% | 1% | 4% | 25% | 69% |
| It is difficult to determine whether conference proceedings are within the scope of the policy | 4% | 18% | 12% | 31% | 35% |
| It is difficult to determine whether conference papers are subsequently published | 1% | 6% | 9% | 42% | 42% |
| It is difficult to identify and obtain the author's accepted manuscript for conference proceedings | 2% | 17% | 15% | 31% | 35% |

62. Respondents had the option to provide comments, noting any additional challenges or issues faced when making conference proceedings complaint with the REF policy. The following paragraphs outline the key concerns:

    a. **In scope.** Some respondents noted that there were difficulties determining which conference proceedings were in scope for the REF 2021 policy. These include:

        i. Respondents noted that authors categorise a range of output types as conference proceedings, and manual process is required to check whether outputs are in scope of policy requirements. Some respondents advised that CRIS systems are unable to determine between conference proceedings with an ISSN, or without, and are therefore reliant on manual checking. One respondent noted that this differentiation is possible using their systems (Pure Instance). It was suggested that services such as

ZETOC were becoming outdated as smaller conferences move to self-publication.

ii. Institutions stated that uncertainty stems from whether the conference proceeding will be published with an ISSN or an ISBN, therefore determining whether the output is in scope of the policy.

iii. A few respondents highlighted that the REF policy and subsequent FAQs cause confusion about the outputs published in a series with an ISBN, where individual outputs within the series have an ISSN. The respondents sought clarification on this in REF guidance.

b. **Disciplinary differences.** It was noted that some units of assessment (UOAs) give very little weight to conference proceedings as an output type and that this impacts author engagement with this element of the REF policy. A few noted that submission of this output type to REF 2014 was low at only 1.4 per cent, and that there is disproportionate burden in complying with the REF policy for this output type given the numbers of these outputs which are actually submitted to the exercise.

c. **Publisher policies.** Respondents suggested that publisher policies on conference proceedings are unclear. It was highlighted that publishers need to provide the same information for this output type as journal articles, and more emphasis should be placed on better provision of metadata.

d. **Date of acceptance/publication.** It was highlighted by respondents that there is a lack of clarity with both date of acceptance and date of publication for conference proceedings due to the iterative development of the output. Potential time lags between acceptance, publication and conferences were highlighted as adding further complexity.

## Institutional approaches to meeting the requirements of the REF 2021 OA policy

63. The survey sought to understand how institutions are guiding authors on the use of subject repositories. Table 13 below demonstrates that 90 per cent of institutions guide authors to put outputs in the institutional repository, regardless of whether a copy is also held in a subject repository.

**Table 13 Institutional policies on the use of subject repositories**

| Institutional policy | Percentage of institutions |
|---|---|
| All full outputs must be deposited in the institutional repository, regardless of whether a copy is also held in a subject repository | 90% |
| All metadata records must be held in the institutional repository/CRIS, but the full text can be deposited in ANY subject repository | 4% |
| All metadata records must be held in the institutional repository/CRIS, but the full text can be deposited in A DEFINED LIST of subject repositories | 2% |
| Outputs recorded in ANY subject repository do not need to be recorded in the institutional repository/CRIS | 0% |
| Outputs deposited in A DEFINED LIST of subject repositories do not need to be recorded in the institutional repository/CRIS | 0% |
| Left to authors' discretion | 4% |
| Other | 0% |

64. Respondents were asked how or whether they replaced or augmented manuscripts following the release of subsequent versions. In accordance with REF 2021, deposited AAMs may be replaced or augmented with an updated peer-reviewed manuscript or VOR documents at a later date (as long as they meet the access requirements). The results are published at Table 14 below.

65. Responses were reasonably split between opting to replace (37 per cent) or augment (41 per cent) the original manuscript with the updated version, with no clear preference between the two. Respondents who chose 'Other' highlighted that the original manuscript was retained but kept hidden from access either permanently (and only to facilitate audit) or until an embargo period was concluded.

**Table 14 How the original deposited document is handled following the release of subsequent versions**

| Approach | Percentage of institutions |
|---|---|
| The original manuscript is replaced by the updated manuscript | 37% |
| The original manuscript is augmented by the updated manuscript | 41% |
| Not applicable/not sure | 7% |
| Other | 15% |

## Summary

66. Over 80 per cent of outputs in the scope of the REF OA policy either met the REF policy requirements in the first year (1 April 2016 to 31 March 2017), or an exception to the policy requirement is known to have applied. This reflects significant progress toward the policy intent to substantially increase the proportion of research that is made open access in the UK. We expect this benchmarking to be useful for institutions to assess their own progress with the policy requirements.

67. Survey responses demonstrate that there are varying levels of compliance with the REF 2021 OA policy across institutions. We hope that this report serves as a benchmarking exercise for institutional compliance and we encourage the sector to work together in order to share best practice.

68. Responses highlighted a number of issues with conference proceedings relating to interpretation of the REF policy, publisher's policies, and the manual intervention required to determine in-scope outputs. The survey responses also called for better metadata to aid compliance. The responses indicate scope for the policy to be clarified in the detailed REF2021 guidance.

69. Responses and comments on the ability to report on compliance demonstrate the need for CRIS systems, and other compliance tools used by institutions, to be reviewed and updated.

70. There is no clear preference apparent from the sector as to how AAMs are augmented or replaced in repositories following the release of later versions. One institution requested clearer guidance with regard to handling updates.

71. Institutions are guiding authors to deposit in an institutional repository, regardless of policy flexibility. No further questions were asked about the use of subject repositories, and further evidence could be sought to understand this practice.

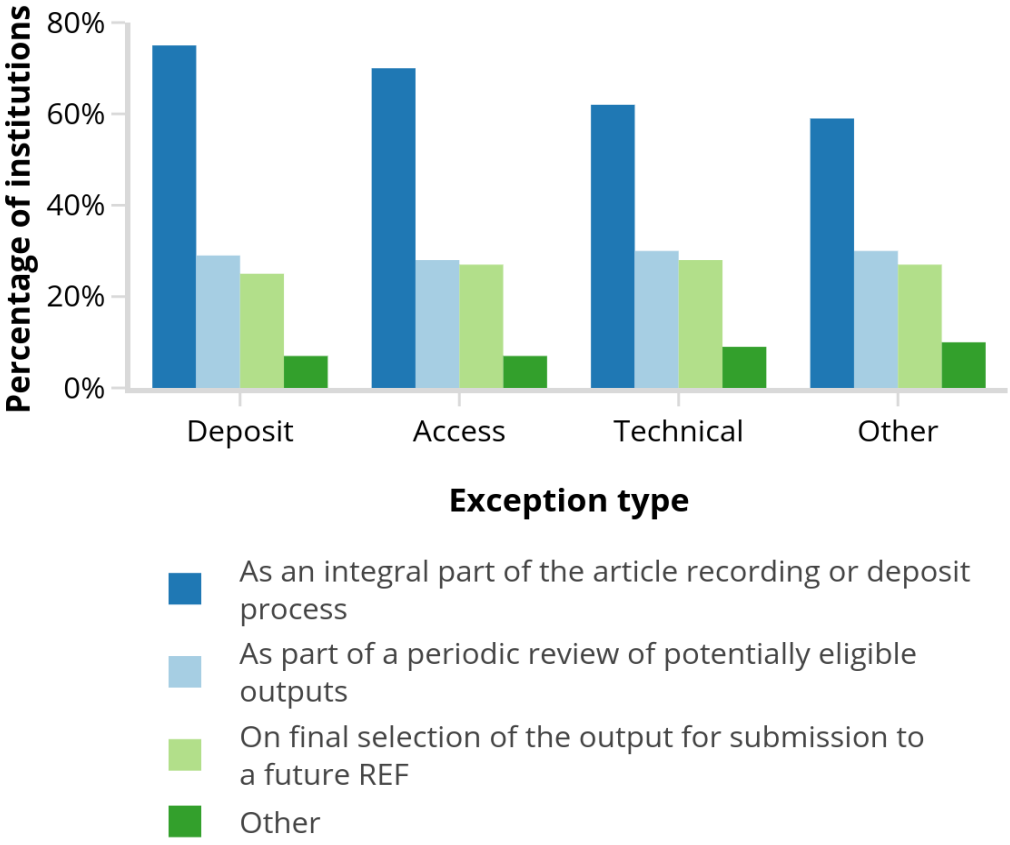
# The use of exceptions to the REF open access policy

(Q18/19/20/21/22/23/24/25)

## Recording exceptions

72. Institutions were asked how they determine and record the application of exceptions to the REF policy (Figure 3). Responses highlight that HEIs are largely tracking exceptions as an integral part of the deposit process. Institutions who provided comments reported that although exceptions are recorded on deposit, they are also subject to periodic internal review as part of the final selection of outputs for REF 2021.

**Figure 3 How institutions determine and record exceptions to the REF OA policy (all options that apply)**



73. Institutions were asked to select the software used to record exceptions for the REF OA policy (see Table 15). The data reflects that a combination of approaches are being used to track policy exceptions. Survey responses show that 41 institutions (36 per cent) use an in-house database (including spreadsheets) to track exceptions. Of these institutions, 15 use in-house databases alongside a CRIS or equivalent, with a further 11 institutions using an in-house database alongside an institutional repository.

74. A cross-reference of survey responses for Questions 3 and 19 highlighted that of the 21 institutions utilising a CRIS platform indicated that they are not recording policy exceptions within these packages.

75. Institutions selecting 'Other' highlighted their use of bespoke systems to record exceptions to the policy. A couple of respondents stated that they are using REF Compliance Checker for these purposes, with one respondent noting that exceptions relating to staff circumstances are recorded outside of the institutional repository for confidentiality.

**Table 15 Software used to record exceptions to the REF OA policy (ticked all that apply)**

| Software | Number of institutions | Percentage of institutions |
|---|---|---|
| CRIS or equivalent | 51 | 45% |
| In-house database (including spreadsheets) | 41 | 36% |
| Institutional repository | 46 | 41% |
| None | 5 | 4% |
| Other | 4 | 4% |

## The extent of exception use

76. Institutions were asked to estimate the use of REF 2021 policy exceptions[4] between 1 April 2016 and 31 March 2017 ('the period of interest'). Institutions who were unable to estimate the use of exceptions were asked to leave the response blank. Where exceptions had not been used within this period, institutions were directed to return '0'.

[4] REF OA policy exceptions can be found at http://www.hefce.ac.uk/rsrch/oa/FAQ/#exeptions10

77. Table 16 below shows the estimated use of policy exceptions for REF 2021. Please note that the number of institutions able to provide an estimate of numbers of exceptions varied by exception type, and therefore this data does not reflect the sector as a whole.

78. The most frequently recorded exception is a 'deposit exception' for outputs published as Gold OA (survey option 'f') accounting for 81 per cent of all exceptions reported. Responses suggest that all exceptions have been recorded at least once, although institutions only reported using technical exceptions 185 times, in comparison to deposit exceptions which have been used over 25,000 times.

79. Respondents were asked to provide observations on the distribution of their current exceptions across REF main panel. Many institutions reflected that it is too early to be able to report on whether there are significant patterns in distribution, due to small and/or incomplete data sets.

80. Respondents highlighted the differences between the type of deposit exception and discipline focus. Significantly:

    a. Deposit exceptions where the output has been published under Gold OA was mostly recorded in disciplines aligned with REF Main Panels A and B.

    b. Respondents suggested that there are a higher number of access exceptions (survey options 'a', 'b' and 'c') across Main Panels C and D due to the inclusion of third-party content.

    c. One institution stated that UOA4 (Psychology, Psychiatry and Neuroscience) has a higher usage of access exceptions that require an embargo period that exceeds REF policy limit (survey option 'b'). They suggested that this is due to a shorter embargo allowance for Main Panels A and B (12 months) in comparison to Main Panels C and D (24 months).

81. Institutions also noted that distribution of REF exceptions by main panels is largely similar to the overall distribution of in-scope outputs.

**Table 16 Estimated use of policy exceptions for outputs to be submitted to REF 2021**

| Exception | | Number of exceptions | Proportion of sector total |
|---|---|---|---|
| **Deposit exceptions** | a. The individual whose output is being submitted to the REF was unable to secure the use of a repository at the point of acceptance. | 95 | 0% |
| | b. The individual whose output is being submitted to the REF experienced a delay in securing the final peer-reviewed text (for instance, where a paper has multiple authors). | 891 | 3% |
| | c. The individual whose output is being submitted to the REF was not employed by a UK HEI at the time of submission for publication. | 529 | 2% |
| | d. It would be unlawful to deposit, or request the deposit of, the output | 22 | 0% |
| | e. Depositing the output would present a security risk. | 5 | 0% |
| | f. The output was published as 'gold' open access (for example, RCUK-funded projects where an open access article processing charge has been paid). | 23,526 | 81% |
| | **Total** | **25,068** | **86%** |
| **Access exceptions** | a. The output depends on the reproduction of third party content for which open access rights could not be granted (either within the specified timescales, or at all). | 101 | 0% |
| | b. The publication concerned requires an embargo period that exceeds the stated maxima, and was the most appropriate publication for the output. | 810 | 3% |
| | c. The publication concerned actively disallows open access deposit in a repository, and was the most appropriate publication for the output. | 686 | 2% |
| | **Total** | **1,597** | **6%** |
| **Technical exceptions** | a. At the point of acceptance, the individual whose output is being submitted to the REF was at a different UK HEI which failed to comply with the criteria. | 79 | 0% |
| | b. The repository experienced a short-term or transient technical failure that prevented compliance with the criteria (this should not apply to systemic issues). | 101 | 0% |
| | c. An external service provider failure prevented compliance (for instance, a subject repository did not enable open access at the end of the embargo period, or a subject repository ceased to operate). | 5 | 0% |
| | **Total** | **185** | **1%** |

| | | | |
|---|---|---|---|
| **Other exceptions** | Other | 2,134 | 7% |
| | **Overall total** | 28,984 | 100% |

## Other exceptions

82. Institutions were asked to provide detail on circumstances in which the 'Other' exception has been used. The most cited reasons include:

    a. Extenuating personal circumstances, including those related to equality and diversity. Mostly this included maternity leave and periods of extended leave;

    b. To track outputs compliant with the policy during the period of flexibility on deposit, but not compliant with policy requirement for deposit within three months of acceptance.[5] Respondents noted that some CRIS and repository systems calculate compliance based on acceptance only. This forces institutions to attach an exception to outputs which fulfil the flexible policy requirements. One respondent noted that this accounts for 91 per cent of their 'Other' exceptions;

    c. Administration error, including: inputting the wrong embargo length, the impact of closure days (including weekends) and the author missing some of the required systematic steps for output deposit;

    d. Difficulty interpreting publisher policy on licensing and embargo lengths. Challenges around licensing are discussed in more detail in paragraphs 105-112; and

    e. Outputs are unlikely to be submitted to REF 2021.

83. Further (but less cited) reasons why institutions used the 'other' exception include:

    a. Author misunderstanding policy requirements;

    b. IT issues which are not covered by the technical exceptions;

    c. Outdated information on SHERPA/RoMEO;

    d. The output is compliant in another repository;

[5] See Decisions on staff and outputs (2017/04). This information is correct as of April 2018.

e. New appointments, specifically when a new member of staff began the publication process at one institution and changed employment after manuscript acceptance but before final publication;

f. Where publication date on an output was inaccurate. Respondents suggested that this was due to different dates being recorded, for example: the early date of publication, the online publication date, and the print publication date. Others noted that publishers have occasionally left the scheduled publication date on the output, with which is then published at a much later date; and

g. As a short term measure until the selection of outputs for REF 2021 have been reviewed.

## Summary

84. Survey responses indicated that institutions have processes in place to record exceptions for submissions to the REF 2021, and that these are tracked and reviewed as an integral part of the deposit process. Free text comments also noted that recorded exceptions are subject to a periodic review as part of the selection for outputs to be returned to the next REF.

85. Although many institutions reflected that it may be too early to report on whether there are significant patterns in distribution due to small or incomplete data sets, the results did reveal some emerging trends. These are:

    a. The most frequently recorded exception is a ‘deposit exception’ for outputs published as Gold OA (for example, RCUK or COAF-funded projects where an OA APC has been paid).

    b. There are disciplinary differences as to why some outputs may have an OA exception under the REF 2021 OA policy. Disciplines aligned with REF Main Panels A and B are more likely to have an exception for outputs published as Gold OA: subject areas more closely aligned with Main Panels C and D, however, reported a higher number of access exceptions.

    c. Approximately a quarter of respondents record REF exceptions in either a CRIS, or a repository and tracking on an in-house database.

86. Institutions have used the ‘Other’ exception for over 2,000 outputs (7 per cent of the sector total). Respondents indicated that the ‘Other’ exception is being applied to outputs where the tools that they are using to track policy compliance do not account for the

flexibility in HEFCE's policy for the first two years. Institutions noted that systems they are using may require software updates to allow them to accurately reflect compliance with the REF 2021 OA policy.

87. Evidence from this survey (including comments received in the 'free text' box) will be used to inform the development of detailed guidance and criteria on the use of exceptions submitted as 'Other'. Main panel criteria will be developed throughout 2018 and will be shared with the sector in due course.

88. As noted in the REF exceptions section above, only one full year of data was available at the time of the survey. However, we are able to see emerging trends in the use of exceptions by panels, noting that the large number of Gold OA exceptions is likely to be reflected in Main Panels A and B.

# Deposit requirements for REF 2021 open access policy

(Q11/12/13/15/16)

## Institutional practices

89. The survey sought information on institutional practice, preference and software approaches to meeting the deposit requirements of the REF 2021 OA policy. A decision was taken by the REF steering group using the evidence gathered and was published in REF decisions on staff and outputs (2017/04) in November 2017. The evidence on which this decision was taken is presented at Annex D.

90. In addition to questions on software approaches (reported in the section 'Approaches to open access software solutions', paragraphs 20 - 34), the survey asked respondents to best describe the institutional policy on the deposit of authors' outputs for the period of interest (from 1 April 2016 to 31 March 2017). Responses are captured in Figure 4. Over a quarter of institutions responded that their policy requires authors to deposit outputs upon acceptance for publication, with a further 60 per cent following requirements for outputs to be deposited within three months of acceptance.

91. Further to this question, respondents were also asked what approach they would like to see implemented in the REF 2021 OA policy from 1 April 2018, assuming that any policy implementation would apply for the remainder of the REF 2021 assessment period. Responses are captured at Figure 5. The majority of institutions (over 70 per cent) indicated that outputs should be deposited no later than three months after the date of publication. Under a quarter of respondents (23 per cent) stated a preference for outputs to be deposited as soon after the point of acceptance as possible, and no later than three months after this date. A handful of institutions did not indicate a policy preference.

92. Institutions were also asked to explain the reasons for their preferred approach. Analysis of the qualitative approaches can be found at Annex D.

**Figure 4 Which of the following best describes the institutional policy on the deposit of author's outputs in place for the period of interest?**

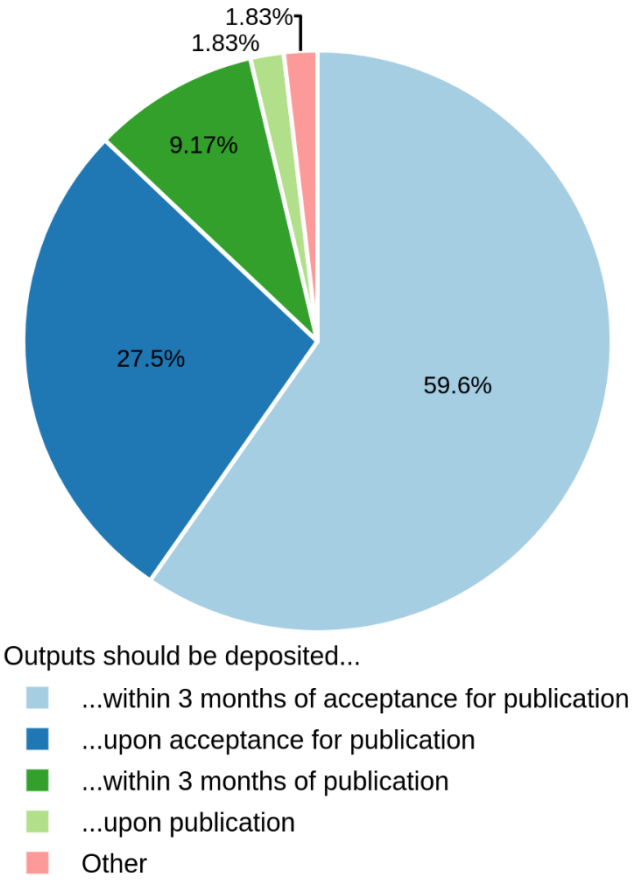
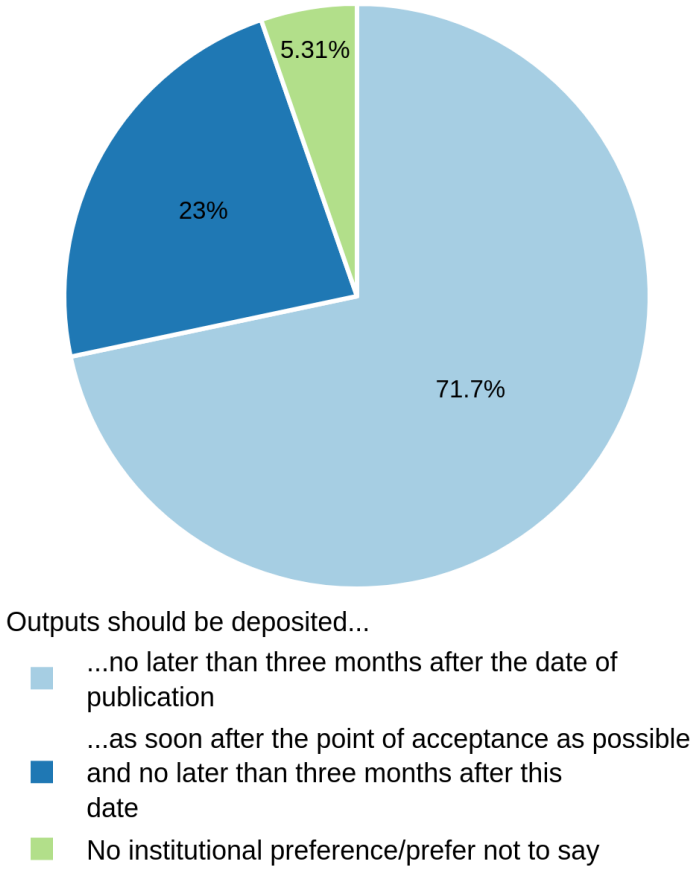



**Figure 5 Which approach would you like to see implemented in the REF OA policy from 1 April 2018?**



## Summary

93. As announced in the REF Staff and Outputs publication (November 2017) the four UK Funding Bodies wish to continue building on the progress achieved to date and to maintain the momentum towards developing new tools to implement deposit as soon after the point of acceptance as possible. It was confirmed in November that the implementation of the REF OA policy will remain as previously set out. The policy will require outputs to be deposited as soon after the point of acceptance as possible, and no later than three months after this date (as given in the acceptance letter or email from the publication to the author) from 1 April 2018. An exception to the policy has been added to take into account the practical concerns raised regarding deposit on acceptance. This exception will allow outputs unable to meet this deposit requirement to remain compliant if they are deposited up to three months after the date of publication.

# Compliance with RCUK and COAF open access policies

Q26/27/28/29

## Institutional compliance with RCUK and COAF OA policies

94. Currently, both RCUK and Wellcome collect data on compliance with their Gold OA polices independently.[6] Project partners were keen to understand what knowledge institutions had on compliance with their Green OA policies as a result of monitoring research outputs for REF. Institutions were asked to provide the number of their publications which fall within the scope of the OA policies of RCUK and COAF. The survey requested publication numbers for the period 1 April 2016 to 31 March 2017 ('the period of interest'), including those that are Gold OA only, Green OA only, and publications not available in OA form. Institutions were also asked to enter the number of outputs falling within the scope of both RCUK and COAF OA policies.

95. Institutions were also asked to record:

    a. Whether the RCUK and COAF compliance figures provided were actual or estimated figures;

    b. The estimated number of APCs paid out in the period of interest from sources other than RCUK or COAF funds;

    c. Further details relevant to these questions.

96. Institutional responses are provided in Tables 17-19. The figures in Table 18 show the number of publications which fall within the scope of RCUK and COAF OA policies. There was widespread double counting of outputs in the responses received, where the value of 'Both RCUK and COAF' was equal to the sum of 'RCUK only' and 'COAF only'. We excluded these instances from the calculation of the total number of 'Both RCUK and COAF' publications. The figures for 'RCUK only' and 'COAF only' for the same institution will be included in the calculation of the respective totals. Totals include both actual and estimated figures, and should be treated accordingly.

[6] An analysis of the RCUK open access block grant (August 2013-July 2017) can be viewed here.

97. The data returned also suggest that there was an inconsistency in institutions returning '0' or leaving answers blank to indicate either no answer or an actual 'zero' response.

98. Only those institutions that responded to both Questions 26 and 28 are included in Table 17.

**Table 17 Proportion of institutions that provided actual or estimated figures on RCUK and COAF funded outputs**

| | Proportion of institutions | | | | |
|---|---|---|---|---|---|
| **Funding Source** | Actual figures | Estimated figures | Combination of actual and estimated | Not applicable | No response |
| RCUK | 55% | 14% | 30% | 2% | 0% |
| COAF | 50% | 13% | 34% | 3% | 0% |
| Both RCUK and COAF | 36% | 14% | 39% | 4% | 7% |

**Table 18 Number of publications that fall within the scope of the OA policies of RCUK and COAF**

| **Category of Publications** | **RCUK only** | | **COAF only** | | **Both RCUK and COAF** | |
|---|---|---|---|---|---|---|
| | **Institutions** | **Publications** | **Institutions** | **Publications** | **Institutions** | **Publications** |
| Gold compliant | 85 | 13,633 | 37 | 2,727 | 28 | 1,005 |
| Green compliant | 74 | 7,642 | 18 | 413 | 10 | 2,420 |
| Non-compliant | 59 | 3,634 | 12 | 238 | 8 | 136 |
| **Total** | 88 | 24,909 | 38 | 3,378 | 28 | 3,561 |

**Table 19 Papers published under Gold OA in receipt of funding**

| Gold OA papers/APCs | Number | Proportion |
|---|---|---|
| RCUK and/or COAF funded | 13,487 | 71% |
| Other funding | 5,446 | 29% |
| **Total** | 18,933 | 100% |

99. The confusion as to the requirements of the survey question and the inconsistent use of zero and blank responses has led to uncertainty in regard to the robustness of the data. This has made any interpretation of the data unsound. However, a discussion of the free text responses alongside select elements of the data is presented in paragraphs 100-102. The survey does reveal the potential under-reporting of non-compliant COAF-funded articles, as Wellcome alone records roughly a thousand non-compliant articles a year.[7]

# Summary

100. Institutional responses received via free text suggested that HEIs employ a range of methods to identify information related to RCUK and COAF grants. These include both in-house systems and commercial services (for example, Web of Science, Scopus and ResearchFish). Several institutions commented that their systems are not currently deployed in a way that enables the easy and reliable tracking of research outputs and OA across the institution.

101. Over two thirds of Gold OA charges in the period of interest were funded by RCUK and/or COAF, highlighting the sector's reliance on these funds to deliver open access. Due to the difficulty for libraries in centrally tracking APCs paid directly by departments, the 71 per cent figure reported in Table 15 is likely to be an over-estimation of the actual proportion of articles funded by RCUK and/or COAF. Within the free text responses only 15 institutions made reference to an institutional OA fund for the payment of APCs,

[7] Volume of non-compliant papers estimated for 2015-2016 based on total number of papers in PubMed acknowledging Wellcome funding compared to number of full-text articles in PubMed Central acknowledging Wellcome funding.

further highlighting the importance of RCUK and COAF funds in enabling Gold OA in the UK.

102. Institutions found it difficult to monitor the compliance of Green OA articles. The use of copyright statements within institution repositories is variable and thus hampered the ability of several institutions to accurately estimate the number of compliant Green articles for articles funded by RCUK.

# Managing metadata

Q30/31/32/33/34/35

## RIOXX Metadata Application Profile

103. Institutions were asked whether or not they had implemented the RIOXX Metadata Application Profile and to indicate if they had a ‘basic’ or a ‘full’ version. We recognise that there was a discrepancy in the wording of this question, as there is only one version of RIOXX and not two, as the survey suggested. This may have caused some misunderstanding amongst respondents. In order to address this mistake we have combined the total responses to ‘Yes – Full RIOXX’ with ‘Yes – Basic RIOXX’. Findings are presented in Table 20.

**Table 20 Number of institutions that have implemented the RIOXX Metadata Application Profile**

| Response | Number of institutions |
|---|---|
| Yes | 57 |
| No | 52 |
| Don't know | 4 |

104. Of institutions who indicated that they did not use RIOXX, 63 per cent noted that that was due to incompatibility with their institutional CRIS or repository (for example, Pure). Respondents also indicated that the use of RIOXX often required software updates to ensure compatibility with other tools (for example, DSpace).

105. Alongside issues with staff resourcing RIOXX use, respondents widely reported that either the benefits are unclear and that there was difficulty mapping some data (particularly funding sources) with RIOXX. Institutional responses are provided in Table 21.

**Table 21 Reasons why institutions have not implemented RIOXX**

| Response | Percentage of institutions |
|---|---|
| Incompatibility with institutional systems | 63% |
| Lack of external support/guidance | 6% |
| Lack of time/resource | 31% |
| Lack of in-house expertise | 19% |
| Benefits of implementation are unclear | 29% |
| Other | 13% |

## Funders' Metadata

106. Institutions were asked if information on funding sources for articles is captured in institutional repositories (Table 22). The majority of respondents (64 per cent) capture this information. Respondents that reported this as 'optional' noted that this to typically owing to:

a. The repository not automatically ingesting this information from a CRIS;

b. Not all articles are linked to funding;

c. The variety of funding sources are not always represented in software choices.

**Table 22 Percentage of institutions that capture information on funding sources for journal articles**

| Response | Percentage of institutions |
|---|---|
| Yes - mandatory field | 14% |
| Yes - optional field | 50% |
| No - but we are working towards capturing this information | 23% |
| No - we have no plans to capture this information | 8% |
| Other | 4% |

107. Institutions were asked if funder metadata (including funder name, funder ID and project ID) is available in their Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) interface (Table 23). The majority of institutions either do not collect funder metadata, or do not know if this information is collected. Reasons for this were varied, with HEIs responding that:

a. Capturing this information is not mandatory;

b. Software restrictions;

c. Adequate systems are not in place, or a system upgrade is not yet complete;

d. Lack of interoperability between CRIS and repository.

**Table 23 Percentage of institutions that include funder metadata in OAI-PMH**

| Metadata | Percentage of institutions | | | |
|---|---|---|---|---|
| | Yes - RIOXX | Yes - other | No | Don't know |
| Funder name | 28% | 7% | 43% | 21% |
| Funder ID | 17% | 3% | 58% | 19% |
| Project ID/grant reference | 24% | 5% | 46% | 21% |

## Documenting licence choice

108. The survey asked institutions about their approaches to licence-selection. Forty-one (37 per cent) of responding institutions indicated that a member of their library staff selects the appropriate licence. This was the most favoured option amongst respondents. Only three institutions indicated that the author selects the licence, with no further review. A breakdown of institutional responses is captured at Table 24 below.

109. Twenty-eight institutions (25 per cent) responded that they have adopted an approach not included in the list of survey options. Respondents mentioned that research office staff are often responsible for selecting the correct licence type. This could be in discussion with library staff or authors.

110. A large number of responses to this question utilised the 'Other' option to highlight that licensing is not always clear (typically for AAMs or for authors) but that it is also funder or journal specific. A small number of institutions also highlighted their use of a standard licence, unless another specifically is requested by the author.

**Table 24 Institutional approaches to selecting licence types**

| Approach | Number of institutions | Percentage of institutions |
|---|---|---|
| Author selects licence, with no further review | 3 | 3% |
| Author selects licence, but it is subject to review by library staff | 27 | 24% |
| Member of library staff selects licence | 41 | 37% |
| No licence is specified | 13 | 12% |
| Other | 28 | 25% |

111. Institutions were asked to "state, or estimate, what percentage of articles and conference proceedings deposited in your repository during the period 1 April 2016 to 31 March 2017 were made available under each of the following licensing arrangements". Responses are captured in Figure 6, with raw data available at Table 25.

112. The data, particularly the large standard deviation (σ) of data for the majority of licence types, highlights the variability of use from institution to institution. This may be due to:

a. The number of HEIs providing default licences for manuscripts;

b. The number of institutions that were unable to provide an answer or only able to provide estimates.

113. The high volume of institutions that were unable to provide an answer or could only provide estimated figures reported that this was due to a lack of, or complexity of, the function within the research management system (RMS). Respondents also noted that it would take a considerable amount of time to answer the question accurately.

114. A number of institutions reported that they were not able to report figures lower than 1 per cent. These were typically CC-BY-ND, CC-BY-NC-SA and CC 0 licences. Other licences mentioned but not recorded included General Public Licence (GPL) and Open Government Licence (OGL).

115. Institutions reported confusion with the licensing terminology as requested in this survey and also more widely. Of those respondents who mentioned how Gold OA licences were recorded, these were interpreted as either ‘CC-BY’ or ‘CC-BY-NC-ND’. Institutions also reported that they applied their own (blanket) licence. There was further confusion with recording publisher licences (particularly with articles deposited as Green OA) with some respondents suggesting they applied either a blanket ‘no licence, in copyright’ or ‘CC-BY-NC-ND’.

Figure 6 Articles and conference proceedings deposited in the period of interest that were made available under the licence agreement options

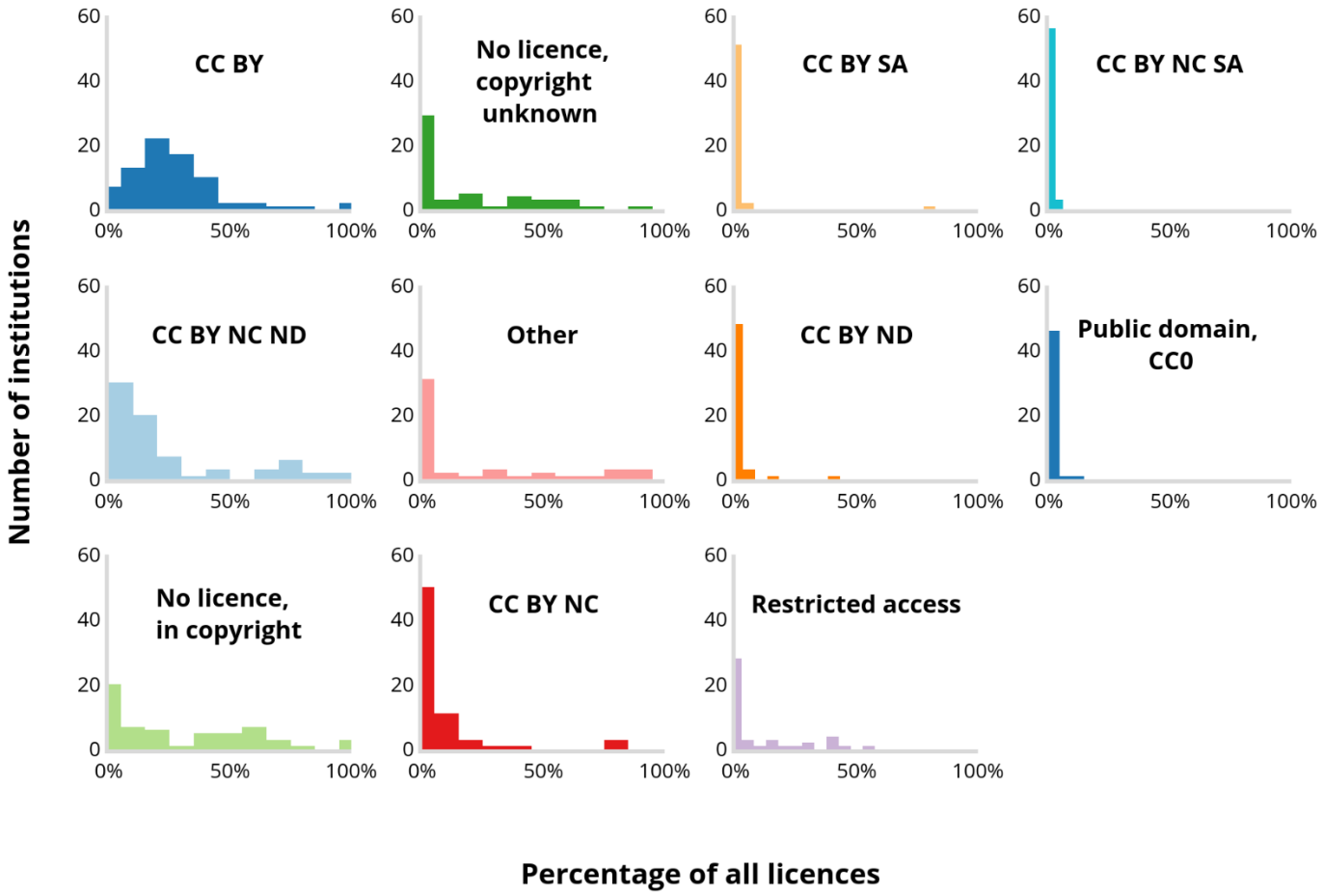


**Table 25 Articles and conference proceedings deposited in the period of interest that were made available under licence agreement options (raw data)**

| Licence | Number of institutions | Percentage of responding institutions | Average | σ |
|---|---|---|---|---|
| CC BY | 75 | 88% | 24% | 20% |
| CC BY SA | 26 | 31% | 1% | 9% |
| CC BY NC | 63 | 74% | 6% | 16% |
| CC BY ND | 36 | 42% | 1% | 5% |
| CC BY NC SA | 38 | 45% | 0% | 1% |
| CC BY NC ND | 74 | 87% | 23% | 29% |
| No licence, in copyright | 42 | 49% | 20% | 28% |
| No licence, copyright unknown | 25 | 29% | 9% | 20% |
| Restricted access | 21 | 25% | 5% | 12% |
| Public domain, CC0 | 20 | 24% | 0% | 1% |

| | | | | |
|---|---|---|---|---|
| Other | 23 | 27% | 11% | 25% |

## Summary

116. The implementation of RIOXX and the capture of funders' metadata largely reflect the choice and/or use of CRIS and repository software, as well as the technical ability and resources to implement this within an institution. As with previous discussions around institutional approaches to software, the uptake and implementation of RIOXX and/or capturing funder metadata would be greatly enhanced by further interaction between stakeholders.

117. The host of licences available, the variation in their use across Gold and Green OA, and the diverse understanding and implementation has made reporting and analysis difficult. In both this section and elsewhere in the survey a number of institutions discussed the potential role of the UK Scholarly Communications Licence (UK-SCL) in simplifying the licencing landscape.

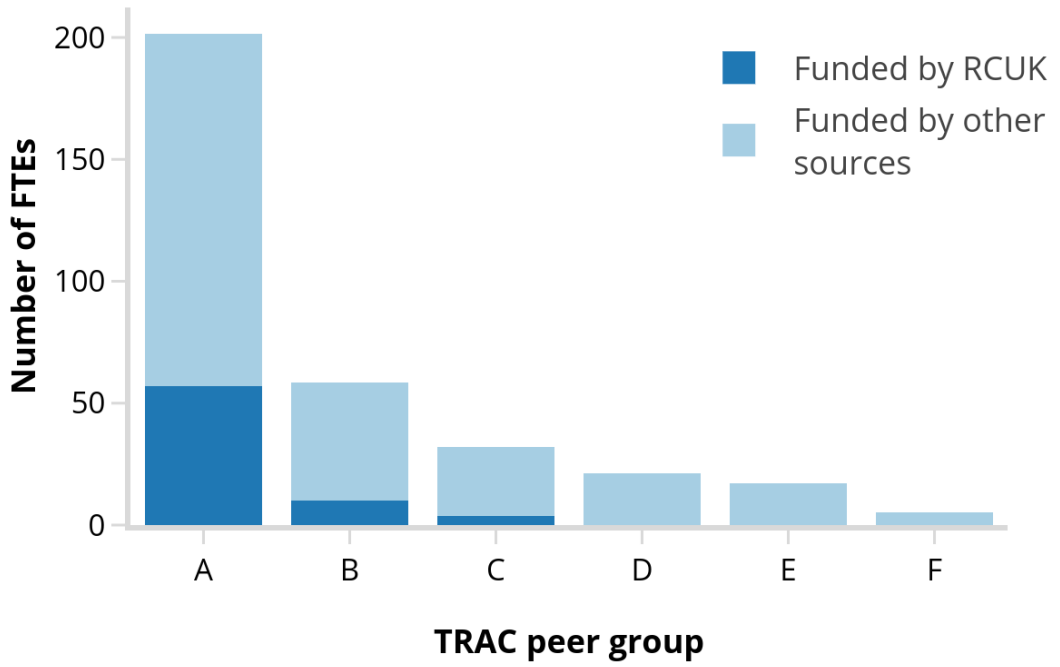
# The institutional cost of staffing open access

## Q37/38

118. Institutions were asked to estimate how many full-time equivalent staff members (FTEs) were directly engaged in supporting and implementing OA at their institution on 1 April 2017, and to allocate these within salary bands. These were further delineated by whether the costs were met by RCUK block grant or 'other' institutional funds, and the salary range of FTE staff. Results are captured in Figure 7.

119. TRAC group A employs the equivalent of 202 full-time members of staff, which is more OA professionals than all the other TRAC groups combined. Furthermore, 28 per cent of TRAC group A OA FTEs are funded via the RCUK block grant. The percentage of OA staff funded via the RCUK block grant drops off from TRAC group A through to C (equating to 79 HEIs) with groups D, E and F not using the RCUK block grant for this purpose. It is worth noting that there are institutions in TRAC groups D, E and F that receive the RCUK block grant, although the value of the grant varies year by year. A breakdown of TRAC groups can be found at Annex D.

**Figure 7 Estimate of how many FTEs are directly engaged in supporting and implementing OA**



120. The majority of staff directly engaged in supporting institutional OA practice are employed in the £20,000-£30,000 salary range whether funded by RCUK block grant or other institutional funding (across figures and salary bins, see Figures 8 and 9).

**Figure 8 FTEs directly engaged in supporting OA at HEIs funded by RCUK block grant. Broken down by TRAC peer group and salary range**

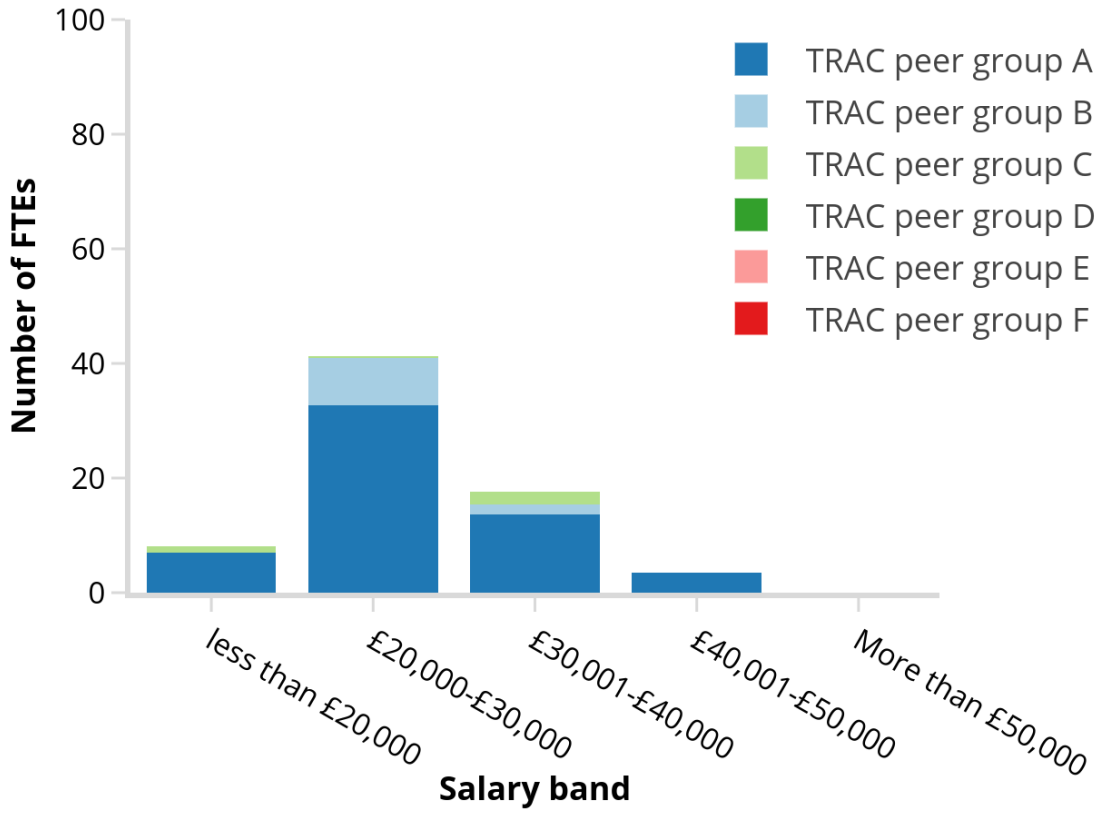
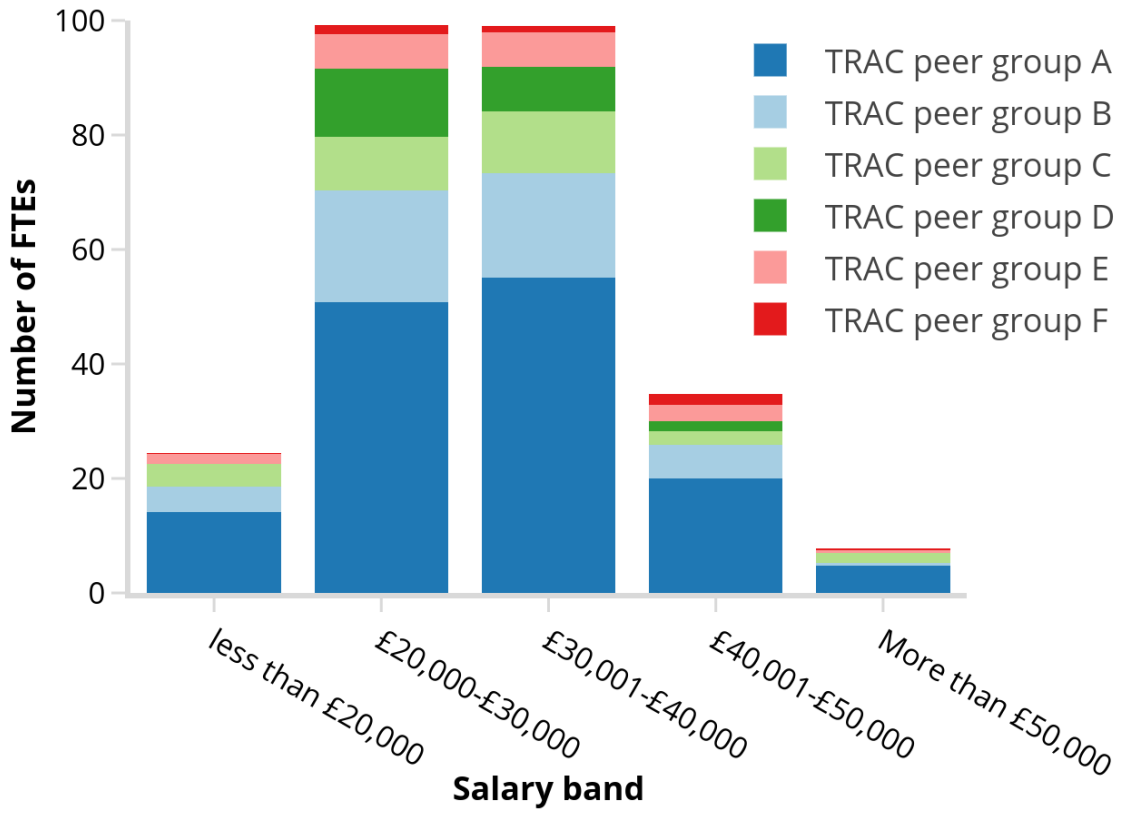


**Figure 9 FTEs directly engaged in supporting OA at HEIs funded by non-RCUK sources. Broken down by TRAC group and salary range**




121. Institutions were also asked to provide details of any additional resource implications of OA beyond those captured in the above figures. It is important to note that the survey is not an exercise to assess overall costs of OA. Respondents noted the following types of additional OA costs:

a. **Academic resource.** The amount of time spent by academics (of all levels) engaging with the requirements of OA policies.

b. **Institutional policy and advocacy.** Senior management involvement in setting policy and professional service/library staff involvement in advocating institutional and funders' policies.

c. **Infrastructure.** This includes the costs around the building and maintenance of OA software infrastructures and subsequent staff training.

d. **REF compliance.** The costs and future resources required in advance of REF 2021.

e. **Legal and financial services.** This included legal advice on licensing and copyrights with regard to third-party images and multimedia research outputs. The resource for the processing of APC charges was also highlighted.

f. **Monitoring.** The cross-departmental resource required to respond to separate funder OA monitoring requirements.

122. Responses received via the free text comments highlighted that there are a range of additional costs incurred within institutions that go towards creating and maintaining the infrastructure for OA. One respondent noted that “Green OA has never been a zero cost route to open research, and is predicated on maintaining substantial and increasingly difficult to sustain levels of academic journal subscription purchases”. Another respondent remarked that researcher time in achieving OA is a “complex landscape’” and that there are a number of ways that this could be made simpler – for example, through the use of standard terminology and licensing options, and an increase in publisher uptake of the Jisc Publications Router.

## Summary

123. Survey responses highlighted the resource-intensive nature of OA, with the equivalent of 335 staff at 1.0 FTE recorded as directly engaged in supporting and implementing OA at HEIs. Respondents in TRAC group A reported significantly higher numbers of staff working on OA than those institutions in TRAC groups B-F.

124. Institutions in TRAC groups A-C reported an allocation from the RCUK block grant towards staff directly engaged in OA. TRAC groups D, E and F did not report using the RCUK block grant for this purpose.

125. Additional costs incurred by universities for OA are wide-ranging. Although the survey is not designed to assess costs of OA, the responses captured give an indication as to where further costs lie. Notably, respondents highlighted the human resources required to ensure compliance with OA policy, citing examples such as time spent by academics engaging with policy, senior management involvement, legal advice, staff training, and monitoring requirements. Responses also highlighted operational costs of OA with a need to maintain appropriate IT infrastructures. Finally, respondents highlighted the need for training and skills at an institutional level to ensure that staff are kept up to date with resources and tools associated with OA processes.

126. The responses to this open question highlight the need for a better understanding of the overall cost of OA and how efficiencies can be driven by stakeholders moving forward.

# Conclusions

## General remarks

127. Institutional responses to the sector survey issued in summer 2017 demonstrate significant progress toward meeting the requirements of funder policies. However, respondents highlighted that systems which support and implement OA are largely manual, resource-intensive processes.

128. The extensive data gathered in this exercise will be considered as part of Wellcome's OA review, as well as UK Research and Innovation's internal review of its OA policies. Where relevant, data will be shared with REF panels to develop their criteria and guidance on submissions. Data will also be interpreted by Jisc in order to understand system use (particularly in relation to the adoption of the Jisc Publications Router) and identity how gaps in provision and uptake may be filled. The importance that institutions attach to the Publications Router service should also be noted by vendors of research information systems (CRISs) as they review how best to meet the needs of their clients, suggesting they should prioritise the development of interoperability with it. These findings should also encourage publishers to contribute to Publications Router.

129. Notably, some areas of the full sector survey recorded poor data quality, restricting analysis and making interpretation of the raw data set difficult. Based on feedback from respondents in qualitative statements, we recognise that greater clarity could have been provided in the survey and in the guidance document in relation to a few of the questions. Institutions also reported that some of the data requested was not available, or was too resource-intensive to collate. In particular, survey responses to the following two areas lacked clarity or questions were misinterpreted by respondents:

    a. **RCUK and COAF Green OA compliance.** There was a widespread double counting of outputs for the number of publications which fall within the scope of RCUK and COAF policies. There was also a low return of data from institutions who provided compliance data on RCUK/COAF Gold and Green OA policies.

    b. **Licensing.** Institutions reported confusion over the use of licences, their terminology, and publisher and funder policies. This may have led to inaccurate responses to the survey

130. The quality of this data will be considered in future OA reviews carried out by Wellcome and UK Research and Innovation.

131. The survey represents one year of data from participating institutions. Although the data provides a broad indication of publication compliance with funder policies, we recognise that this is only a 'snapshot' of the information held on record by HEIs.

## Compliance with funder policies

132. Over 80 per cent of outputs either met REF policy requirements in the first year (61 per cent) or an exception to the policy requirement is known to apply (20 per cent). These figures account for over 100,000 research outputs produced by UK universities. There are varying levels of compliance with the REF 2021 OA policy across institutions, and we would encourage the sector to work together in order to share best practice.

133. The intent of funders' OA policies is to make as many outputs freely available as possible. The survey responses demonstrate that, for example, some institutions are fulfilling the policy requirements only for the outputs they expect to submit to REF 2021 rather than those in scope of the policy.[8]

134. From 1 April 2016 to 31 March 2017, RCUK and COAF funded approximately two thirds of Gold OA APCs. With only 15 respondents making reference to an institutional OA fund for APCs, RCUK and COAF funding plays an important role in the Gold OA agenda.

135. Based on survey responses it would appear that additional article indexing and tracking for REF compliance is not currently leading to increased understanding of compliance with other funders' policies on Green OA.

[8] For example, in REF 2014, just 40 conference proceedings were returned to Main Panel A; in comparison, Main Panel B submitted 2097 of this output type, with 1898 from UOA 11 (Computer Science and Informatics). This may account for some of the 22 per cent of research outputs from subjects aligned with Main Panel A that were not compliant with funder policies. See Tanner, Simon (2016) An analysis of the Arts and Humanities submitted research outputs to the REF 2014 with a focus on academic books: An Academic Book of the Future Report, King's College London, November 2016. http://doi.org/doi:10.18742/RDM01-76

## OA infrastructure: challenges and opportunities

136. The report recognises the ongoing operational challenges faced by institutions to meet funders' OA policy requirements. The limitations imposed by resource constraints combined with myriad and complex software solutions to support OA, were a consistent feature of survey responses, particularly in regard to the following:

a. **Depositing AAMs.** Processes and solutions for depositing AAMs and checking licences are wide-ranging and principally carried out by dedicated research support staff. The complexities of publisher policies, variations in embargo periods and the extent to which they allow open access, combined with different funder policies, a lack of publisher/funder policy alignment, and a lack of machine-readable licences from either publishers or funders, prompt extensive cross-examination and compliance checking for outputs.

b. **Tracking APCs.** Institutions are likely to use more than one tool to track APCs and to ensure compliance with funder policies, increasing administrative burden. Institutions highlighted that their systems are not currently deployed in a way that enables the easy and reliable tracking of research outputs and OA.

c. **Monitoring RCUK Green OA.** Institutions noted the difficulties in monitoring compliance with Green OA, particularly in regard to licence compatibilities with RCUK's self-archiving policy. More generally, respondents highlighted the need to standardise funders' licensing policies to reduce administrative burden.

137. These challenges need to be addressed in order to reduce strain on institutional resources. Greater workflow efficiencies can be made between publishers and institutions, for example, by improving interoperability between the publisher's digital outputs and institutions' system solutions through the wider support of services such as the Jisc Publications Router. Vendors of research information management systems especially could be instrumental in addressing this by prioritising development of interoperability with this service.

138. RIOXX Metadata Application Profile was implemented to improve data quality and collection across common and key metadata fields. In order to support compliance with a funder's policy and ensure relevant and robust metadata is being collected and shared, the uptake and implementation of RIOXX would be enhanced by further interaction between stakeholder groups.

139. This survey has demonstrated the need for publishers, funders and research institutions to work jointly towards reducing burdensome manual processes. Through the uptake and improvement of automated digital workflow solutions between authors, publishers, institutions and funders, the sector can both ensure resources are directed to tackling other priorities and raise compliance for both publishers and funders.

140. There is an opportunity for the evidence presented in this report to be considered as part of UK Research and Innovation and the Wellcome Trust's OA reviews that are taking place over the next year. By addressing some of the challenges of OA the benefits of open access research and innovation may be further realised.

# Abbreviations

| | |
|---|---|
| AAM | Author accepted manuscript |
| APC | Article processing charge |
| API | Application programming interface |
| COAF | Charity Open Access Fund |
| CREST | Consortium for Research Excellence, Support and Training |
| CRIS | Current Research Information System |
| FTE | Full-time equivalent |
| HEFCE | Higher Education Funding Council for England |
| HE | Higher education |
| HEI | Higher education institution |
| ISBN | Individual Standard Book Number |
| ISSN | International Standard Serial Number |
| OA | Open access |
| RCUK | Research Councils UK |
| REF | Research Excellence Framework |
| RMS | Research Management System |
| TRAC | Transparent Approach to Costing |
| UKRI | UK Research and Innovation |
| UOA | Unit of assessment |
| UUK | Universities UK |

# Glossary

| | |
|---|---|
| Authors Accepted Manuscript / Post Print / Accepted Version (AAM) | The author's final, accepted manuscript is the one that has been agreed with the editor at that point. The accepted manuscript is not the same as the copy-edited, typeset or published paper – these versions are known as 'proofs' or 'versions of record' (VOR) and publishers do not normally allow authors to make these open-access. |
| Article Processing Charge (APC) | A single payment made to the publisher to make an output open access. This does not guarantee author retains copyright or a publication made available under a Creative Commons licence. |
| Charity Open Access Fund (COAF) | The Charity Open Access Fund (COAF) is a partnership between six health research charities, including the Wellcome Trust, to enable free and unrestricted access to the published outputs of the research they support. |
| Digital Bibliographic Library (DBLP) | A computer science bibliography website. |
| Directory of Open Access Journals (DOAJ) | A website that lists open access journals. |
| Inspire High Energy Physics (Inspire-HEP) | An open access digital library for the field of high energy physics (HEP), formerly Stanford Physics Information Retrieval System (SPIRES) database. |
| Jisc Publications Router | Jisc's Publications Router is an alerting service that automatically sends notifications about research articles to institutions' systems such as their repositories or CRISs. These notifications indicate, for example, that an article has been accepted for publication or that it has been published. They often include the articles |

| | |
|---|---|
| | themselves in the version agreed by the publisher, or they may just consist of metadata. |
| CC – Creative Commons | A free public copyright licence that enables the free distribution of an otherwise copyrighted work. A CC license is used when an author wants to give people the right to share, use, and build upon a work that they have created |
| BY – Attribution | Licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they give the author or licensor the credits (attribution) in the manner specified by these. |
| SA – Share-Alike | Licensees may distribute derivative works only under a license identical ("not more restrictive") to the license that governs the original work. I.e. without share-alike, derivative works might be sublicensed with compatible but more restrictive license clauses, e.g. CC BY to CC BY-NC. |
| NC – Non commercial | Licensees may copy, distribute, display, and perform the work and make derivative works and remixes based on it only for non-commercial purposes. |
| ND – Non derivative works | Licensees may copy, distribute, display and perform only verbatim copies of the work, not derivative works and remixes based on it. |
| CC0 – Public domain | Creative Commons Zero is a way to release work through to public domain, i.e. all rights expired. |
| General Public License (GPL) | A free software license, which guarantees end users the freedom to run, study, share and modify the software |
| Open Government License (OGL) | A copyright licence for Crown Copyright works published by the UK government. Other UK public sector bodies may apply it to their |

| | |
|---|---|
| | publications. It was developed and is maintained by The National Archives. It is compatible with the Creative Commons Attribution (CC-BY) licence. |
| United Kingdom Scholarly Communications Licence (UK-SCL) | The UK-SCL is a model open access policy with a standard set of licence terms designed for adoption by UK HE Institutions. It has been drawn up in response to researcher concerns about growing requirements to assign their copyright to a publisher at the point of acceptance, and in response to funder calls for a transition to a more open access environment. Implementation of the UK-SCL ensures that authors retain the right to share their manuscripts freely, and to reuse their research outputs in their own teaching and research. Authors retain copyright and, by extension, moral rights and are free to publish in the journal of their choice and, where necessary, to assign copyright to the publisher. The model is seen as an interim measure until a sustainable open access publishing model is implemented that facilitates sharing of scholarly outputs without delays or barriers. See website. |
| Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) | A low-barrier mechanism for repository interoperability. |
| Open Researcher and Contributor ID (ORCiD) | A unique identifier code for academic authors. |
| PubMed | Is a free search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. |
| RIOXX | A Metadata Application Profile that provides a mechanism to help institutional repositories comply with the RCUK policy on open access. |

| | |
|---|---|
| Research Papers in Economics | Provides links to over 1,200,000 full text economics articles. Most contributions are freely downloadable, but copyright remains with the author or copyright holder. It is among the largest internet repositories of academic material in the world. |
| Social Science Research Network (SSRN) | Preprint (AAM) service circulating scholarly research in social sciences and humanities. |
| Version of Record (VOR) | The copy-edited, typeset and published academic output. |