

informationpower

Usage of higher education institutional repositories and “Green” open access for longform publications

Gali Halevi, Dave Jago, Lorraine Estelle, Mikael Laakso,
Alicia Wise, and Thom Barnes-Wise

Contents

Foreword	3
-----------------	----------

Executive summary	5
--------------------------	----------

Introduction	7
---------------------	----------

Interviews with librarians	9
Strategy and funding	9
Academic perceptions and practice	13
Practical challenges	14

Quantitative analysis	16
Data sources	16
Limitations of the data	17
Volume of longform material	18
Use of longform content	19
Publisher data	26

Survey of institutional repositories	30
---	-----------

Academic views	38
Sharing via repositories	38
Discovery of longform materials	39
Using longform materials in repositories	40
Future outlook	41

Analysis	42
-----------------	-----------

Conclusions	44
--------------------	-----------

Acknowledgements	46
-------------------------	-----------

Annexe 1 – Case study: London School of Economics (LSE) Institutional Repository	47
---	-----------

Foreword

As a research funder, a publisher, and a national policy body, the British Academy has long supported the principle of open access (OA). Done well and sustainably, it has the potential to greatly broaden inclusion and accessibility, increasing access to knowledge and understanding. At the same time, we have been clear about our concerns that funding for and access to OA is currently unevenly available to SHAPE (social sciences, humanities and arts for people and the economy) research outputs, which are often longform publications. These books or monographs have a longer development period, higher production costs, and the ‘version of record’ plays a distinctive role in scholarly communication compared to other fields of research.

In 2024, Research England rightly recognised the distinctive challenges of achieving OA for longform publications and chose not to introduce a requirement for such outputs in the Research Excellence Framework (REF) assessment in 2029. This was an important acknowledgement that the sector needs clearer evidence about how OA for longform publications works in practice and more robust infrastructure before mandating change at scale.

The Academy remains committed to contributing constructively to this debate and to supporting the government’s ambition through evidence-based policy development. With that aim, we commissioned Information Power to examine the potential and limitations of “Green” OA for longform publications.

This study brings together interviews and surveys with repository leaders, quantitative usage analysis, and the perspectives of academics in SHAPE fields—to produce an unusually holistic view of how accepted manuscripts of monographs, edited collections and chapters are currently deposited, discovered and used through UK institutional repositories.

The picture that emerges shows great variability. Many institutions have extended journal-oriented OA policies to books, yet few possess a dedicated longform strategy; repository systems often struggle with the formats, rights and reading practices associated with books; and authors remain reluctant to share author accepted manuscripts they consider unrepresentative of their finished work. Meanwhile, the costs of fully open publication continue to challenge university budgets and risk entrenching inequalities if left to adhoc funding. These are not problems of principle but of design, resourcing and coordination and point to three priorities for all of us in the system to act upon:

A clear strategy for longform OA. Progress will require a joined-up approach that is co-designed by funders, universities, libraries, publishers and scholarly communities across the country. Such a strategy should clarify aims, roles and timelines; align repository standards and metadata requirements to improve searchability; and provide for monitoring and evaluation that recognises the specificities of monographs, edited volumes and scholarly editions.

Sustainable funding and infrastructure. Longform OA cannot rest on unfunded mandates or on charges that are unaffordable for most institutions. Targeted investment—in Diamond and community-led models, and in repository capabilities (rights management, accessibility, usability and persistent identifiers)—is essential if OA books are to be both publishable and usable.

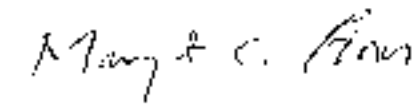
Meaningful author engagement and protection. Policy must address legitimate concerns about version quality, third-party rights, career incentives and international readerships. Early-career researchers, people with fewer financial means, and those publishing image-rich or archival material should not be disadvantaged by the route to openness we collectively choose.

Thanks to Information Power and the participants in this study, we can now start to chart a course through this complex landscape. We want to encourage, invite and convene conversations to grapple with this and begin to assemble the conditions that would allow longform OA to succeed.

A handwritten signature in black ink, appearing to read "Lindsay Farmer". The signature is fluid and cursive, with the first name being more prominent.

Professor Lindsay Farmer FBA

Vice-President (Publications)

A handwritten signature in black ink, appearing to read "Margot Finn". The signature is cursive and somewhat stylized, with the first name being more prominent.

Professor Margot Finn FBA

Vice-President (Research and Higher Education Policy)

Executive summary

Open access (OA) has become mainstream for journal articles, but progress for longform outputs—particularly monographs, edited collections, and book chapters—has been slow.

A consultation in 2024 resulted in a decision not to proceed with the proposal that longform submissions to the 2029 Research Excellence Framework (REF) should be OA. In its submission¹ to this consultation, the British Academy emphasised its support for the principle of OA and also the need for any requirements for longform OA to be properly funded and to protect research integrity and excellence. The British Academy asked that there be a full impact assessment of proposed changes informed by an assessment of the likely behaviours of key stakeholders including universities, authors, publishers, and libraries. The British Academy also emphasised how important it is to promote inclusivity and maintain a positive research culture and environment.

In 2024 UK Research and Innovation (UKRI), in partnership with the British Academy, Association of Learned and Professional Society Publishers (ALPSP), and Open Access Scholarly Publishing Association (OASPA), commissioned a report from Information Power on longform OA publishing, which was published in February 2025². In March 2025 the British Academy invited proposals for a study specifically focused on the ways longform scholarly publications are made available and used through UK higher education institutional (HEI) repositories as part of the Green OA landscape. The British Academy sought this evidence to fill key evidence gaps and inform policy, particularly because the proposals for longform OA in the 2024 REF consultation would have relied heavily on Green OA. This report by Information Power is the result.

We were asked to examine how accepted manuscripts of longform works are deposited, discovered, and used, and what institutional, financial, and cultural barriers impede broader adoption. This report provides one of the first comprehensive analyses of both qualitative and quantitative data concerning the deposit, discoverability, and use of longform OA materials in repositories.

The project integrated four complementary strands of evidence: interviews with library and repository leaders from UK institutions, a national survey of repository managers, quantitative analysis, and focus groups with academics in SHAPE (social sciences, humanities and the arts for people and the economy) fields.

We conclude that Green OA for longform works is contested and not yet mature in practice. There is evident support for openness in principle, but without a national strategic framework and cross-sector coordination progress will remain slow and fragmented. Key stakeholders who need to be engaged in this process include authors, employers, funders, libraries, and publishers. A coordinated, evidence-based roadmap is essential if the UK is to achieve equitable, effective, and durable OA for longform scholarship.

Main findings

Most universities lack a dedicated longform OA strategy. Existing policies, developed for journal articles, are inconsistently extended to books. Most focus on encouraging academics

1 https://www.thebritishacademy.ac.uk/documents/5456/OA_and_REF-British_Academy_Position_Paper-July2024.pdf

2 Estelle, L., Jago, D., Jones, R., Laasko, M., Snijder, R., and Wise A. (2025), 'Supporting learned society, subject association, and smaller specialist publishers to transition to open access book publishing (SPA OPS 4.0 project)'. <https://zenodo.org/records/14679068>

to voluntarily embrace longform OA and on supporting case-by-case compliance with funder policies.

OA publishing is viewed by many of the stakeholders with whom we engaged as the preferred approach, but there are problems. Gold OA is viewed as financially unsustainable, with book-processing charges (BPCs) averaging £10,000–£20,000. Institutions are experimenting with community-led and Diamond OA models, but budgets remain constrained and long-term funding mechanisms are unclear. Green OA is perceived as a less desirable way forward.

Many academics, especially in SHAPE (social sciences, humanities, and the arts for people and the economy) disciplines, view Green OA as a “last-resort” access route rather than a valued dissemination route. Authors support the principle that everyone should have access to their book or book chapter. However, they are much more positive about Gold OA dissemination of the publisher’s version of record (VoR), and are often reluctant to deposit author accepted manuscript versions in repositories, perceiving this avenue to OA as potentially harmful to professional reputation and inferior in format and reading experience to published versions. For early-career researchers especially, aligning with well-known publishers is still seen as critical for career progression, tenure, and professional recognition, creating an incentive for Gold OA and disincentive to deposit a submitted manuscript in a repository.

Repository environments do not provide strong support for the engagement and extended reading and annotation that SHAPE scholars desire. Institutional repositories aspire to take on functions of the broader publishing ecosystem, but without sufficient funding this ambition risks creating fragile, fragmented, and unsustainable services. Most institutional repositories are optimised for journal articles, not books, which bring specific challenges including complex formats, metadata inconsistencies, and copyright complexities. It is nearly impossible to meaningfully track how publishing and sharing practices for longform materials are changing over time.

Extending Green OA mandates to longform outputs should be approached with caution. A rapid push in this direction, or one lacking in nuance, risks undermining the fragile progress made toward building sustainable Diamond and community-supported OA publishing models. Policymakers and funders should recognise that, without careful design, Green OA for longform materials could inadvertently weaken the business and community foundations now enabling fully open publishing.

Cross-sector collaborative dialogue is essential. As the UK higher education sector navigates a period of acute financial constraint, it is vital that all stakeholders come together to explore sustainable approaches to longform OA. Bringing together universities, research funders, libraries, learned societies and other publishers, authors, and readers can help to identify the models that both strengthen scholarship and ensure equitable access to knowledge.

Longform publishing is an international activity. The publishing model for monographs relies on international funding and is not sustainable without this. International collaboration is essential. Policy development offers a timely opportunity to reinforce the UK’s leadership position in SHAPE research and in academic publishing, and to maximise impact and return on investments. By aligning collective interests, it is possible to design policies and services that safeguard research quality, expand readership and engagement, and bolster economic resilience.

Introduction

The British Academy, as the UK's national academy for the humanities and social sciences, is a research funder, a publisher, and a policy voice contributing to debates on behalf of the important SHAPE disciplines (i.e., social sciences, humanities, and the arts for people and the economy). It has a longstanding interest in OA³.

In response to growing national and international interest in the future of OA for longform scholarly publications—particularly monographs, edited collections, and book chapters—the British Academy commissioned Information Power to investigate the availability and use of author's accepted manuscripts (AAMs; i.e., Green OA versions) of longform publications in UK HEI repositories. Many Green⁴ OA policies encourage that these versions of longform outputs be made freely available through institutional or subject-specific repositories.

The British Academy recognised that while OA deposit in repositories is now well-established for journals, less is known about how longform OA content is being deposited and used. There was concern that policy decisions might be made without sufficient evidence about how such content is accessed, by whom, for what purposes, and with what impact.

To address this gap, the British Academy sought to gather robust data and qualitative insights on the real-world use and value of Green OA longform outputs. The aim was to explore not only usage metrics but also institutional readiness, user behaviours, and perceptions across the academic sector. In doing so, it was hoped that the project would inform policy development so that OA policies reflect the distinctive practices and challenges of the SHAPE disciplines, and that they support, rather than undermine, the continued production of high-quality scholarship in these fields.

Over the past two decades the OA movement has transformed the landscape of scholarly communication, particularly in the journal article domain. Driven by the principles of transparency, accessibility, and public accountability, especially where research is publicly funded, OA has gained wide institutional and policy support in many parts of the world. Major funders such as UKRI, the European Commission, and the Wellcome Trust have adopted OA mandates, and an increasing number of journals now offer OA publishing models.

However, the pace and scope of OA adoption have been uneven across publication formats. While shortform outputs such as journal articles benefit from more established infrastructure, funding models, and mandates, longform formats have lagged. The reasons are multifaceted, ranging from higher production costs and the lack of sustainable processing charge models to more complex intellectual property arrangements and less consensus on what constitutes a viable OA model for books.

Longform scholarly works remain a cornerstone of knowledge production and dissemination, especially in the SHAPE disciplines. Monographs, edited volumes, and book chapters provide the space for deep, sustained argumentation, complex narratives, and engagement with primary sources that cannot be easily condensed into journal-length outputs. These works are

³ <https://www.thebritishacademy.ac.uk/programmes/open-access/>

⁴ Throughout this report we have used labels such as Diamond, Gold, and Green OA. In a nutshell, the first two are approaches that result in the publisher's VoR of the book itself being published OA. Green OA is a compromise in many ways, as the book needs to be sold in traditional ways to cover costs and the author's final manuscript submitted to the publisher is made available OA with a reuse licence. However, because OA journal policies have often just been extended to longform outputs, basic things that happened in advance of OA mandates for journals—like the identification of different stages that articles pass through in production, and consensus building across stakeholders about which of these should be deposited as the final manuscript—has not happened for books. Labels relevant for journals fit awkwardly in discussions of OA for books.

central to academic advancement in the SHAPE disciplines and serve as critical resources for cultural institutions, educators, independent researchers, and policymakers. Despite this centrality, OA for longform content remains significantly underdeveloped. Unlike journal articles, longform works often lack dedicated funding streams for open publication, and the purchasing behaviours of librarians have not yet evolved to drive a change toward a sustainable OA publishing ecosystem for books.

During this project we compiled or generated four strands of evidence: quantitative usage data, qualitative data drawn from interviews with library leaders and repository managers, a national survey of institutional repositories, and focus groups of academics. These work packages were undertaken in parallel, and in this report we begin by presenting the interview data.

Interviews with librarians

After desk research and consultation with the project Steering Group, we reached out to request interviews with colleagues across the UK. We spoke to 12 colleagues—library directors, scholarly communications experts, REF managers, and repository managers—across a range of institutions. All interviewees had responsibility for institutional or cross-institutional repositories. These discussions offered valuable insights into institutional strategies and policies, repository operations, the discoverability of longform works, purchasing decisions, and rights retention practices. Participants also shared their thoughts on the evolving landscape around OA for longform outputs.

We also interviewed a representative from IRUS-UK⁵, which is a national, standards-based aggregation service that collects and provides COUNTER-compliant⁶ usage statistics for UK institutional repositories. The aim of IRUS-UK has been to enable repository managers to monitor and compare usage data consistently across institutions, supporting evidence-based decisions and reporting. The service is being discontinued and will end in April 2026 but users can transition to the OpenAIRE Provide⁷ service, which provides similar data.

The interviews were semi-structured, following a core script tailored to each participant’s role. This approach allowed for consistency across conversations while also accommodating the specific experiences and responsibilities of different stakeholders. Each interview lasted approximately 1 hour and was conducted virtually. With the interviewee’s informed consent, all sessions were recorded and transcribed to ensure accuracy and to facilitate detailed analysis. Following each interview, the draft transcript was shared with the interviewee for review, allowing participants the opportunity to clarify or elaborate on their responses.

The key themes that emerged from these interviews relate to strategy, funding, academic perceptions and practice, and practical challenges.

Strategy and funding

Institutional circumstances and vision

Library aspirations for longform OA policy generally came through in some of the discussions, and included building institutional collections, cost savings, and improving services such as discovery, access, and even AI and text and data mining.

‘Ideally, we’d make everything open, because it broadens the reach of our research.’

‘There needs to be a clearer message about the value of Green open access beyond just REF compliance. We need to articulate how repositories serve discovery, sampling, mining, and broader access—because that’s where the real benefit lies.’

Most institutions and libraries, however, did not yet have a dedicated longform OA strategy. All had extended their journal-oriented policies and practices to longform where possible, and provided support to academics to comply with funder policies on a case-by-case basis. Rights retention policies were proving tricky to extend from journals to books, and when in place were being handled sensitively so as to keep both academic authors and their publishers on board.

5 <https://www.jisc.ac.uk/irus>

6 <https://www.countermetrics.org/>. The COUNTER standard means that publishers, aggregators, and technology providers around the world can deliver credible, consistent, comparable usage metrics.

7 <https://provide.openaire.eu/home>

'We have a university research publications and copyright policy, but it only applies to journal articles and conference proceedings, shortform outputs. Longform works, like monographs, are strongly encouraged to be open access, but there's no institutional mandate.'

'We have a suite of open research policies, and open access is part of that, with encouragement for open access. We also have a rights retention policy, but we made a deliberate decision to keep longform out of scope.'

'In our open access policy, which we revised last year, we incorporated rights retention, including the possibility of applying it to monographs. However, we're transparent with our academics that publisher reactions are uncertain.'

'While we've always encouraged depositing content in the repository, it's now formally part of university policy. The longform inclusion is relatively new, and we see it as a long-term goal to embed the policy more fully.'

Funding concerns were mentioned by most, although not all, interviewees. Sometimes this was a concern about university or library funding generally. More often the concerns related to an unresolved tension between aspiration and means. Interviewees expressed a vision for a world in which all books were published OA and also available through repositories that had evolved to rival or even surpass the platforms and user experiences of large commercial publishers, all while spending less money than they do now. However, Gold OA for longform outputs was viewed as prohibitively expensive and it was difficult to elicit any clear views on the potential funding for, or costs of, Green OA for longform outputs.

'The university is currently in a difficult financial situation, it's been in the news. So any hopes for internal funding [for longform OA] have basically disappeared.'

'We've done some modelling to estimate the cost if the university mandated all monographs be open access via book-processing charges (BPCs). A couple of years ago, our back-of-the-envelope calculations suggested we'd need an additional £2–3 million per year. The issue is that BPCs vary greatly from £5,000 to as high as £20,000. Most hover around £12,000. UKRI only provides up to £10,000, but in reality, the fund they provide for BPCs only covers about £8,300 per book. So, researchers are stuck between insufficient funding and high publisher fees. That's why we believe BPCs aren't a sustainable long-term solution. Instead, we're focusing on supporting alternative publishing models.'

'We're already trying to shift toward more open purchasing, but it's tough. Budgets are being cut, and most of our book budget goes toward textbooks, not monographs.'

'We just don't have the resources, staff time, or funding to provide fully accessible versions like flowable XML. So I truly believe Green open access for longform works is not desirable for authors, readers, or institutions.'

'We're hoping our new repository system will allow us to mint [Digital Object Identifiers] DOIs for other research outputs, including accepted manuscripts. That would significantly improve discoverability and compliance with UKRI's persistent identifier requirements. If we're allowed to use the final published version, we usually close off the accepted manuscript initially. We keep the accepted versions mostly for preservation and long-term storage. But since we're dealing with longform content and we want our academics to engage with the policy, we aim to meet their needs too. They generally prefer the final published version to be available online, if permitted.'

UKRI policy

The UKRI policy⁸ is that from January 2024 longform research outputs acknowledging funding from UKRI or any of its councils need to be made OA within 12 months of publication. UKRI set up a central fund of £3.5 million per annum to support publication costs for in-scope longform outputs, and there is a cap, including VAT, of £10,000 for books or £1,000 for book chapters. Compliance can also be achieved by depositing an AAM in a repository and making it available within 12 months of online publication under a Creative Commons⁹ licence.

The UKRI OA policy for longform outputs is both contested and underfunded: authors remain reluctant to share accepted manuscripts, and current funding levels do not enable Gold OA publishing or cover the true cost of compliance.

‘Authors feel strongly that accepted manuscripts aren’t equivalent to the published version, especially in book publishing, and they resist uploading them.’

‘It was a tough sell, but since it’s externally mandated, our approach was to emphasise that if you’re accepting UKRI funding, you need to meet the requirements tied to that funding.’

‘Right now, we’re running a £20,000 deficit because of that cap. Ideally, I’d love to implement a rights retention policy for UKRI-funded books and chapters, similar to what we do for journal articles. That would allow us to make accepted manuscripts openly available without additional cost.’

‘The UKRI policy has changed, and we’ve had maybe 20 monographs acknowledging UKRI funding, where we’ve applied to the central pot. But the policy’s impact has been relatively modest, it’s not going to drive a sea change. A REF requirement would be the real game-changer.’

REF policy

In 2024, a consultation resulted in a decision not to proceed with the proposal that longform submissions to REF 2029 should be made available OA, with Green OA as the default mechanism to achieve this.

The REF exercise was mentioned frequently during the interviews, with many interviewees signalling that this would be important to drive an increase in longform OA. There was, however, a widespread concern should this happen in a rush for 2029 or if the approach were to mirror OA policy for journals.

‘While most authors support equitable access to their work and want researchers and the public to be able to access it, I think they’re struggling with the disempowerment they feel. Their work becomes a pawn in a game between institutions for REF.’

‘It’s interesting that there’s a contradiction between what policymakers are trying to do and how disconnected that is from academic and publisher behaviour.’

‘One issue is that REF and UKRI policies are largely modelled on journal articles, with little awareness of how books differ.’

‘My concern is that the market isn’t ready. These smaller publishers won’t be able to adapt, and funders will start pushing for Green open access for monographs. Anecdotally, authors have told us they don’t want their books available as a 400-page Microsoft Word document. They want their work to look polished and professional.’

8 <https://www.ukri.org/publications/ukri-open-access-policy/>

9 <https://creativecommons.org/>

'We anticipated REF 2029 might include longform in its open access requirements, which worried us and many others. Without that driver, we didn't feel the need to rush policy changes, which gives us time to explore what a longform policy might look like and talk with our community about it.'

'Unless something changes, whether out of necessity or because repository infrastructure improves, most libraries don't see Green OA as the most desirable route for monographs. It may be the only affordable route, and it might be what ends up happening, but it's not what people would choose if given other viable options.'

Publisher policies

While librarians have responded to publishers on longform OA strategy, this is in a different way to their historical engagements with publishers around OA journal strategy. The focus is infrequently on dialogue with long-established and more commercial book publishers (e.g. through trade associations) and most often takes place in the context of collaborative community-based initiatives that advance new library and university presses (e.g. COPIM) or open infrastructures (e.g. OAPEN). Libraries and library consortia are beginning to experiment with co-funding OA books through collective models or subscribe to OA publishing initiatives.

'We're cautious about using funds for large publishers asking £20,000 to make a book OA. But if someone's working with a smaller, ethically aligned press, we're much more likely to support that.'

'We've also started using our materials budget to invest in open infrastructures, including monograph publishing initiatives. We participated in an RLUK/Russell Group study and were the second-highest investor in open access monograph publishers.'

'We invest in community-driven models like Open Access Community Framework, MIT Press initiatives, Subscribe-to-Open, and similar. These models are more sustainable for us. We don't have any funding for BPCs, it's just not feasible. I can't imagine how we'd fund our entire REF submission through a BPC model.'

'There's a lot of resistance in the humanities. People see significant value in traditional publishing workflows. The challenge is to show them that open models aren't about stripping away those values, they're about achieving the same outcomes in different ways. That's why mandating open access is key, it kickstarts the shift in mindset. People start engaging with it, realise it's not as bad as they thought, and gradually become more open to change.'

'We also need to keep emphasising the benefits: discoverability, impact, rights retention. And to frame it not as a rejection of traditional publishing, but as an expansion of how scholarship can reach audiences.'

'We're comfortable with embargoes, I genuinely believe the monograph is a distinct form. We publish open monographs, so I understand both the effort and cost involved. It's not comparable to a journal article, and I do have some sympathy for that difference. Ideally, I'd love for us to offer open access versions of longform works. I'm particularly interested in establishing a secondary publication right with a reasonable embargo; 2 years seems fair. If funders support that, we could have a sustainable approach that doesn't unduly harm publishers.'

'We work within the limits set by publishers. Real change has to come from them. If publishers start mandating certain formats or accessibility standards, that's when we'll

see real improvement. Diamond open access publishing has existed for a long time, it's not just commercial publishers. Some of the Diamond presses in Europe offer great platforms.'

'To be honest, I think the way to transform longform outputs is to see certain emerging publishers who are prepared to take risks and demonstrate that it's viable; actually even undermine the already failing books market by grabbing a significant percentage of the publishing. I don't know about the business model, but if someone came in with a new model and was able to open up all their longform outputs, do a really good job of producing them, make people think, "Oh, actually, that's a really nice venue for publication," and that started to undermine Routledge, Palgrave, etc., then maybe that would start to gain some momentum.'

Academic perceptions and practice

Interviewees expressed the concern—voiced by authors at their institutions—that Green OA versions do not fully or respectfully represent their work. These versions often lack the final typesetting, illustrations, formatting, or design features of the final publication, leading some scholars to feel that their intellectual contributions are shared in a diminished or incomplete form. Instead many scholars viewed versions in repositories as incomplete or unofficial drafts, and lacking the editorial rigour and perceived authority of the publisher's VoR. This perception can reduce the willingness of authors to deposit their work, even when policy mandates require it and/or publishing agreements allow it. For early-career researchers, especially, aligning with well-known publishers is still seen as critical for career progression, tenure, and professional recognition, creating an incentive for Gold OA and disincentive to deposit a submitted manuscript in a repository.

Authors feel strongly that accepted manuscripts aren't equivalent to the published version, especially in book publishing, and they resist uploading them.

'Many authors don't like the idea of their work being made available in a "less-than-perfect" form.'

'A poor format reflects poorly on the author and their institution.'

'If authors were forced to choose from a limited list of publishers because of funding or infrastructure restrictions, that would be very unpopular. They've spent years on a book—they want to choose where it's published. Also, there's concern around depositing author-accepted manuscripts. With journal articles, the difference between accepted and published versions is smaller. But with monographs, there's often a big gap between the submitted and final version. Authors might worry that earlier deposits look like inferior versions.'

Librarians reported that when academics access longform repository materials, they often do so for specific purposes and not because they wish to read the entire text. For example, one interviewee reported that academics may read sections of a monograph in an institutional repository to decide whether it was useful for their research, which could then be used as evidence to help justify purchasing a book for the library.

'No one wants to read an unformatted, unedited 400-page document.'

'Realistically, I don't think people are reading entire books published in a Word format on a repository. But we do have experience with people reading parts of books... and then realising, "Actually, this is useful for my research," which helps justify the purchase.'

Librarians were increasingly aware of author concerns about automated data scraping, particularly when content could be harvested without attribution or used in ways that might compromise intellectual property or academic integrity. These concerns could introduce a new layer of hesitation around making longform works freely accessible, even among those otherwise supportive of OA principles.

'Some authors don't want to upload to repositories if it means [the texts will] be harvested by AI companies.'

'We're seeing a lot of worry from authors about people using it for AI. And it's fascinating, because from my perspective, I would much rather AI was trained on legitimate content than the rubbish on the rest of the Internet.'

Practical challenges

The interviews suggest many university libraries face ongoing practical challenges in managing repositories generally and longform content in their repositories particularly. Many of the institutions that we interviewed operate with lean repository teams, often composed of generalist librarians or part-time staff who juggle multiple responsibilities.

Deposits

Most institutional repositories were designed with journal articles and shorter outputs in mind. Unlike journal articles, monographs are larger, more complex files that can be awkward to import, require tailored metadata, where there are not well-established protocols for version naming and tracking, and where readers benefit from good formatting. The ingestion of longform content can be a labour-intensive process requiring manual effort in areas such as format preparation, metadata entry, version control, rights management, and quality assurance. Several interviewees emphasised the disproportionate amount of time and attention required to process a single longform work compared with other repository items. These issues are compounded when dealing with publications with non-standard file formats or hybrid media objects.

Librarians expressed unease about potential copyright and intellectual property issues with Green OA. Longform publications frequently contain third-party content such as images, archival materials, or co-authored chapters, which pose legal complexities. Rights clearance for these elements can be time-consuming and legally complex, often requiring specialist support. Librarians also expressed unease about potential copyright and intellectual property issues with Green OA, particularly where contracts with publishers often do not permit open sharing of full works unless funding has been found to enable the book to be published Gold OA.

Interviewees emphasised the importance of shared advocacy materials, such as author-facing guides, policy templates, and institutional case studies, to build awareness and foster engagement across academic communities. These resources were seen as critical for aligning expectations between researchers, administrators, and support staff, while also reducing the burden on individual institutions to develop materials independently.

Despite all this effort, many librarian interviewees were uncertain about how many monographs, edited collections, or book chapters were hosted in their institutional repositories. In several cases digital monographs were embedded within larger collections, inconsistently labelled, or missing key metadata, making them difficult to identify or track. Some repositories could not distinguish between versions of the book (e.g., to distinguish simple PDF uploads from enhanced versions that include additional features, structure, or interactive elements).

While AI is still in an early stage of adoption within library workflows, one interviewee expressed a sense of optimism about the potential for automation—through artificial intelligence—to ease some of the administrative burdens of uploads, content structuring, and metadata normalisation. Several interviewees expressed interest in participating in pilots of these tools to see if they could reduce manual labour.

Use

A recurring theme was the general inability to track usage of longform content within institutional repositories. Most respondents acknowledged that usage metrics—such as downloads, views, or engagement with longform works—are either not systematically collected in the repository or are difficult to interpret due to platform limitations. As a result, libraries and repository managers often cannot provide clear insights into how longform works are being accessed or used. This lack of visibility can make it challenging to demonstrate the impact of this sharing and of these works and to advocate for continued investment in longform publishing infrastructure.

While one interviewee cited specific steps being taken to improve the discoverability of material in their repository (e.g., the addition of metadata, keywords, and links to improve search engine optimisation and the inclusion of DOIs where possible), others were concerned about effective discovery and use. Inconsistent metadata, lack of integration with discovery platforms, and unintuitive interfaces contributed to concerns that repository contents could be difficult to locate and were less likely to be used.

Longform materials were known to benefit from more sophisticated interfaces that support sustained reading, structured navigation, and logical sectioning. Unlike commercial eBook platforms, most repositories do not support in-text annotation, bookmarking, or integration with e-reader devices. This is likely to hinder usability for readers who need to reference specific parts of a text or who rely on interactive reading tools for research and learning. Many repositories lack the infrastructure or workflows to produce content in formats compatible with screen readers or structured navigation aids for users with disabilities. Without clear standards and tools for accessibility, institutions risk falling short of legal and ethical obligations to provide inclusive access.

Several interviewees highlighted the pressing need for evaluation frameworks specifically designed for longform scholarly outputs. Unlike journal articles, longform works lack clear, standardised metrics for assessing their reach, usage, and scholarly impact. Tailored evaluation tools would enable institutions to make more informed decisions about where to allocate resources, demonstrate value to stakeholders, and report on the effectiveness of OA initiatives.

Quantitative analysis

In parallel with the library interviews, we gathered data to understand the volume of longform material available in UK institutional repositories and how this is used.

Data sources

Four data sources were used for this project: IRUS-UK, Dimensions, data available from institutional repositories, and data shared by two publishers.

IRUS-UK

IRUS-UK¹⁰ is a service provided by Jisc that collates usage statistics for repositories across the UK. A total of 143 UK HEIs provide data to IRUS-UK, along with 19 institutions such as Tate, the British Museum, the OAPEN Foundation, UK Data Service, the Natural History Museum, and the National Trust. Some HEIs have more than one repository covering differing types of content; separate repositories are sometimes maintained specifically for data or for theses. Where this was the case all were included, as in some instances relevant usage was reported from repositories that were seemingly out of scope such as data repositories. As we were primarily concerned with usage in HEI institutional repositories, we filtered out the data for non-HEI repositories.

IRUS collects data from repositories and processes them into COUNTER R5 usage statistics¹¹. COUNTER R5 data are available for January 2020 onwards. We used Unique_Item_Requests, which counts the number of views of a full content item but de-duplicates when multiple file types are downloaded from the same item record in the same user session. The Custom Platform Master Report gives an overview of usage across the repositories, and we looked at these data over the period 2020–2024. The Custom Item Master Report gives usage data for each specific content item; for practical reasons, we limited this analysis to item accesses in 2024.

Some data fields were disappointingly incomplete. For example, there is a field for the publisher of the content, but much of this data was missing. A sample of over 1,000 longform item uses gave publisher information for only 35% of the records. A further limitation of the data is that the version of the material (AAM, VoR, etc.) was not available; these data are not regularly recorded by the contributing repositories.

IRUS categorises material used from repositories into 31 item types¹², covering a very broad range of materials such as lectures, audio, maps, chemical structures, and so on. We looked at usage data for two longform item types: 'Book' ("A complete, stand-alone work published in one or more volumes, often identified with an ISBN") and 'Book Section' ("A defined chapter or section of a book, or anthology/collected work and usually with a separate title or number"). For comparison of longform with shortform content, we also looked at the data for 'Articles' ("An article on a particular topic and published in a journal issue.")

10 <https://www.jisc.ac.uk/irus>

11 https://irus.jisc.ac.uk/r5/support/guides/understanding_r5/

12 https://irus.jisc.ac.uk/r5/uk/about/policies/item_types/#current_item_types_policy

Dimensions

We used Dimensions AI¹³ from Digital Science to understand the volume of academic books being published, and the proportion of these books that are made available through different OA pathways.

Institutional repositories

As IRUS only monitors usage data, we looked at individual repositories to gather data on the number of items deposited. Examination of individual repository records allowed us to understand what users were actually seeing in the repository; for example, whether a book record in fact contained the full content of a book.

Publishers

Two publishers very kindly reviewed the IRUS data relating to their own titles and provided their own data to us for comparison. These data enabled us to better understand the content available in the repositories (e.g. whether a draft of some kind or the final version was available, whether the book had been published Gold OA, and whether the items were aligned with author contracts and publisher policies). We were also given 1 year of usage and/or sales data from the publisher's own platforms. Please note, however, that while a publisher may make the books and metadata available on their own websites, these platforms will not be the primary way the books are discovered, accessed, and used.

Limitations of the data

It is important to note that this quantitative analysis was limited in a number of ways, including:

- The supply chain for books is distributed. By this we mean that the full text of books and their metadata are distributed, as opposed to widespread use of redirection technologies guiding users to a primary version, as happens in the journal supply chain. Books are distributed to consumption and discovery partners such as aggregators, bookstores, distributors, library platforms, repositories, sales channels, and so on. A publisher may also make the books and metadata available on their own websites, but these will not be the primary way they are accessed, very visible, or greatly used. So too with repositories. The book supply chain is explained in detail in our 2025 report¹⁴.
- Books that are currently shared via a repository may be atypical of scholarly books as a whole (e.g. they may be more likely to be associated with research grants).
- The data relate to a wide range of disciplines, and differences between these may be masked.
- Usage patterns now may not predict future usage patterns.
- Our sample sizes were small because OA for longform, and especially Green OA for longform, is in its infancy. Any conclusions drawn were necessarily tentative.

13 <https://www.dimensions.ai/>

14 Estelle et al. (2025), 'Supporting learned society, subject association, and smaller specialist publishers'. <https://zenodo.org/records/14679068>

Volume of longform material

It is very difficult to know how many books are published each year, much less how many of these are academic books, books in scope for funder OA policies, books published OA, or books retrospectively converted to OA. Surprisingly, there is no central data source for this information, and different databases approach this challenge in different ways. The Dimensions database reported the total number of edited books and monographs published each year as around 89,000 in 2019 and steady at around 101,000 per annum from 2020 to 2023. However, this underestimates the numbers because the Dimensions database is more focused on journals than on books and does not include print-only resources. Only titles with DOIs assigned are included.

By any measure, the longform content with repository records represents only a small proportion of the titles available, and the proportion of those records with full text is even tinier. Book records for approximately 52,000 titles were accessed in 2024, along with 72,000 book sections. As will be seen later, longform content continues to be accessed long after publication, so these uses represent relatively small numbers of books published in each year. For example, just over 3000 records for books published in 2021 were accessed, and this represents the peak.

We initially took the description in IRUS for "Book" at face value; this was supported by a few random checks looking at the source records. However, when we cross-checked IRUS data with data from publishers, we discovered that in fact many of the items recorded as a "Book" did not include the full text of the book in question. A manual examination of 932 individual records showed that only 165 (18%) of the "Books" actually included full text. The remainder typically consisted of bibliographic data and sometimes an introduction, sample chapter, or cover image. Manual examination of 100 book section records selected at random revealed that just one record contained only a cover image; the remaining 99 were indeed book chapters. Four of these book sections had been embargoed.

Despite this drawback, we considered the data worth analysing; we were primarily concerned with user behaviour, and in general there is no way for a user to know in advance of accessing the institutional repository record whether the full text is available there. It seems likely that a typical user accessing a book record in a repository is looking for the full text rather than for bibliographic data, which is more reliably sourced elsewhere. It also seems likely that a typical user accessing a book section record in a repository is looking for the full text rather than a metadata record noting the text is embargoed.

The date of deposit for items is not usually given. However, we hypothesised that this date is closely related to publication date, which is readily available. Figure 1 shows that the volume of longform and shortform content both increased at a steady rate; the rate of addition of articles, however, was greater.

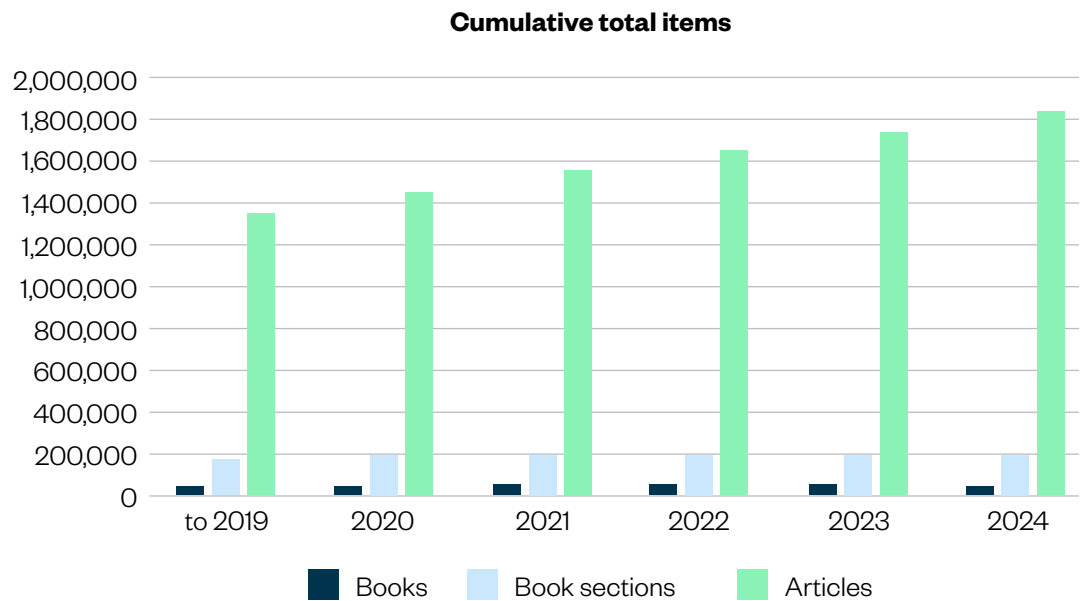


Figure 1: Cumulative numbers of items available for use in UK Higher Education institutional repositories (estimated by publication date).

Use of longform content

The use of articles in UK repositories is much higher than that of longform content.

Nonetheless, the use of longform content in UK institutional repositories grew steadily between 2020 and 2024, although patterns differed between books and book sections. Book section usage increased consistently from just over 2 million in 2020 to more than 3.5 million in 2024, while book usage was more irregular, probably due to the limited availability of full-text content—only 18% of book records contained full text, while almost all book section records included the full text of the chapters.

Use of journal articles grew faster than use of longform content over the period 2020–2024. This was driven by two factors: the growth in article deposits has been greater, and usage per item has grown more for articles.

Use was highly concentrated in a small number of repositories. For books, the top 20 repositories accounted for 60% of usage, with LSE Research Online alone responsible for over 25%. A similar concentration was seen for book sections, with the top 20 providing 70% of usage.

Access was international in scope. For books, the UK represented the largest share of usage (29%), followed by the USA (14%) and India (5%). As shown in Table 3, book section usage showed a similar pattern, with the UK at 25%, the USA at 14%, and India at 7%.

Older titles continue to attract strong interest. Both books and book sections displayed a long-tail usage pattern, with substantial engagement for works published many years earlier. Books tended to generate higher immediate use, while book sections saw slower but sustained discovery over time.

By discipline, 61% of books were classified as SHAPE fields, while 54% of book sections fell into STEM fields. Citation patterns were more evenly distributed, with both SHAPE and STEM longform content demonstrating significant scholarly impact.

Longform versus shortform use

Figure 2 shows the overall usage of longform content on HEI repositories over the period 2020–2024. It can be seen that use of book sections increased steadily over this period, from just over 2 million uses in 2020 to over 3.5 million uses in 2024. Usage of books was more random. The difference between the two may reflect the fact that for the latter the required content (i.e. full text) was very often not found, which would discourage future searches.

IRUS-aggregated Book and Book Section usage in HEI repositories

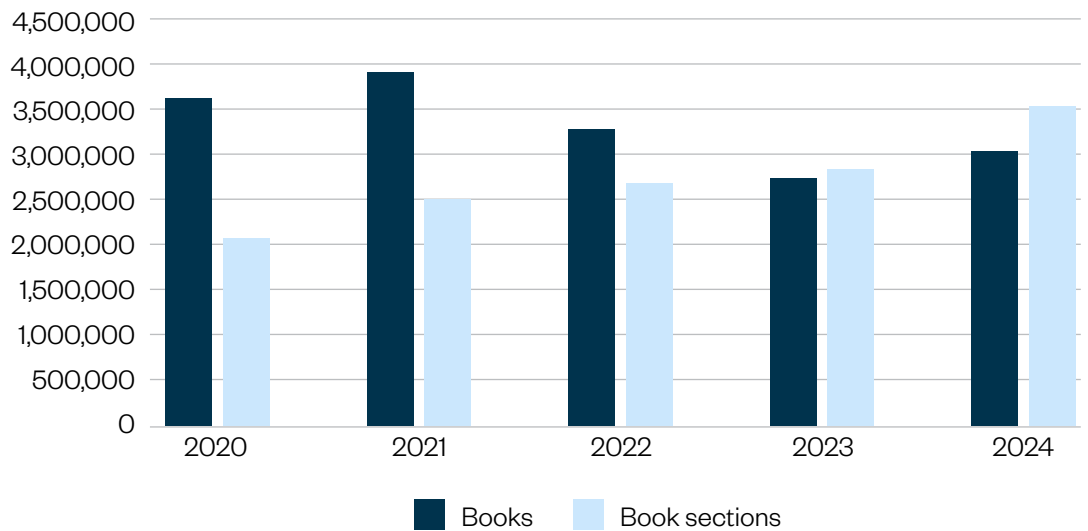


Figure 2: Book and book section usage in HEI repositories, 2020–2024.

Figure 3 compares the usage of longform and shortform content over this period. Article usage increased dramatically, at a rate greater than the increase in longform usage, which raises the question of whether this was driven by an increase in usage per item or by the increase in the number of items available for use.

Usage by document type

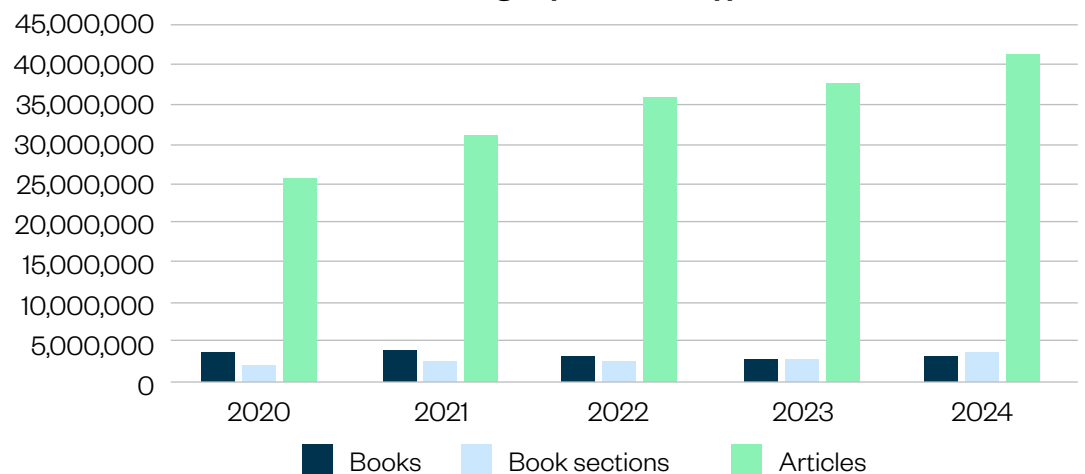


Figure 3: Longform and shortform usage compared.

Comparing the usage data with the number of items in the repositories allows us to see how the average usage per available item has changed over the past few years (Figure 4). Unsurprisingly, the number of (attempted) accesses of individual books was considerably higher than for individual book sections or articles. However, while the number of uses per item for the latter two types of content increased steadily over the period, the number of book accesses varied considerably with drops in 2022 and 2023. Again, it seems reasonable to draw a line between that finding and the fact that (in most cases) users were not actually finding the full text of books they were looking for.

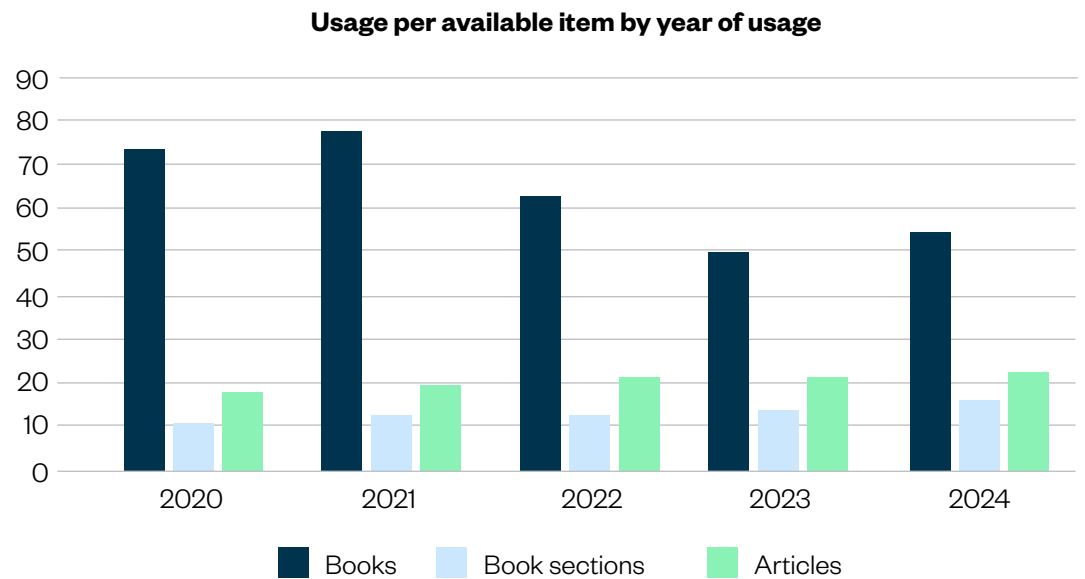


Figure 4: Average usage per item, 2020–2024.

Looking at both the number of items deposited and the number of uses per item, we can conclude that the pattern of usage seen in Figure 2 is driven in part by both factors.

Variations between repositories

There was a striking variation across the repositories in terms of usage.

For book records, the LSE repository recorded 4.2 million accesses (over 25% of the total usage), and the top 20 repositories accounted for 60% of the usage. The next 91 accounted for most of the remaining 40% of usage, with the remaining 60 driving just 0.9%. The pattern for book sections was similar, with the top 20 accounting for 70% of usage, the next 84 for 24%, and the remaining 60 repositories reporting just 0.6% of the total usage. Twelve repositories fell within both top 20s (Table 1).

Given the scale of usage from the LSE repository, we repeated some of the analyses here, excluding the LSE data in case they were skewed by this one repository; the results were broadly unchanged. We also dove deeper and created a case study about the LSE repository, which can be found in Annex 1.

Repository	5-year usage (%)	
	Books	Book sections
LSE Research Online	25.6%	9.1%
Manchester Research Explorer	13.5%	8.1%
UCL Discovery	7.5%	6.5%
White Rose Research Online	4.4%	6.5%
Strathprints: The University of Strathclyde	0.6%	4.1%
Online Research @ Cardiff—ORCA	3.5%	1.2%
City Research Online	1.1%	3.4%
ORO: Open University Open Research Online	0.7%	3.3%
SOAS Research Online	1.6%	3.2%
University of Southampton—ePrints Soton	3.0%	1.6%
SHURA (Sheffield Hallam University Research Archive)	1.7%	2.2%
Goldsmiths Research Online	0.7%	2.0%
Kent Academic Repository—KAR	1.3%	1.9%
Bournemouth University BURO	1.9%	1.5%
STORRE: Stirling Online Research Repository	0.1%	1.8%
e-space: MMU's Research Repository	0.2%	1.7%
Enlighten: Publications (University of Glasgow)	0.4%	1.6%
Open Research Exeter (ORE)	0.2%	1.5%
Royal Holloway	1.5%	0.3%
Edinburgh Research Explorer	1.1%	1.5%
Apollo—University of Cambridge Repository	0.4%	1.5%
Aston Publications Explorer	1.2%	1.5%
Oxford University Research Archive: ORA	1.0%	1.4%
Kingston University Research Repository	1.3%	0.2%
Explore Bristol Research	1.3%	0.7%
QUB Research Portal	1.1%	1.2%
Lancaster EPrints	1.0%	0.0%
Roehampton Research Explorer	0.9%	0.2%
Totals	78.6%	69.9%

Table 1: Institutional repositories generating the most usage for either books or book sections.

Usage by country

Table 2 shows the number of book records accessed by country across the 5-year period. Approximately 29% of usage came from the UK, with 14% from the USA. India was the third biggest user with 5%.

Usage by country: Books	5 year total	% of total
United Kingdom	4,842,089	29.4%
United States	2,345,172	14.3%
India	844,677	5.1%
Germany	508,458	3.1%
Philippines	418,390	2.5%
Australia	396,691	2.4%
Canada	358,916	2.2%
France	326,581	2.0%
Netherlands	315,382	1.9%
Italy	250,580	1.5%
China	249,080	1.5%
South Africa	208,745	1.3%
Ireland	201,339	1.2%
Pakistan	175,832	1.1%
Malaysia	171,316	1.0%
Spain	170,265	1.0%
Nigeria	168,290	1.0%
Indonesia	161,114	1.0%
Turkey	155,454	0.9%
Russia	152,107	0.9%
RoW	4,028,797	24.5%
Total	16,449,275	

Table 2: Book usage by country.

Table 3 shows the same breakdown for book sections. The top countries follow a very similar pattern: the UK accounted for 25%, the USA for 14%, and India a little less than 7%.

Usage by country: Books	5 year total	% of total
United Kingdom	3,339,875	24.7%
United States	1,886,272	13.9%
India	903,723	6.7%
Germany	425,378	3.1%

Australia	397,787	2.9%
Philippines	349,519	2.6%
Canada	337,970	2.5%
Netherlands	250,528	1.9%
South Africa	237,336	1.8%
China	236,699	1.7%
Italy	225,918	1.7%
France	209,514	1.5%
Pakistan	195,663	1.4%
Indonesia	164,737	1.2%
Ireland	155,823	1.2%
Spain	150,412	1.1%
Malaysia	148,109	1.1%
Nigeria	147,348	1.1%
Turkey	142,093	1.0%
Hong Kong	134,967	1.0%
RoW	3,493,988	25.8%
Total	13,533,659	

Table 3: Book Section use by country.

Individual title usage

We looked at the publication dates of book records accessed in 2024 (Figures 5a, 5b). These figures (and the following two) include data that are aggregated for the pre-2000 period. There are uses recorded for books to be published in 2025 that have presumably been deposited prior to formal publication. There is a long tail of usage, and while the uses per item tails off a little for the earlier content, it is still significant.

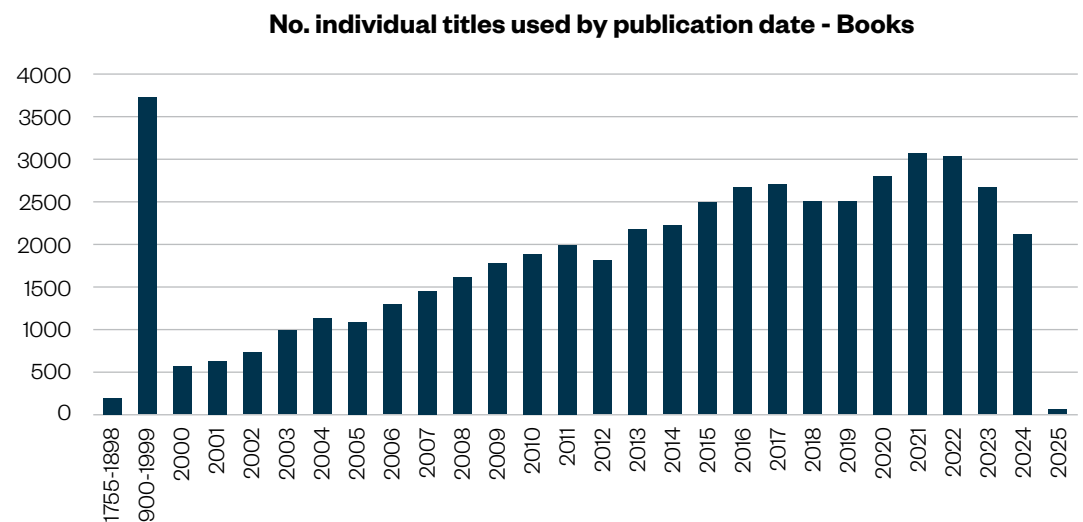


Figure 5a: Number of individual titles used in 2024 by publication date.

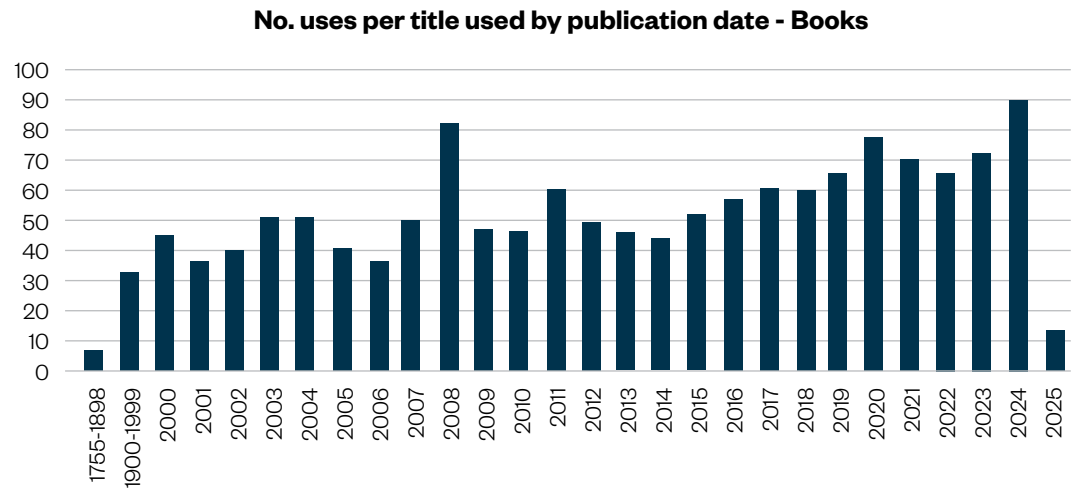


Figure 5b: Average number of uses per title in 2024 by publication date.

A similar pattern is seen for book sections (Figures 6a, 6b). There is more of a tail off of both the number of items used and the average use per item toward recent years, which suggests that complete books have a more immediate impact while book sections take a while to be discovered. Clearly, though, content in repositories remains useful for extended periods of time.

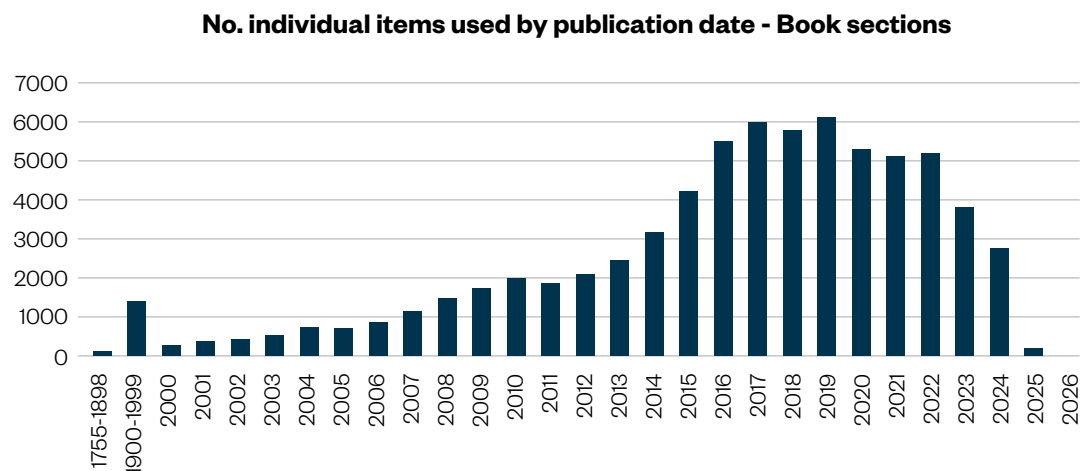


Figure 6a: Number of individual book sections used in 2024 by publication date.

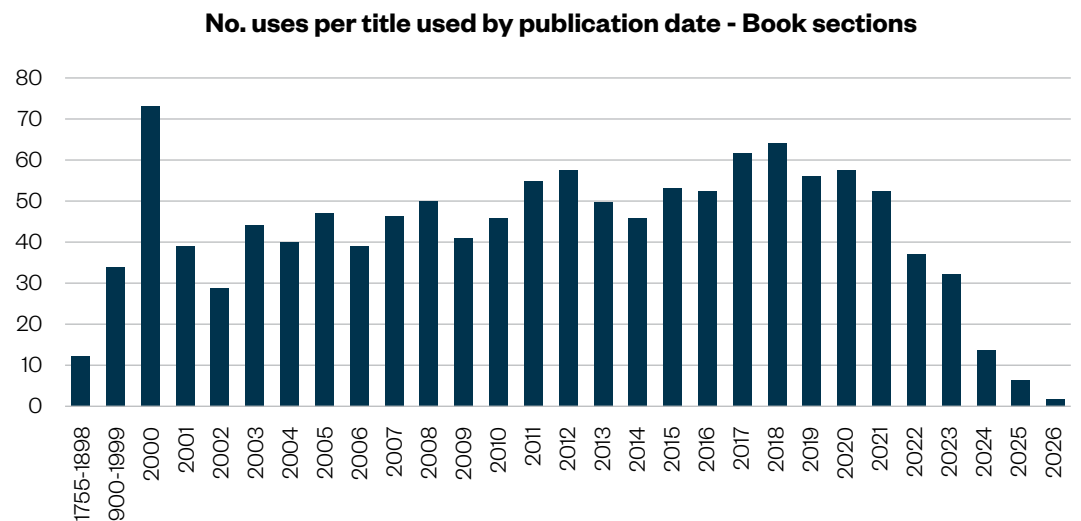


Figure 6b: Average number of uses per book section in 2024 by publication date.

Similar graphs were produced for the data examined in the section on individual publisher data (see below), looking separately at the "Green" and Gold titles. While the sample sizes were too small to allow a proper statistical comparison, there was nothing to suggest that there were any major differences in the patterns seen.

Repository content by subject

IRUS does not provide any breakdown of usage by discipline. We attempted to generate a breakdown using ChatGPT, employing several language learning models, including the most recent (at time of writing)—GPT-5. We selected 100 random titles, assessed each one manually, and then attempted to obtain a match for those assessments using AI. The matches were insufficiently accurate to justify using ChatGPT to analyse the whole sample. This was partly due to the AI's inability to make a reasonable assessment with the limited data, and partly due to the ambiguity of some of the titles.

Instead, we turned to Dimensions.ai¹⁵, a bibliographic database, to look at the breakdown of Green longform content (Publication Type = Chapter, and Publication Type = Edited Book + Monograph). Dimensions allows these to be categorised by the Australian and New Zealand Standard Research Classification¹⁶. Consolidating these into either SHAPE or STEM, we obtained the results given in Table 4.

Books		Number		Citations	
	STEM	52,082	39%	1,487,780	50%
	SHAPE	81,804	61%	1,515,606	50%
	Total books	133,886		3,003,386	
Chapters		Number		Citations	
	STEM	133,983	54%	2,677,190	72%
	SHAPE	115,075	46%	1,046,630	28%
	Total chapters	249,058		3,723,820	

Table 4: Green longform content listed in Dimensions.ai, categorised as STEM or SHAPE. For comparison, 52,048 individual books were accessed through HEI repositories in 2024, and 72,101 chapters.

Approximately 61% of books listed were SHAPE, but interestingly the number of citations driven by these books was split evenly between STEM and SHAPE. The number of STEM chapters (54%) exceeded the number of SHAPE chapters (46%). It should be noted that Dimensions.ai covers a range of sources, so these results are global and not limited to UK repositories.

Publisher data

A total of 647 items published by Publisher A and categorised as a book showed up in the IRUS Individual usage data for 2024. However, only a limited number of these proved to include full

¹⁵ <https://www.dimensions.ai/>

¹⁶ <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-research-classification-anzsro/latest-release>

text of the item in question. Examination of each individual record showed that the records comprised the following:

- 100 Gold OA books available in full text.
- 20 "Green" OA books available in full text. We have labelled these as "Green"; however, the publisher does not permit repository posting for these titles, so these books were shared without the correct licences and permissions in place. From the user's point of view, however, they would be functionally identical to Green OA titles. They comprised a mixture of AAMs and author proofs.
- 26 records which stated that full text had been deposited and embargoed. No embargo dates were given; it is possible that these were permanent embargoes.
- 8 titles that had been removed by the time the records were examined. Reasons for removal were not given, but it seems plausible that they may have been removed because of copyright considerations.
- The remaining 493 records turned out not to include full text. The records instead included basic bibliographic information sometimes accompanied by a cover image, a brief summary of the book, or a sample chapter.

Publisher A kindly supplied 2024 usage data for the "Green" titles from their own eBook platform, using COUNTER 5 Unique_Item_Requests so as to be consistent with the usage reported through IRUS. (It is important to note that these books are available from an array of other platforms, so this does not in any way represent the total usage). The publisher also kindly provided comparable data for a number of paywalled titles similar to the "Green" titles (matched individually by publication date and subject) and—where available—similarly matched Gold titles. A summary of the usage figures is given in Tables 5 (all titles) and 6 (only those titles with Gold OA comparison titles).

Full sample	n	Mean repository use	Publisher platform use	
			Mean	sd
"Green"	20	153	61	127
Paywalled	59		92	154
Gold	24		417	235

Table 5: Publisher A: summary statistics. NB: Books labelled as "Green" were shared without the correct licences and permissions in place as the publisher does not permit repository posting for books unless they are published Gold OA. From the user's point of view, however, they would be functionally identical to Green OA titles. They comprised a mixture of AAMs and author's proofs.

Titles with Gold comparators	n	Mean repository use	Publisher platform use	
			Mean	sd
"Green"	9	111	35	45
Paywalled	26		85	155
Gold	24		417	553

Table 6: Publisher A: restricted to titles with Gold OA comparators.

The available sample sizes were small, so any conclusions are necessarily tentative. However, we can make a number of observations, while noting that these don't reach statistical significance at $p < 0.05$.

Taking the 20 "Green" titles as a group, there were 3,060 uses in 2024 (an average of 153 per title) through HEI repositories. The publisher reported 1,218 uses for these same titles over the same period (61 per title)—only approximately 40% of the repository usage. While there was quite a lot of variation at an individual title level, repository use exceeded publisher platform use for 75% of the sample. Remember, these titles will have more usage on other platforms that are licensed by the publisher.

Looking at the paywalled control group ($n = 59$), the average number of uses per title on the publisher's platform in 2024 was 92—higher than the 61 per title in the "Green" group. This tends to indicate that availability of a free-to-access copy in a repository may draw usage away from the paywalled version on the publisher platform (although, as noted, the confidence in this conclusion is limited by the sample size and the majority of book usage will be on entirely different platforms).

As might be expected, the usage of Gold titles on the publisher's platform (a mean of 417 uses per item) greatly exceeded the usage of either "Green" or paywalled titles. (Again, there is additional usage of these titles from other platforms). Interestingly, the IRUS data for use of Gold items from Publisher A in repositories showed a mean usage of 96 per item ($n = 100$). Unsurprisingly, this is considerably lower than usage on the publisher's site.

A total of 285 titles published by Publisher B were recorded in IRUS as books showing usage in 2024. As with Publisher A, inspection of the individual records showed that only a small proportion of these records included full text. The records included:

- 21 Gold OA books available in full text versions.
- 24 "Green" titles available in full text, comprising a mixture of AAMs and author proofs. The publisher does not allow deposit of any form of full text for paywalled titles, so these files were posted without the correct licence or permissions in place. Again, we labelled these as "Green", as this best matched the users' understanding of their status.
- 23 records which stated that full text had been deposited and embargoed. No embargo dates were quoted, and it is possible that these were permanent embargoes.
- The remaining 217 records included only basic bibliographic information, sometimes accompanied by a cover image, a brief summary of the book, or a sample chapter, but full text was not available.

It became evident that establishing the correct formal status of the books isn't always straightforward. Several titles were stated in their repository records to have Creative Commons licences applied (or were presented on the repository website with an unlocked padlock OA symbol); however, the publisher confirmed this was incorrect. A couple of other titles were co-publications where the publishing partner had retained the right to make copies free to read and download online from their website (i.e. Bronze OA), but these titles would only be open on one co-publisher's site and not the other. Although the agreement did not extend to posting to other websites such as institutional repositories, it may be that this distinction had not been fully appreciated by the authors (or had simply been ignored).

Publisher B kindly supplied (all-time) sales data for the 24 "Green" titles. Repository downloads were highly correlated with sales of digital versions of the same titles (Pearson's correlation coefficient $r = 0.95$) and less so with sales of print editions ($r = 0.67$).

We hypothesised that the use of "Green" versions might be correlated with the price of the VoR—that is, a higher price might provide more impetus to look for a free alternative. This proved not to be the case—any correlation was weak and negative (i.e. to a small extent the cheaper books received more usage). As cheaper books are more likely to be aimed at a wider market, this negative correlation makes sense. The publisher offers three digital formats—XML, PDF, and ePub—priced differently; this result holds for all three formats.

Publisher B also supplied usage data from their own platform for the Gold titles that had received usage in repositories. For logistical reasons, the data supplied by Publisher B ran from May 2024 to July 2025, as opposed to Jan–Dec 2024 for the repository usage. Two titles had been recently acquired from a publisher without their own platform; these books had received significant repository usage (1072 and 286 downloads, compared with an average of 58 for the remaining 19 titles), indicating that HEI repositories may be a useful platform for Gold OA titles in the absence of a publisher platform. The publisher data were more granular than the repository data, but combining whole book downloads, title page views, and chapter uses gave an average of 359 uses per title, so HEIs contributed 14% of the combined usage—a similar pattern to that seen with Publisher A. (Again, it should be noted that the titles are available on a wide array of other platforms.)

Survey of institutional repositories

To build upon the interviews, and broaden our understanding, we issued a survey to explore how longform publications are used and managed in HEI repositories. It was circulated via a number of JISC Mail lists (e.g. glamlabs, goal, Jisc-repositories, lis-e-resources, and lis-link) as well as the UK user groups for DSpace, Fedora, EPrints, and Figshare. The aim was to gain insights into current practices, the roles of repository staff, and the types of institutions engaging with longform publications in their repositories.

There were 57 responses, with 35 respondents describing themselves as "repository administrator, curator or manager". Four responses were received from Library Directors, three from Data Librarians, two from Library Assistants or support staff, one from IT staff, and 12 from people who described themselves in a variety of ways including Digital or Electronic Resources Librarian (n = 2), Research Support Librarian (n = 3), Metadata Manager, OA Manager, and several other variations on librarian job titles. The majority of responses (n = 54) were from higher education institutions, with one respondent from each of a national library, a library consortium, and a consortium of HEIs.

A total of 47 respondents were based in the UK, with a further 3 from Belgium, 2 from the USA, and 1 each from Austria, Colombia, India, the Netherlands, and Sweden.

We asked which Jisc band the UK respondents' institutions fell into. There was a wide spread across the bands (Table 7). The 10 non-UK institutions were similarly diverse in size, with a wide range of staff numbers.

Jisc band	No.
1	2
2	1
3	6
4	7
5A	2
5B	3
6	7
7	4
8	1
9	1
10	2
Don't know	9
Total	45

Table 7: Jisc bands for UK respondents.

In most cases (44 respondents) the the institutional repository (IR) was managed in terms of day-to-day operations and decisions by library faculty or staff; 2 were managed by IT staff, 2 more by the Research Office, and 4 were managed jointly by the library and the Research Office.

Policy

We asked respondents to rank in order of importance a set of institutional objectives with regard to longform materials in their repository. Increasing the visibility of research outputs ranked most highly, followed closely by enabling authors to comply with research requirements and collecting institutional scholarship. Supporting open education and promoting the institution scored significantly lower (Table 8).

Objective	Average score
Increase visibility of research outputs	3.90
Enable authors to comply with funder requirements	3.52
Collect institutional scholarship	3.48
Support open education	2.19
Promote the institution	1.90

Table 8. Weighted score rankings for institutional objectives for longform materials in repositories.

One respondent added that compliance with the REF was the most important objective. An additional objective reported by three respondents was to provide a platform for content not published elsewhere:

‘Where a longform publication is not formally published elsewhere, we consider it “published” through the repository and have a policy to ensure this is treated seriously (i.e. compliance with the law regarding personal information, minting of DOIs, etc.)’

Just under half of the sample (21/45) had a formal policy regarding the deposit of longform publications. Approaches varied. Only four institutions mandated deposit of full text of longform materials; seven required faculty to notify the repository when they published longform works and create a record including bibliographic details; and nine simply encouraged deposit. One respondent noted that their rights retention policy applied to longform materials, and another reported that they actively sought permissions from publishers to post content:

‘We encourage researchers to create records for all their longform publications and we commit to actively seeking permission to make as much of the actual content Green open access as possible. We also provide guidance on Gold open access publication of longform content.’

Fewer than half (44%) devoted any resources to getting authors to deposit longform works. For those that did, the level of engagement varied. Provision of online guides and recommendations was standard, and some institutions offered some form of training. A number reported that they actively solicited deposits:

‘Weekly alerts are scanned and authors contacted where output has not already been self-submitted to the institutional repository.’

In a couple of cases, this had been automated:

‘[A] bot that searches web for OA publications and sends email to consider deposit.’

‘We have created an automatic upload into the repository for Thoth Archiving Network, part of Open Book Futures.’

Five repositories replied that they devoted resources only when deposit was mandated by the funder:

'Our staff facilitate the UKRI and Wellcome Trust open access policies and support and encourage our authors in depositing to Pure or requesting Gold open access funding from UKRI.'

Several repositories reported that they were not proactive, but responded to requests for help from authors. Four repositories commented that they did not have the resources to allocate to longform material:

'Our focus in our resourcing is on journal articles due to external policy requirement. We advocate longform deposit as part of general advocacy around research outputs management, but do not actively seek this.'

However, one noted:

'The drive to get authors to deposit journal articles results in many of them also depositing book chapters.'

Fewer than a third of repositories required authors to state whether a longform deposit was an AAM or VoR. Twenty commented that their repository staff undertook checks, ensuring that copies with the correct permissions were used. In general, authors were not expected to understand the details:

'Authors here typically have no idea whether they can deposit content to our repository until we inform them. Some, in contrast, blithely upload the VoR (or more commonly proofs) of book content because "that's what they're supposed to do".'

Just over a third of repositories (17/45) asked authors to state whether there was an embargo period. Again, this was more typically checked by repository staff, and authors were not necessarily expected to understand the requirements.

Discoverability, acceptability, and usability

We asked whether authors had reported any concerns about depositing their AAMs (Figure 7). Concerns that users would not be able to cite the book were most widely reported (41% of respondents said that authors had expressed these concerns). Formatting and presentation issues were close behind, and perceived credibility, access concerns, version control, usability concerns, and perceived prestige were all noted by significant numbers of respondents as issues that had arisen.

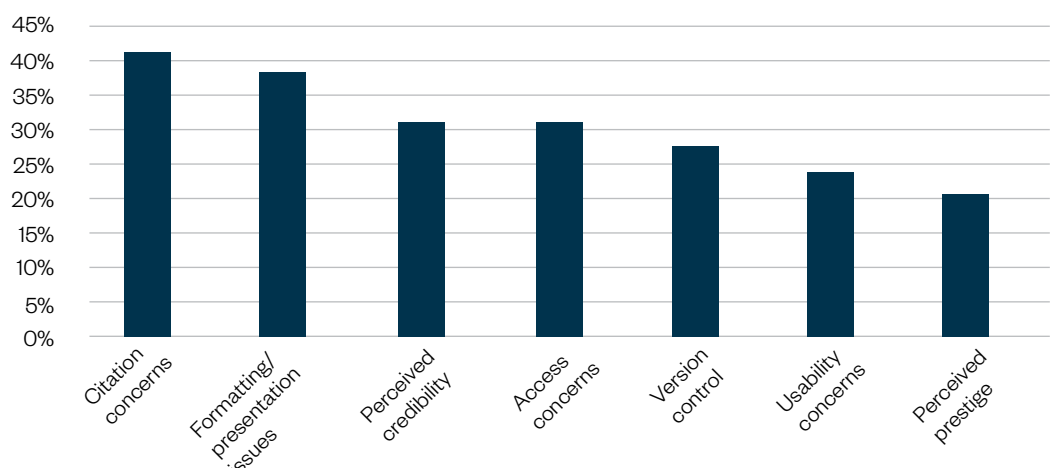


Figure 7: Reported author's concerns about depositing AAMs.

Six respondents also reported concerns about copyright issues:

‘Authors have had concerns around the inclusion of third-party copyright material, and how this can either be made open or redacted effectively.’

We asked how repositories enhance the discoverability of longform material in the repository—for example, did they enhance the metadata? Nine repositories did not do anything along these lines, while 12 reported that they checked metadata for accuracy and added missing metadata where possible. Additional actions, each reported by only one or two respondents, included:

- Adding ORCID¹⁷ data to identify authors
- Indexing in CORE¹⁸
- Linking related records (e.g. chapters within a single text)
- Tagging with Sustainable Development Goals
- Supporting harvesting through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)¹⁹
- Adding keywords (“Our keyword strategy is to add three keywords”)
- Citing the VoR in a header in the AAM
- Adding digital object identifiers (DOIs)
- Posting about the documents on social networks

A third of respondents thought that AAMs were less discoverable than VoRs, while 42% did not know (Figure 8). One respondent commented:

‘With good metadata, AAMs are more discoverable. This is because publishers do not make VoRs consistently easy to access on their websites and discovery platforms, even if they are open access. Repository content is much more easily discoverable by search engines than outputs on publisher book summary pages.’

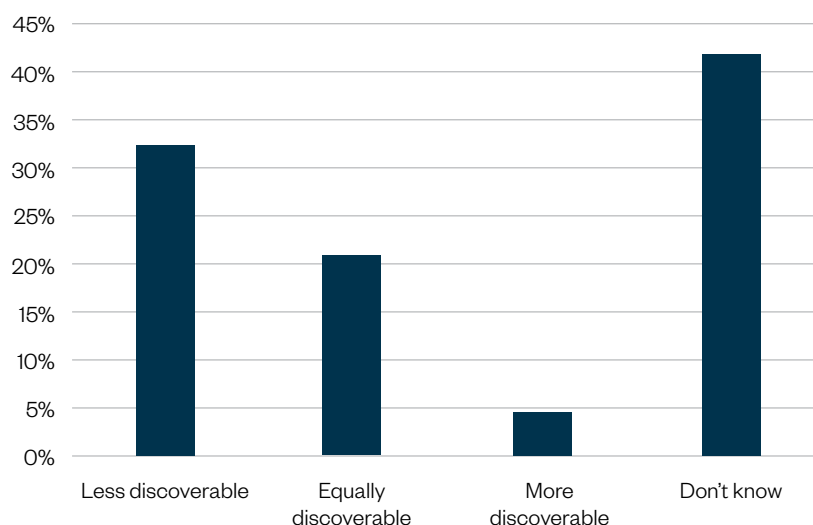


Figure 8: Are AAMs more or less discoverable than VoRs?

17 <https://orcid.org/>

18 <https://core.ac.uk/>

19 <https://www.openarchives.org/pmh/>

Only a few repositories had any feedback or evidence regarding the acceptability and usability of AAMs in the repository. One noted:

‘Generally staff are quite keen to upload and share AAMs where they can, and download stats have seemed quite good.’

Another reported a difference between users and authors:

‘There is anecdotal evidence from our librarians that students don’t know if they can trust and use AAMs. A small selection of academic staff were asked what they thought of AAMs and they all seemed happy to trust and use AAMs as much as they use the VoR.’

On the other hand, another commented:

‘There is concern, usually around images and image quality, dependent on the raw materials the author has or can deposit. Many would like their book to be used as a printed book would be, edited and laid out.’

And for one respondent, neither user nor author satisfaction was important:

‘Our main focus is REF compliance—if they state AAMs are acceptable, that’s good enough for us.’

Usage

Apart from beyond the standard geographical data, the majority of repositories (79%) had no data on who was using longform content (Figure 9). One repository commented:

‘Anecdotally, we receive enquiries on access from all of the above groups, albeit heavily weighted towards academia (UK and abroad).’

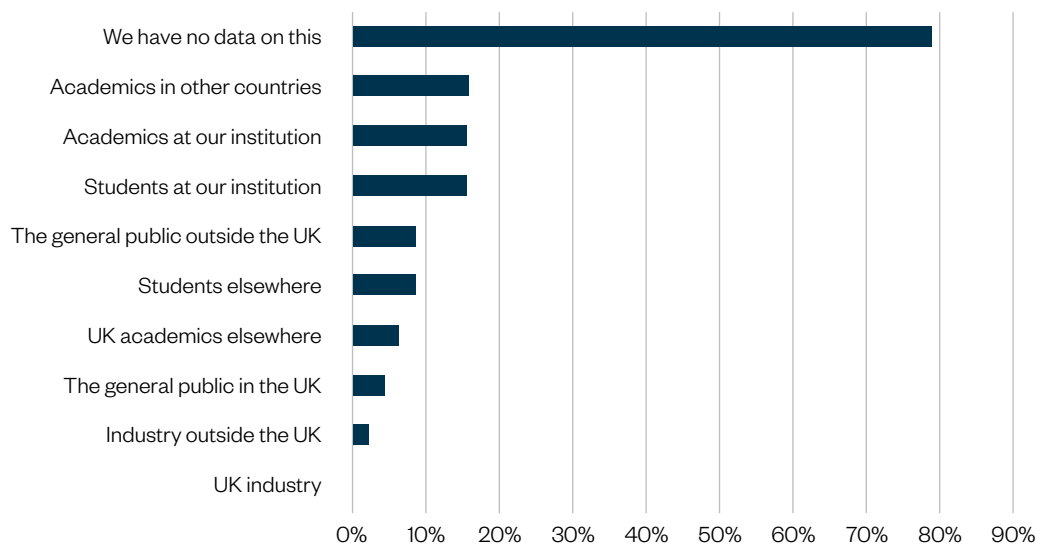


Figure 9: Who uses longform content?

A total of 30% of the sample reported that scraping of longform materials from their repository causes problems. Eight respondents reported that scraping activity had led to performance issues on the repositories (such scraping was general, and not specific to longform content). One commented that it can upset usage stats, and several more expressed concern about the use of material in AI training. Views varied; while noting that their authors might legitimately disagree, one commented:

'We in the library are generally persuaded that use of AAMs by large language models is legal and appropriate and (potentially) enhances the reliability of outputs from those models.'

The majority of respondents did not have any evidence about use of longform content in their repositories apart from the standard data forwarded to IRUS. There were, however, a few extra sources of data used by individual repositories:

- Altmetric data²⁰.
- Data derived from the Current Research Information System²¹ for comparison with IRUS.
- 'We have a stats package installed across the repository, and plugins which track use across the social web.'
- 'We integrate with Unpaywall.'²²
- 'There is a project starting this year to bring together IRUS-UK data with this repository data and a third database that records some additional information about the submissions.'

Resources

We asked whether repositories were able to quantify the administrative, infrastructure, and financial resources they currently required specifically for longform publications. A total of 32/36 replied that they did not; of these, only 8 were interested in developing this capability. Of the remaining 6, only 1 reported that they had detailed data, with the remaining 5 having estimates only (unfortunately, the repository with detailed data did not choose to share it.) A few respondents gave Full Time Equivalents (FTE) totals for all content types, ranging from 0.3 to 1.5 FTE, which they were unable to break down by content type. Three estimates were provided specifically for longform content: 0.05 FTE, 0.08 FTE, and 0.5 FTE.

Unsurprisingly, given these low resourcing levels, most respondents agreed that if the proposed OA policy for longform publications had been implemented for REF 2029, additional administrative, infrastructure, and financial resources would have been required. Eighteen (47%) felt that the additional requirements would be significant, with a further 13 (34%) saying that some extra support would be needed. Five were unsure, leaving only 2 who felt that their present resources would be sufficient.

Seven respondents expressed concerns that such a policy would require their institutions to fund BPCs or publish titles themselves if publishers were unwilling to authorise deposit of AAMs. Several noted that longform deposits required more resources than journal articles:

'This is mostly due to the time it takes to hunt down publisher policies. The process is also more drawn out so submissions are hanging around for much longer than shortform publications, and need monitoring/chasing by repository staff to ensure that they don't drop off the radar. Training our authors up on this new requirement would have taken a huge amount of time and staffing resource.'

'It is generally understood that processing longform materials for the repository would require more work than articles due to their nature and author's wishes (more individual and specific than an article).'

20 <https://www.altmetric.com/>

21 Current Research Information Systems which are databases used by institutions for managing and tracking research activity, such as publications, projects, and grants.

22 <https://unpaywall.org/>

Just over half the sample felt that dealing with such a policy would have been completely impractical (n = 4) or would present significant challenges (n = 15). Others thought it would be somewhat practical with modest effort (n = 10) or mostly practical (n = 4). Only 3 felt it would be fully practical with existing or slightly expanded resources. Several respondents would have welcomed the policy:

‘Book chapters could have been incorporated into the policy with relatively little change, given existing policies that the major academic publishers already have in place.’

Others were less positive:

‘There were several potential barriers to making this transition successful. A top-down approach would have been necessary to effectively communicate the change, but given how much our authors value the REF, it’s likely they would have engaged well with sessions on this topic. The most significant challenges for us would have been the financial and administrative burdens.’

Effect of book deposits on purchasing decisions

When asked: ‘Would an acquisition librarian in your institution purchase a book title if an AAM were available, e.g. in another institution’s repository?’, almost half of the sample (16/33) said it would have no bearing on purchasing decisions. Eight respondents said they would be less likely to buy a book if a Green OA version were available, while 9 thought that the availability of a Green version may help to preview the title and inform the purchasing decision:

‘The cost of the book would still inform the purchasing decision: if a Green OA version is available it may help us review expensive titles to decide whether purchasing is worthwhile.’

The answer may also vary according to the specific need:

‘This would depend to some extent on what the use of the book would be. If for research by an individual, we would refer to the OA version first as an option. If for teaching, we wouldn’t, as there can be specific teaching needs that are met within published longform materials.’

One respondent commented:

‘Of course we would be less likely to buy it *if* a fully OA version were available—although that relies on the purchasing librarian finding that version, which is not part of their current workflow.’

We also asked what factors might change that answer (Figure 10). Embargo periods were ranked highest, followed by policy changes; interestingly, the volume of books in repositories was considered less important.

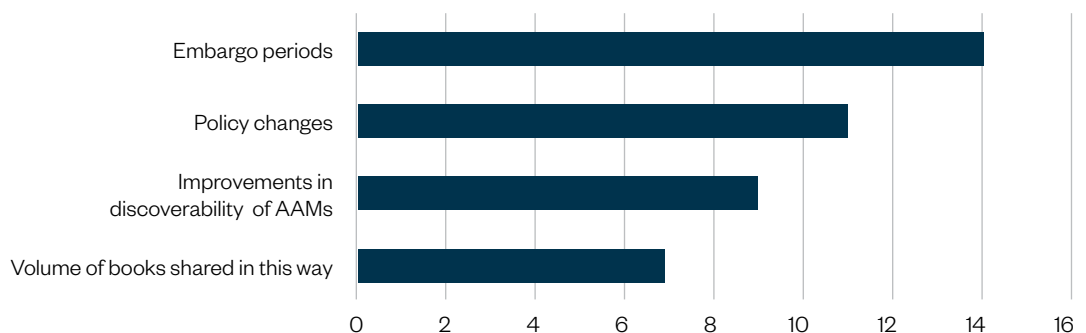


Figure 10: What factors might change your response?

Other factors mentioned included:

- Format of AAMs—whether they are accessible and usable by assistive technology such as screen readers.
- Reliability of AAM access.
- Discoverability via Google Scholar and LibKey Nomad.
- If the eBook is not available for solo purchase but only available online as part of some large and expensive (and unwanted) package.
- AAMs of longform outputs need to become the norm, and to do so we need a critical mass of availability, which can drive increased use.
- Budget.

Academic views

The primary focus of this project was on institutional repositories, but any consideration would be incomplete without an understanding of the views of scholars. We were interested in academic views both as authors depositing their own works into a repository and as researchers discovering and using materials available on institutional repositories.

It was not possible with the funding and time available to conduct a systematic and rigorous survey, and so our objective was to lay the groundwork for more systematic work in future.

We initially planned to conduct academic focus groups with:

- Academic users from several SHAPE disciplines, including early-career/established/retired researchers and independent researchers.
- Users in the galleries, libraries, archives, and museum sector with an interest in the SHAPE disciplines.

The lack of interest from most academics in even a conversation about Green OA for longform publications was, we believe, telling. Despite inviting participation on listservs, via libraries, and through the British Academy networks, we found it extremely difficult to recruit academic participants. In the end we held one online focus group and had conversations with members of the Academy's Open Access Advisory Group and an additional Fellow. We were surprised by the number of participants at the online focus group who did do research but also had some sort of administrative role with an institutional repository or who were tasked by their institution with a role in ensuring compliance with funder policies. It was these administrative responsibilities that motivated their engagement with our focus group.

Sharing via repositories

Participants were asked about their experience sharing materials via institutional repositories. There were some differences in perspective depending on whether the individual also had a role in championing or delivering a repository service. For example, those with library roles had insight into how publishers influence repository content and use. There was also quite a lot of discussion of other ways of openly sharing scholarship that did not involve an institutional repository.

'Limited repository experience. Don't support Green OA as it's a worse form of access—institutions often require deposits, but I'm never sure that anyone really engages with them. The major exception to this is PhD theses, in which case this ethos is essential, especially since the British Library's central service is not working at the moment.'

'Not as an author, although I have shared journal preprints; however, as a repository manager, I try to help colleagues obtain access to content via repositories if we cannot afford it otherwise.'

'In Business we tend to publish more in journals than books. ResearchGate and Academia provide good access, so repositories are not the only way that we can openly share our papers, and the other ways are easier.'

'Author satisfaction is higher when the final version of a book can be shared. Uploading to repositories, and therefore the motivation to upload to repositories, is even worse for longform than for articles.'

'I work for a doctoral training partnership, and we get a lot of questions from PhD students who ask about putting their theses in repositories. We encourage this in case they never go on to publish books, so their research is still available. I am less familiar with repositories; perhaps this leads to duplication if they do publish. I myself have contributed to edited collections, but these tend to be less visible if they go into a repository than putting my research into a journal and then having it discoverable via ResearchGate.'

'When I finished my PhD, many of us were concerned about the duplication of material, yet this dissipated when I went on to write a book as part of the editing process focused on evolving the book to make it distinctive from the theses. Having my thesis on the repository actually helped me get the book contract because it made my research more visible. Humanities Commons is really fantastic, as I can share material (e.g. my recent book chapter) even though I don't have an institutional affiliation. Where we are in terms of the policy landscape, it is still early on re books. DOAB contents don't seem very discoverable via library discovery tools.'

'I am cautious about sharing my research on repositories because my ideas could be used by others before I have a chance to publish, and it might undermine my chances of getting published.'

'A lot of people are worried about their ideas being stolen and make use of embargoes. I did that with my thesis because I was worried too, even though it seems somewhat irrational now.'

'There should be a lot of consideration given to the impact on career development. Simply transposing STEM policies to SHAPE isn't appropriate. We work on our theses for 3-6 years, and there should be protection as we need additional time to continue forming our ideas and an opportunity to secure a book agreement and then to publish.'

Discovery of longform materials

Participants were asked about their experiences discovering longform materials in institutional repositories. Materials in repositories were not perceived to be very discoverable; instead, repositories were searched individually to find specific works when other access options had been exhausted.

'Institutional repository contents are not very discoverable via a Google search, so you cannot find them if they are even there.'

'When we set up our IR, we worked hard in terms of search engine optimisation. For a while, most of the articles on the repository were higher ranked on Google than the actual published version. So, it is possible to do that.'

'Google is unlikely to send me to a repository, so I rarely ever use one. Recently, I was sent to a repository, but when I tried to download the item I found that only authorised users could. I don't get the purpose of these repositories if only some people can access them.'

'Normally I can find longform publications, particularly theses, but I can rarely find relevant books or book chapters on repositories.'

'If I find myself navigating to a repository, it is via an old-school traditional trail: following a citation, hitting a dead end, and then doing a web search to find out which institution the author is located in. This is just not how readers behave anymore.'

'Visibility to search engines is critically important. It's not clear if institutional repositories have really thought about, for instance, their visibility through Google search.'

'Repositories support a slightly outdated search strategy, in most cases. Large language models allow us to discover paragraphs of text, and institutional repositories are not plugged into this discovery channel.'

'I am involved with societies that put materials on a platform called Concordance. I don't know if that is a repository, but it is like Humanities Commons, and so it has curated collections around a specific academic subject. I don't care where research is—if it has got the information, then I want to be able to access it.'

Using longform materials in repositories

Participants reflected on their experience using longform material discovered on institutional repositories. The formatting of final versions of book and book chapters discovered on repositories was viewed as superior to the experience of using draft versions, but any form of access was appreciated if no other route to the full text was available.

'Repositories are better than nothing, but they contain the ghosts of real publications.'

'In the Global Majority, what I like to call the Global South, a lot of the central universities will lack access to many SHAPE materials. I have used a lot of items in the LSE repository [before I moved to the UK].'

'With the current infrastructure, I don't think repositories are very easy to use. It needs to be easier for scholars to search for papers.'

'It feels like institutional repositories do not really help with finding readers. Nor do they make the content look and feel nice. Without these things, they are not going to be positioned to challenge the dominance of publishers.'

'The problem with longform author's accepted manuscripts (AAMs) is that page formatting is different, and so to reference the works, you still need to find the published version. It means that you can browse and read through an IR, but if you want to reference a book, you will have to purchase it. Thinking about readers in the Global South, I'm not keen on an approach where we invite them to read a poorer version. It preserves the two-tier approach. As an author, I want to engage with my readers. Repositories are better than nothing, but they don't facilitate engagement.'

'I'm less likely to read a huge block of text. I like things to feel like a book. Reading is hastened with proper paragraphs and layout. It signals that you've really worked on the ideas and worked toward a longform book.'

'I want my work to speak for itself, so I'm happy if it can be accessed. I don't mind if people get Word documents in a repository, as I back my content to deliver on its own.'

'Publishers are often more lenient with what is permitted to be uploaded to personal websites rather than repositories. There is perhaps a bit of power in this, should someone wish to be less limited with regards to what they share. I recall a friend who works in physics not being at all fussed about embargoes because he knew that he could share his work more freely on his website. I appreciate that personal websites don't have the same potential discoverability and preservation benefits as repositories.'

Future outlook

Some participants believed that institutional repositories might evolve to become more relevant if discoverability and usability were improved, and if they could develop into community environments for supporting new forms of scholarship and facilitating engagement between authors and readers. There was some scepticism about whether this would be a sensible strategy, or even in an institution's control.

'If the REF drove a lot more institutional repositories Green then the driver is coercion and it decreases author motivation to publish open access. Not ideal. More money should go into funding for Diamond or Gold OA.'

'If repositories are curated in such a way that everything is set out nicely, is easy to click through, easy to link back to the source, et cetera, then IRs could be useful. At the moment, there are a lot of different ones—they are a bit disparate.'

'There is only so far it will make sense to push on discovery because, like it or not, library systems are just not where academics are searching.'

'The current library standards are not designed for practical academic use. Not just the content available, but also the inbound and outbound links.'

'Rich links between works, and info about the authors, related datasets, et cetera, might be a strength of repositories in future. This is a strength of Zenodo, for example, which links to the author and their institution and a whole raft of other things. It's the rich network that also makes this platform more friendly to search engines.'

'I think IRs will get better and be better integrated. This will, however, only come from engaging more with the commercial reality of publishers. And not all libraries (even most) will be able to support these sorts of higher functions. However, I am optimistic about improving discoverability, but at the cost of loss of authorial and institutional control without enormous legislative change.'

'Institutional repositories are a symptom that something is really wrong with the current system, rather than a solution in and of themselves.'

At times during the discussion, there had been an implicit comparison of institutional and subject repositories, and so we asked about this explicitly. Views were mixed:

'In terms of actual subject repositories, I don't really know them on the same level. I just know Humanities Commons and Knowledge Commons, etcetera. I just come across them as and when I'm doing things. I love them; they speak to the way the humanities work.'

'Subject repositories can work as an approach with a clearly defined subject focus, but I do cross-disciplinary browsing and reading in a subject area such as Business and so subject repositories are not so helpful.'

'Subject repositories as a future focus for community? Yes, we would be more aware these existed and more motivated to deposit. A subject focus would create loyalty and a sense of belonging. Might be more drawn to search there too. I'm not convinced that subject repositories would be better than Academia.edu.'

Analysis

We did not uncover a clear, coherent strategy for longform OA, much less longform OA in institutional repositories. This is perhaps unsurprising, as in his seminal work in this area Geoffrey Crossick noted that there was no particular crisis that needed to be solved and in many ways discussions around OA for books are still entangled in discussions about digital for books. He wrote that:

'finding successful business models for OA publication of monographs is the greatest and most important challenge. There is, after all, no such thing as free open access. 'Author pays' (usually the author's funder or institution) has become the key model for Gold OA journal articles, but the much higher costs of OA books and the lower funding for these disciplines make it hard to see how author pays could work were OA to be mandatory.'²³

Despite a great deal of activity in the last 5 years, it appears that little has changed. There is gridlock in the transition to longform OA publishing without strategic and practical alignment across authors, employers, funders, libraries, and publishers. Significant changes are needed in funding and infrastructure²⁴. Most universities and libraries do not yet have a dedicated longform OA strategy. Where longform policies exist, they are extensions of journal-oriented frameworks that are not always fit for purpose, especially in the SHAPE disciplines. Institutions take a "softly softly" approach to implementing policies for fear of straining author–publisher relations.

Publishing longform outputs OA is clearly the preferred approach but Gold OA is seen as prohibitively expensive, with charges ranging from £5,000–£20,000 per book. This far exceeds available funds. Some universities are experimenting with new university and library presses and alternative community-driven funding models, but the number of books published OA in these ways remains small. Many institutions face budget cuts, especially in libraries, reducing flexibility for new OA initiatives. Funding is not available, and is unlikely to be, to support the full spectrum of library aspirations to provide content and reading services that rival those of long-established commercial publishers.

Sharing via institutional repositories is seen by libraries as perhaps a less desirable, and possibly more pragmatic, way to achieve OA for longform outputs. Some aspire for institutional repositories to take on functions of the broader publishing ecosystem. There is a notable disconnect from scholars here. Authors dislike uploading accepted manuscripts, arguing these are inferior to final published versions. There is resistance from scholars, especially within SHAPE disciplines, where longform publications remain deeply entwined with prestige and central to community engagement. Sharing the final version of a longform output resonates strongly, but sharing an inferior version does not. There is support in principle for access to all to be available as a last resort and if no alternative avenues are available for it. There is not support for the level of investment that would be required to populate institutional repositories with all longform outputs, to integrate repository infrastructure into researcher workflows, to support rich reader engagement, and to support extended reading and community engagement.

23 Crossick, G. (March 2016), 'Monographs and Open Access', UKSG Insights, Volume 29, Issue 1, pp. 14-19. DOI: [10.1629/uksg.280](https://doi.org/10.1629/uksg.280)

24 Estelle et al. (2025), 'Supporting learned society, subject association, and smaller specialist publishers'. <https://zenodo.org/records/14679068>

It is not clear whether the purpose of longform OA via repositories should be to provide a cheap and cheerful access mechanism of last resort with no bells and whistles, or an all-singing, all-dancing environment that will rival the services of even the largest publishers. It is not clear where the funding would come from to support the first of these objectives, much less the second

Established book publishers do not appear to have been constructively engaged as potential strategic partners in this discussion, despite the alignment of their interests with author interests. This is perhaps in part because of an understandable loss of trust within the library community toward large commercial publishers around OA generally, but there needs to be more effort and will from libraries to engage with mission-aligned publishers on a strategic rather than a transactional level. If a win-win solution can be found that suits authors, employers, funders, libraries, publishers, and readers, then everyone self-evidently benefits. Where are the strategic conversations to explore whether such a world is possible?

And if we are wrong and this has been tried to no avail, then it is time to compete with those publishers. In addition to an ecosystem of well-funded new university and library presses and institutional repositories, it may also be time to invest in powerful competition—a sort of PLOS (Public Library of Science) for books—that can help change the game.

Without strong support from authors in particular, it will be difficult to populate repositories and negative incentives such as mandate compliance can backfire. Until the strategy for longform OA is really clear, and the funding for it secured, then the myriad problems with current practice will remain unaddressed.

When we asked library interviewees to imagine future policy and funding possibilities, they articulated the need to think more freely. However, their ideas were contorted around current funding and policy rules and limitations, and were sometimes contradictory. Their aspirations soared far above the material ground they stand on, leaving a gap between what they believe should be possible and what their budgets allow. A REF policy for longform OA was seen as the best, possibly only, way to cut through the Gordian knot. A future imagined in full colour, financed in greyscale, and contested by key stakeholders may provide difficult to realise.

Conclusions

We were impressed by every individual we worked with on this project, and the evident passion for OA. However, the bottom line is that without a clear strategy and sufficient funding, no amount of effort and professionalism is going to make Green OA for longform work via HEI institutional repositories.

We found that:

Most universities lack a dedicated longform OA strategy. Existing policies, developed for journal articles, are inconsistently extended to books. Most focus on encouraging academics to voluntarily embrace longform OA and on supporting case-by-case compliance with funder policies.

OA publishing is viewed by many of the stakeholders with whom we engaged as the preferred approach, but there are problems. Gold OA is viewed as financially unsustainable, with BPCs averaging £10,000–£20,000. Institutions are experimenting with community-led and Diamond OA models, but budgets remain constrained and long-term funding mechanisms are unclear. Green OA is perceived as a less desirable way forward.

Many academics, especially in SHAPE, view Green OA as a “last-resort” access route rather than a valued dissemination route. Authors support the principle that everyone should have access to their book or book chapter. However, they are much more positive about Gold OA and are often reluctant to deposit accepted manuscript versions in repositories, perceiving this avenue to OA as potentially harmful to professional reputation and inferior in format and reading experience to published versions. For early-career researchers especially, aligning with well-known publishers is still seen as critical for career progression, tenure, and professional recognition, creating an incentive for Gold OA and disincentive to deposit a final manuscript in a repository.

Repository environments do not provide strong support for the engagement and extended reading and annotation that SHAPE scholars desire. Institutional repositories aspire to take on functions of the broader publishing ecosystem, but without sufficient funding this ambition risks creating fragile, fragmented, and unsustainable services. Most institutional repositories are optimised for journal articles, not books, which bring specific challenges including complex formats, metadata inconsistencies, and copyright complexities. It is nearly impossible to meaningfully track how publishing and sharing practices for longform materials are changing over time.

Cross-sector collaborative dialogue is essential. As the UK higher education sector navigates a period of acute financial constraint, it is vital that all stakeholders come together to explore sustainable approaches to longform OA. Bringing together universities, research funders, libraries, learned societies and other publishers, authors, and readers can help to identify the models that both strengthen scholarship and ensure equitable access to knowledge. A rapid push to extend Green OA mandates, or one lacking adequate funding for institutional repositories, risks undermining the current progress made toward building sustainable Diamond and community-supported OA publishing models.

Longform publishing is an international activity. The publishing model for monographs relies on international funding and is not sustainable without this. International collaboration is essential. Policy development offers a timely opportunity to reinforce the UK's leadership position in SHAPE research and in academic publishing, and to maximise impact and return

on investments. By aligning collective interests, it is possible to design policies and services that safeguard research quality, expand readership and engagement, and bolster economic resilience.

We conclude that Green OA for longform works is contested and not yet mature in practice. There is evident support for openness in principle, but without a strategic framework and cross-sector coordination progress will remain slow and fragmented.

Acknowledgements

British Academy Open Access Advisory Group and Officers

Professor Lindsay Farmer FBA, University of Glasgow
Professor Margot Finn FBA, University College London
Mr Richard Fisher, Formerly Cambridge University Press and Royal Historical Society
Professor Simon Green, University of Salford
Professor Jane Ohlmeyer FBA, Trinity College Dublin
Mr James Rivington, The British Academy
Dr Catherine Souch, Royal Geographical Society
Ms Destina Thompson, The British Academy
Dr Joanna Thornborough, The British Academy
Professor Chris Wickham FBA, University of Oxford

Focus Group, Interview, and Workshop Participants

Dr Theo Andrew, University of Edinburgh
Professor Katherine Brickell, King's College London
Dr Christopher Bundock, University of Essex
Ms Veronica Cano, University of Edinburgh
Ms Suzy Cheeke, University of Bristol
Ms Rebecca Clarke, Queens University Belfast
Dr Harry Dimitropoulos, OpenAIRE
Dr Richard Espley, Victoria and Albert Museum
Ms Emma Francis, University of Aberdeen
Dr Alice Gibson, The Francis Crick Institute
Dr Rosie Higman, The London School of Economics and Political Science
Mr Neil Jeffries, University of Oxford
Dr Jade Jenkinson, University of Nottingham
Mr John Kaye, Jisc
Professor Geoffrey Khan FBA, University of Cambridge
Professor Petr Knoth, The Open University
Ms Bethany Logan, University of Sussex
Ms Ruth Mallalieu, University of Oxford
Ms Rachel Persad, GuildHE
Dr Devina Sarwatay, City St George's, University of London
Dr Hui Situ, Cardiff University
Ms Isobel Stark, University of Southampton
Ms Niamh Tumelty, The London School of Economics and Political Science
Dr Steven Vidovic, University of Southampton
Mr David Watson, University of Oxford
Dr Chris Wingfield, The University of East Anglia

Publishers

We would also like to thank our two anonymised and very helpful publishers who kindly shared confidential data with us for this project.

Annexe 1 – Case study: London School of Economics and Political Science (LSE) Institutional Repository

Our analysis of usage data across UK HEI institutional repositories showed that usage, both domestic and international, of longform material from the LSE repository was much higher than any other. This was partly due to the quantity of longform material available in the repository, but also a result of increased usage levels per item. We re-interviewed staff from LSE to try to determine why this might be. The two interviews disclosed several key areas where LSE distinguishes its institutional repository from others, especially in the context of longform deposits.

LSE stands out in its usage, and it is beneficial to understand why. Therefore, we present this as an example rather than a model for other UK higher education institutions to follow.

First and foremost, LSE has a clear institutional strategy. The institution's leadership is committed to advance and support OA generally and the LSE repository specifically. LSE's OA policy now explicitly includes monographs and is supported by a rights retention framework and a pilot fund for book-processing charges (BPCs). In addition, the institution provides funding for BPCs directly rather than through the library budget. It was emphasised that this distinction is significant, as it signals a top-level institutional endorsement for OA and removes financial barriers that often discourage faculty participation. Another key component of the LSE OA ecosystem is LSE Press, which publishes high-quality OA books and journals that are automatically included in the repository. This partnership between the press, the library, and senior management creates a coordinated and sustainable model that advances open scholarship.

Second, LSE's disciplinary focus and international status are key differentiators. LSE's international status as a leading social science institution has played a major role in the success of its institutional repository. The repository aligns closely with the research culture of the institution, offering a natural home for social science scholarship and reinforcing LSE's global identity as a centre for research addressing social, political, and economic issues. Many LSE scholars deliberately embrace OA as a way to reach and engage the communities they study, particularly in regions such as South Asia. Analytics reveal that repository users often arrive through targeted searches and citations rather than browsing, indicating that LSE's OA outputs are integrated into scholarly and policy-related networks. Because much of LSE's research is policy-relevant and intended for non-academic audiences, including NGOs, government agencies, and international organisations, the open access model broadens readership and amplifies public impact.

Third, unlike many institutions where self-deposit is expected, the LSE Library plays an active role in filling the repository. Researchers send their files directly to the repository team who handle permissions, metadata creation, and the deposit process. The library also proactively identifies new research outputs using databases such as Scopus and Web of Science, and then contacts authors to obtain copies. LSE has managed to integrate OA books into its existing repository workflows with minimal additional staff time or new technical infrastructure. Most of the effort lies in policy development, communication, and cross-department coordination rather than in costly platform investment. The repository is currently transitioning from an internally hosted to an externally hosted system, reflecting a pragmatic approach to long-term sustainability.

Despite its successes, LSE faces challenges familiar across the sector. Some researchers express concerns about the implications of OA for relationships with commercial publishers, fearing that OA requirements could limit future publishing opportunities. Others are wary of how AI models might use OA content without proper attribution or compensation. Metadata is also a perennial challenge. The LSE repository hosts several working paper series and many reports. These are assigned the item type 'monograph' and are therefore included in usage data for longform materials. Other repositories may describe similar deposits in other ways.

About the British Academy

The British Academy is the UK's national academy for the humanities and social sciences. We mobilise these disciplines to understand the world and shape a brighter future. From artificial intelligence to climate change, from building prosperity to improving wellbeing - today's complex challenges can only be resolved by deepening our insight into people, cultures and societies. We invest in researchers and projects across the UK and overseas, engaging the public with fresh thinking and debates, and bring together scholars, government, business and civil society to influence policy for the benefit of everyone.

The British Academy
10-11 Carlton House Terrace
London SW1Y 5AH

Registered charity no. 233176

thebritishacademy.ac.uk

Published March 2026

© The British Academy. This is an open access publication licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported License

[10.5871/hepolicy/he-repositories-green-oa-longform-publications](https://doi.org/10.5871/hepolicy/he-repositories-green-oa-longform-publications)