AGEH consulting

THE PRESERVATION OF KNOWLEDGE IN THE DIGITAL AGE

An analysis of public and research libraries' perpetual access to licensed electronic literature.

September 2024

Tony Ageh
Michael Bayler
Chris Durlacher
Julian Turner

Prepared for the Arcadia Fund by Ageh Consulting

Contents

Original Commission

Proposal for provision of a report in relation to an analysis of public and research libraries' perpetual access to licensed electronic copies of scholarly books

Arcadia is a charitable foundation which works to protect nature, preserve cultural heritage and promote access to knowledge. Arcadia's interest in open access to knowledge includes the ability of members of the public to access online, without charge, scholarly books (mono-graphs and edited volumes), and the ability of libraries to provide such access in respect of out-of-copyright scholarly books held as physical copies in their holdings through digitisation. Arcadia also supports efforts by libraries to provide online public access to scholarly books held as physical copies which are within copyright but out-of-commerce.

As many academic and public libraries shift from purchase of print copies of scholarly books to the licensing of electronic books (e-books) from publishers, Arcadia is concerned at the future impact on provision of access to out-of-commerce and, in time, out-of-copyright copies of these works.

In order to understand the scale and extent of this issue, and potential remedies, Arcadia wishes to commission a research report which will determine whether there is a meaningful public domain in e-books and, if so, how it can be enforced and what mechanisms are need-ed to ensure that libraries possess the same or better opportunities to provide public access to out-of-copyright e-books as they do for physical copies.

In particular, the report will cover:

> **1.** A survey of the kinds of licensing agreements which currently govern academic and public library access to scholarly e-books, and what – if any – restrictions they impose on libraries' future access.

> **2.** A summary of technical and legal issues which may be imposed or arise naturally which constrain either the survival of e-books into the public domain (such as format compatibility) or which prevent public domain criteria being fulfilled (for example, a publisher's electronic version which is continually updated and 'republished', without a static copy being created that can be deemed out-of-copyright).

> **3.** A survey of some of the provisions which libraries, publishers or third parties (such as dark archives providing escrow services) currently make to ensure that e-books are preserved once licenses lapse and before works enter the public domain, highlighting areas of best practice or innovation.

> **4.** Recommendations for future actions by libraries, publishers, and other organisations to ensure that libraries possess the same or greater opportunities to provide public access to e-books once they enter the public domain, highlighting any areas where funders such as Arcadia might take action.

The report should cover research libraries and public/national libraries in the US, EU and UK, identifying where issues and solutions are common to all jurisdictions and where differences occur.

Executive Summary

As academic and public libraries shift from owning, storing, and controlling access to printed books and journals to the licensing of electronic versions, this report examines the impact on the preservation of scholarly publications and how access will be sustained as digital works fall out of commerce and, eventually, out-of-copyright.

The authors initially set out to determine the range, volume and types of publications that fall within a practical definition of scholarly books, recognising there are significant variations in content, style, format, audience, and access modes that present differing challenges for long-term preservation and future access, potentially requiring distinct solutions. At the same time, the working definition needed to be broad enough to encompass obvious similarities across all publication types, allowing for common observations and shared objectives.

Although the following framework is not explicitly referenced throughout the report, we observe that the approach to the long-term preservation of scholarly e-literature varies significantly depending on two factors: subject matter – whether Science, Technology, Engineering, and Mathematics (STEM) or Arts, Humanities, and Social Sciences (AHSS) – and the publication model, which we classify into four categories:

1. Journals and serials published under commercial licenses
2. Scholarly monographs
3. Academic and trade books
4. Works published under Open Access models

We analyse and assess the various preservation arrangements for STEM and AHSS works in each category presently in use across the sector, the criteria for determining which approach to take, currently being applied by libraries and archives, and the implications that these arrangements may have on the future preservation of knowledge. A Red/Amber/Green 'traffic light' assessment of preservation activities is provided in Appendix E.

We find that STEM publications, particularly as commercially published e-journals, are generally considered at lower risk of loss due to the quantitative nature of their content, specialised audiences, and commercial viability, often bolstered by government funding. In contrast, AHSS publications face greater volatility, given their focus on qualitative critical analysis, appeal to more diverse audiences, and limited commercial viability, along with occasionally complex issues surrounding copyright and licensing.

This makes AHSS publications somewhat more prone to long-term preservation risks. Additionally, the authors have paid particular attention to scholarly monographs, which, for the purpose of this report, are identified as in-depth, single-volume books that often serve as foundational texts within academic fields. While all scholarly monographs fall under the broader category of academic books, the latter term includes a wider range of publications, such as textbooks, edited volumes, and some more popular works. Scholarly monographs are distinguished by their focus on original research and are aimed at a specialised audience, whereas academic books can serve both specialised and general readerships.

We particularly consider the changing role of libraries in this process and how this challenge is being considered and addressed differently across various types of libraries. We also look at the provision of access within the various preservation systems and the implications this may have for the effectiveness and subsequent value of libraries to future scholars and readers. We consider

how such arrangements might change the nature of libraries and their role as custodians, safeguarding long-term preservation and ensuring access to knowledge, and what new arrangements might be evolving across the sector as a whole.

Our enquiry into the preservation of published knowledge and ongoing access to it, led us to the separate but related question of whether there is or will be a meaningful public domain into which electronic literature will pass once copyright and all other legal restrictions have expired. However, we believe that the increasingly critical issues surrounding the definition of, access to, and the control of, digital public space requires a separate, substantial investigation, and therefore with regard to this study we attempt to answer just two key questions; i) Is there likely to be a meaningful public domain for electronic literature? ii) Are today's efforts by themselves going to be enough to ensure it?

The report is based on in-depth interviews with key stakeholders as well as quantitative research in the form of two surveys, one of librarians' perceptions of these issues and another of actual licensing arrangements between publishers and libraries. These arrangements have evolved within a context which for libraries is characterised by the constraints of limited budgets, increasing demands and rising costs.

We find that the various arrangements for long-term preservation are all partial solutions and that there is near unanimous concern that some digitally published literature will be lost. Every interviewee thinks this is a problem – even after having completed interviews with representatives of organisations that were specifically created to be long-term solutions. We find no comprehensive and universal approach to the preservation of electronic literature, nor the provision of ongoing access to it, from the moment of its creation right up the point it enters the public domain and beyond. All the solutions deal with only a part of the problem and/or are contingent on arrangements between publishers and libraries that may or may not endure.

It has taken centuries to evolve systems that ensure printed works are preserved and can be accessed long after they have been published. The recent and relatively sudden shift to digital formats is threatening to upend those systems. Whereas there were typically hundreds of copies of a printed work, distributed across many organisations and archives around the globe, there may in future be as few as two master-copies of a digital publication, one under the publisher's control, that libraries and individuals may only access for the duration of an acquired licence, and another held in a national legal deposit archive, where access is intentionally very restricted. So, whereas when a printed work enters the public domain there might often be many surviving copies still somewhere in existence, that can then be shared and made freely available, this is not yet certain to be the case for electronic literature.

This report considers the very long-term persistence of these two established points of access to a digital publication, as well as assessing the viability of preservation initiatives that aim to act as safeguards and provide failover access to e-literature, introduced elsewhere in the electronic publishing ecosystem, most notably in third party dark archives. We also look specifically at Open Access publications and assess the variations in arrangements put in place to safeguard these works for the long future. We commend the efforts of the many people behind all these initiatives while also observing vulnerabilities in them.

We conclude that knowledge will almost certainly be lost unless new robust and intentional arrangements are put in place to preserve and provide future access to electronic literature. The challenges to be addressed are numerous, interrelated, and often intricate, and we recommend further in-depth research into specific areas that we have identified as either pressing or that have the potential to deliver significant impact. As such, this report serves as a high-level, preliminary overview of a multifaceted concern, highlighting what appears very likely to result in a critical

outcome. We offer sixteen key recommendations. Nine of these we present as general recommendations, and seven that are largely specific to scholarly monographs. We find that, until now, most of the considerations around long-term preservation have been primarily focussed on e-journals, and an ill-prepared shift from today's default preference for printed editions of books to digital-only formats may ultimately lead scholarly monographs to become more threatened and subsequently at greater risk than many have assumed.

Recommendations – General

To address the gaps in the long-term preservation of digital literature, we propose the following actions to safeguard all types of electronic publication:

1. The wholesale adoption of common international standards for preservation of electronic publications.

2. Exploring the development of more affordable and user-friendly archival software systems, such as LOCKSS, to enable libraries to effectively archive digital literature and create collaborative redundancy. Such systems would allow libraries to preserve digital works independently and within the scope of their legal permissions, ensuring long-term access and safeguarding against potential loss.

3. Further discussion, research, and collaboration with the dark archives to discover new opportunities to grow sector-wide support for, and fund additional resources to build upon, their excellent preservation work.

4. Research into developing distributed archiving solutions specifically tailored to the needs of Open Access publications. Building on foundations laid by initiatives like JASPER, this should also aim to foster greater awareness that the act of publication alone is not enough without a plan to ensure that published knowledge is actively being preserved in order to made permanently accessible as an integral part of the publishing process.

5. The development of copyright risk assessment to encourage libraries to provide access to works that are out-of-commerce but not yet in the public domain.

6. That all preservation arrangements lessen contingency so that works can survive shifts in the balance of power and/or withstand the passage of time until the emergence of a sustainable system that is fully resourced and enforced by both publishers and libraries.

7. Raising awareness about this important issue so that governments, as well as librarians, are concerned about the loss of knowledge and will be more prepared to support plans to preserve it.

8. Advocating that libraries be granted the legal right to download and print preservation copies of digital literature for which they have already paid and secured perpetual access, potentially through a collective legal challenge to current contractual limitations.

9. Our key recommendation concerns national libraries, which we find are the only organisations systematically, via legal deposit, currently preserving electronic literature. However, this preservation is often undertaken in tandem with copyright registration, particularly in the United States, rather than as intentional cultural

preservation for broader public access. We recommend advocacy for a change in that approach as the first step. The next step would be for national libraries to systematically provide public access to this material as it falls out of commerce and eventually out of copyright, so that national libraries are actively preserving knowledge, providing access and, in so doing, creating a meaningful public domain for e-literature.

Recommendations – Specific to Scholarly Monographs

Recognising the particular challenges posed to scholarly monographs in digital formats, we recommend the following targeted actions:

10. Conducting comprehensive research to determine the current share of scholarly monographs published exclusively in digital or digitally enhanced formats, and to track the growth rate of this trend.

11. A very directed project to identify if any such born digital scholarly monographs have already been lost and if so, how many to date.

12. A comprehensive re-evaluation of preservation strategies for digital scholarly monographs, including the feasibility, implications, and costs of developing capabilities to systematically acquire, print, and secure these works.

13. Research into how to develop and set aside a capacity within the successful dark archives e-journal practices, so that scholarly monographs start to be systematically preserved within these systems, in anticipation of a move to digital-only formats.

14. Further research into the actual number of orphaned, born digital e-books and consideration of how to create an archive facility alongside an accessible library, into which they can pass while simultaneously seeking the permission of the rights holder.

15. Similar research into other academic and trade e-books that are currently published exclusively in digital formats, including self-published works. The research should also assess whether there are established plans to identify and preserve these works as cultural artefacts. (Our understanding is that much of this e-literature is regarded only as lower-quality reading material and not as potential cultural artefacts, so therefore unworthy of preservation.)

16. The investigation into the potential need for, and the feasibility of the creation of, a centralised data source on the preservation status of e-books – the equivalent of the Keeper's Registry for e-journals.

Electronic literature has crashed into the print-based knowledge economy, with its highly developed commercial relationships, practices, norms, laws, and safeguards. We note that many digital publishers may have been quick to exploit the potential of this new technology, most often for profit rather than the public good. Unfortunately, we find that robust systems for preserving this knowledge and making certain that future scholars, generations hence, will have a guarantee of access to it, have not yet been much considered and certainly not established. As a result, knowledge – in particular knowledge discovered or created since the turn of the millennium – is in serious danger of being lost.

1. Introduction

The ever-evolving digital landscape presents both opportunities and challenges for the creation, dissemination and preservation of culture and knowledge. This report, commissioned by the Arcadia Fund, investigates a specific concern: the potential erosion of public access to written knowledge, as libraries transition from physical book collections to licensed electronic materials, in particular scholarly e-journals and monographs. Arcadia's commitment to Open Access and cultural heritage preservation aligns perfectly with this inquiry. In particular, their ongoing commitment to, and support for, access to out-of-print and out-of-copyright scholarly materials, both physical and digitised, underscores the importance of ensuring unfettered availability of knowledge for the public good.

The following Introduction sets out how this report will delve into the critical issues of long-term preservation and future access, and then explores the question of whether there is a need to reimagine the concept of the public domain within the context of e-literature.

1.1 Scope

This report examines a pivotal moment in the evolution of knowledge creation and dissemination, driven by the rise of born-digital literature. This category encompasses e-journals and e-books that exist solely in digital form, lacking any physical equivalents. Due to the specificity of the brief, our primary focus falls onto scholarly and academic e-journals and scholarly monographs, as they constitute the bedrock of future research.

Although our central emphasis remains on the critical realm of academic discourse and study through the preservation of scholarly materials, inquiries with trade publishers and public libraries have yielded additional insights that have informed our understanding of the broader e-book landscape, including trade and general-interest academic publications. These categories are often considered a lower priority due in part to the fact that many continue to be produced at sufficient volume in print, with little or no evidence that the entire market will transition to exclusively digital formats any time soon. Moreover, these publications are frequently viewed as ephemeral or of lesser importance by the academic community.

However, certain genres within these categories are increasingly being produced exclusively in digital formats, which puts them at particular risk of being lost. The growing appetite for these digital-only publications, combined with limited preservation efforts from libraries and archives, raises significant concerns. Although these works may lack immediate scholarly value, they capture cultural, linguistic, and creative elements that could be of great interest to future generations. The vulnerability of these digital-only publications, exacerbated by a combination of resource constraints and institutional disinterest, places them among the most at-risk categories. This evolving trend calls for greater attention, even though it currently lies outside the scope of mainstream preservation efforts.

While we have intentionally avoided an in-depth focus on this specific area in the main report, we have instead reserved our findings for a brief set of remarks in Appendix D, which outlines our key observations regarding the preservation of these publications.

1.2 Context

The long-term preservation and accessibility of e-literature is intricately woven into the economic models of various stakeholders within the publishing ecosystem. Scholarly journals, academic e-books, and commercially-driven e-books (e.g. genre fiction) each function within distinct financial structures, presenting unique challenges and implications for their long-term viability. Unravelling these disparities is crucial for crafting sustainable solutions that would ensure the health of the entire corpus of scholarship and recorded knowledge. Directing limited resources at isolated problems might offer temporary succour, but it almost intentionally neglects the bigger picture, hidden as it is by the illusion of progress. The aim instead should be to develop comprehensive solutions that guarantee the long-term viability of the entire written record, not just address isolated issues and 'pet projects', leaving the broader landscape just as fragmented and unstable.

1.3 Charting the course: A roadmap for inquiry

This report intentionally steers clear of two specific areas: i) addressing arguments concerning value for money in arrangements between libraries and publishers and ii) the question of fairness within the current publishing landscape. Our focus instead centres on a broader and more existential concern – the long-term survival of e-literature. While we may cite specific instances for illustrative purposes, our investigation does not delve into whether particular works have been lost or are currently inaccessible. Our concern is on why this is happening, or may happen, and what might be done to reduce or prevent this from happening in the future. We have taken an analytical lens to an examination of current systems and licensing practices. In doing so, we aim to project the potential for a positive future for the scholarly and academic e-literature that is being published today, with a particular emphasis on the future state of these works once they fall outside of their licensing agreements, or become commercially unavailable (out-of-commerce), lose the attention or interests of rights owners and, ultimately, as they enter the public domain.

The core questions guiding this inquiry are as follows:

- **Preservation challenges:**

    To what extent do current practices ensure the perpetual preservation of e-literature?

- **Future access landscape:**

    How and where will users access any such preserved e-literature in the near and distant future?

    To what degree will libraries and/or other entities facilitate such access?

- **Evolving role of libraries:**

    What roles are libraries currently playing in safeguarding and providing access to e-literature?

    How might these roles change in the future?

- **Public domain in jeopardy?**

    Beyond legal definitions, are there genuine efforts to ensure the existence of a meaningful public domain for e-literature?

By addressing these critical questions, this report seeks to illuminate potential challenges and pave the way for the consideration of sustainable solutions that will guarantee the enduring accessibility of e-literature for generations to come.

1.4 Methodology

To uncover the complex landscape surrounding the long-term future of e-literature, the team of four researchers compiling this report employed a multifaceted research methodology. Here, we detail the various tools utilised to gather the information and insights that inform our analysis.

1.4.1 Stakeholder Interviews

We conducted in-depth interviews, both in-person and remote, with a diverse range of stakeholders across the e-literature ecosystem. These included librarians from research, public and national libraries across the US and Canada, the UK and EU, Australia and NZ, alongside publishers and other key players within the sector. The interviews were designed as open-ended dialogues, fostering a rich exchange of viewpoints on the economic realities of e-literature and, more specifically, perceptions regarding the prevailing attitudes and behaviours around preservation and access. A list of interviewees can be found in Appendix A.

1.4.2 Survey of Licences

Complementing the qualitative data gleaned from stakeholder interviews, we embarked on a quantitative research initiative. This involved analysing thousands of contracts governing library-publisher agreements for e-literature access. Our objective was to identify recurring patterns and key variations within these contracts, paying close attention to factors such as jurisdiction, publisher, and publication type. Appendix B includes examples of typical licensing agreements encountered during this analysis.

Throughout the contract review process, we sought answers to critical questions concerning the nature of access provided:

- **Access duration:**

    Do contracts guarantee perpetual access?

    Do contracts guarantee continued access after license termination?

    If so, under what conditions?

- **Archiving rights:**

    Are libraries permitted to retain archival copies of licensed works?

    If not, do contracts outline robust alternative preservation arrangements?

- **Public domain access:**

    How do contracts address access scenarios for works entering the public domain?

1.4.3 Librarian survey

To further complement our qualitative data, we designed and distributed a global questionnaire to which responses were received from 97 librarians. This survey aimed to gauge levels of concern among practitioners regarding the issues of e-literature preservation and to assess their understanding of current practices for ensuring long-term access. The Findings in Section 2 of the report offers a detailed analysis of the librarian survey results, while Appendix C provides a statistical analysis of the responses received.

1.4.4. Synthesis and analysis

The culmination of our investigation involved a rigorous analysis of the data gathered through contract reviews, librarian interviews, and global surveys. The research team, comprised of individuals with substantial experience in knowledge economy leadership and advisory roles, actively participated in synthesizing the findings. Brief biographical sketches of the team can be found in Appendix G.

1.5 Scholarly and academic e-journals

The traditional model of scholarly and academic journal publication rests on a foundation of intellectual exchange rather than financial gain. Authors, driven by career advancement, contribute their research for publication, usually retaining copyright but with minimal expectation of monetary rewards. Publishers, often for-profit entities, step in to manage the dissemination process, acquiring publication rights in exchange for editorial services and distribution channels.

Readers within the academic realm rely heavily on access to these journals to remain abreast of current research developments. However, the financial burden of individual subscriptions falls not on the readers themselves, but on the libraries of their institutions. These libraries, recognising the indispensable role of journals, dutifully subscribe to the comprehensive packages offered by publishers.

This lucrative business model, with its guaranteed market of academic institutions, fuelled the rise of giants like Pergamon Press in the 1960s. Elsevier's subsequent acquisition and conglomeration strategy led to the now-ubiquitous 'Big Deals', where libraries are pressured into large-scale subscriptions encompassing vast swaths of academic titles, often at standardised rates. This approach has transformed the scholarly publishing landscape into a nearly $19 billion-a-year, almost entirely digital, enterprise. The vast majority of these journals exist today solely as e-journals, a stark shift from the traditional print collections.

Previously, libraries subscribed to physical journals, meticulously building and preserving their own permanently accessible collections. The digital shift has fundamentally altered this dynamic. Libraries now pay for licenses to access e-journals hosted on publisher servers. The libraries have relinquished ownership in return for a more precarious access model. Biologist Michael Eisen aptly described this economic arrangement as *"a perverse and needless obstacle to the optimal use of scientific knowledge"*.

Open Access (OA) presents an alternative model where research institutions, usually via library budgets, subsidise the publication process, making the resulting content freely available online. However, while offering the research communities distinct financial advantages, OA also introduces its own set of challenges regarding preservation and ongoing access, which will be explored in a later section.

The relationship between libraries and publishers transcends mere commercial transactions; it underpins the very foundation of academic discourse and knowledge dissemination. However, this vital partnership is undergoing a period of significant transformation. As libraries relinquish ownership in favour of licenses, the long-term implications of this changing arrangement on collection development, secure storage, diligent preservation of, and the guarantee of future access to, this critical body of knowledge, demands careful consideration.

1.6 Scholarly monographs and other academic e-books

In the context of this research, our focus regarding 'other academic e-books' extends beyond digital academic textbooks to include born-digital works such as edited volumes, collections, anthologies, reference works, technical reports, government papers, case studies, dissertations, and theses.

Much like e-journals, the motivation behind the production of scholarly monographs and other academic e-books is driven by more than just the potential for sales revenue. Authors are primarily focused on disseminating their research and enhancing their academic standing, often with the aim of securing future career opportunities or promotions. Similarly, the intended audience – researchers and academics – turns to these e-books as a means of advancing their own work. Consequently, the business models for these publications hinge on the crucial partnership between publishers and libraries, with libraries functioning as the primary conduit of access.

All the librarians interviewed for this report indicated that the digital portion of their book collections continues to grow, as does the share of their budgets dedicated to e-books. Only one respondent from a research or academic library indicated that they do not currently subscribe to e-books. However, our interviews with both librarians and publishers reveal that, unlike the nearly universal shift to digital seen with e-journals, academic books – particularly scholarly monographs – still retain a substantial presence in print alongside their digital counterparts.

Many librarians stated that, under normal circumstances, they purchase a print copy in addition to licensing the digital edition, or in some cases they might opt for a print version entirely. Budgetary constraints in recent years, however, have begun to limit these options, raising growing concerns about the long-term integrity of library collections.

While some interviewees remained optimistic that scholarly monographs and other academic e-books will continue to be available in both print and digital formats, others were less confident. This divergence in perspective was not tied to a specific group (e.g., publisher, librarian, or archivist), but there was consensus that the digital shift that has transformed journal publishing could, in time, similarly impact books.

A rising concern expressed during our research was the possibility that future editions of scholarly monographs might evolve into value-added, digitally-enhanced versions incorporating features like embedded media or live datasets. Although we were alerted to this possibility, we found no clear examples or evidence that this trend is currently accelerating.

Despite the increasing prevalence of e-books and the diversification of formats, print editions of these works remain sufficiently available. This indicates that many printed copies are likely to endure well beyond the period of commercial viability (when they go out of print), and will continue to be preserved in libraries until they enter the public domain.

E-books also inherently differ from e-journals in their suitability for subscription models. Unlike the ongoing, serial nature of journals, e-books require a wider range of licensing options to accommodate individual purchases or lend themselves to entirely different licencing models.

Consistent with other areas of our inquiry, this diversity in purchasing and licensing arrangements for e-books by itself introduces complexity and potential vulnerabilities for long-term preservation and access. We will explore these vulnerabilities in detail later in the report.

1.7 Trade e-books

Trade e-books occupy an entirely distinct economic space compared to their scholarly and academic counterparts. Publishers in this realm derive the bulk of their revenue from direct book sales, with libraries constituting a relatively minor, and often unwelcomed, market segment.

While academic libraries may not be central players within the trade e-book market, valuable insights can be gleaned from this sector. These insights hold significant implications for the broader conversation surrounding the preservation and accessibility of all e-literature, regardless of category.

Although trade books are often viewed as having a limited lifespan as commercial products, they should also be recognised as potentially significant cultural artifacts. Over time, these works may reveal valuable and often overlooked insights into the era in which they were produced – the people, societal norms, challenges, and nuances of the historical period. Therefore, their preservation warrants attention, not just as literary works but as essential records of cultural history.

One particularly noteworthy aspect of the trade e-book landscape is the presence of Amazon, a company that dominates the digital realm and currently chooses not to offer its publications to libraries. This presents a different and altogether unique challenge while underscoring the need for a more holistic approach to the entire ecosystem in order to ensure long-term survival of all categories of e-literature and future access for everyone who needs it.

1.8 A meaningful public domain

A key element of this report revolves around the question of whether there is now, or will in the future be, a meaningful – or enforceable – public domain for e-literature. Currently a great many e-publications exist outside of this realm, as they are, and will remain for decades to come, under copyright protection. The likelihood is that the vast majority of these works will not enter the public domain within the authors' lifetimes.

Further to this, the inquiry also briefly took time to extend beyond the question of a public domain for e-literature and contemplated the very nature of the concept of a much broader 'digital' public domain, one that would take advantage of the many new and emerging affordances of digital technology. Is a digital public domain inherently different from its analogue-era counterpart? If so, how?

This report considers the challenge of reimagining these ideas within a digital context, considering potential evolutionary trajectories and proposing guiding principles for the construction of a robust public domain tailored for the unique characteristics of the world that we will be living in, rather than restricted by the one that we came from.

1.9 Structure of the report

Following this Introduction, in the Findings section we present our research and analysis. We then present our Commendations, highlighting selected examples of good practice in preservation and/or provision of access. We then offer Recommendations, in which we set out some ambitious

proposed solutions to the vulnerabilities observed. Finally, in our Conclusion we summarise our analysis of the long-term vulnerabilities of e-literature under present arrangements.

1.10 The headlines

Key discoveries, analyses and recommendations have been introduced in the executive summary, and will be explored in more detail. They bear reemphasis and some repetition:

- **Universal concern:** A resounding consensus emerged from our interviews and surveys – the prevailing practices for preserving and ensuring ongoing access to electronic publications require urgent scrutiny. The pervasive lack of confidence in the current system is deeply concerning.

- **Knowledge in jeopardy:** Because much e-literature remains locked behind pay-walls, there is no exhaustive survey capable of identifying already lost or inaccessible works. However, many librarians interviewed for this report indicated that they have been unable to retrieve e-literature they are certain was previously available. This presents a serious risk that significant knowledge is being lost, with the potential for even more to vanish. If this trend continues unchecked, future generations may find themselves deprived of important insights, collective wisdom, and cultural practices from previous eras. The risk is not inherent to digital technology itself, but rather to the inadequacy of the current systems for preserving and providing enduring access to e-literature.

- **An ongoing conversation:** This report is not intended as the definitive statement on this critical issue. On the contrary, we view it as the catalyst for a continuous and essential dialogue.

Some may argue that these concerns are overly focused on a distant future. We vehemently disagree. The example of London's Thames Barrier serves as a potent analogy. This colossal structure protects a vast area of central London from storm surge flooding. When first proposed in the wake of a devastating flood in 1953, the barrier was not yet a pressing necessity. However, visionary planners recognised its future indispensability, anticipating its potential use two to three times per year in the coming decades.

Their proactive approach stemmed from a profound commitment to safeguarding the well-being of future generations, underscoring the importance of foresight and taking action before a crisis erupts. As it turned out, the Thames Barrier's value was demonstrably real, with a staggering 50 deployments in 2013 alone. Had construction been delayed until such frequent activation became commonplace, it would have been far too late.

The potential consequences of inaction are difficult to envision in the absence of immediate disaster. However, waiting for irrefutable proof of a problem guarantees a missed opportunity for proactive solutions. We urge the adoption of a Thames Barrier-inspired approach, taking decisive action in the present to preserve our knowledge heritage for the benefit of those who come after us. The time to act is now.

2. Findings

A wicked problem is a term used to describe a complex, multifaceted issue that is difficult (or impossible) to solve because of incomplete, contradictory, and changing requirements. The concept was introduced by design theorists Horst Rittel and Melvin Webber in their 1973 treatise "Dilemmas in a General Theory of Planning".

Key characteristics of wicked problems include:

- **Complexity and interconnectedness:** Wicked problems have many interdependencies, making them highly complex.

- **No clear solution:** There is no single, definitive solution to a wicked problem. Solutions are not true-or-false but better-or-worse.

- **Unique nature:** Each wicked problem is essentially unique, meaning that there is no set procedure for solving it.

- **Uncertainty and ambiguity:** The problem and its causes are often not fully understood, and the information needed to understand it may be missing or conflicting.

- **Stakeholder conflicts:** Multiple stakeholders with different values and priorities are involved, making consensus difficult.

- **No stopping rule:** There's no definitive end point or stopping rule for solving a wicked problem. Solutions are iterative and ongoing.

- **Consequential impact:** Solutions to wicked problems often have significant and far-reaching impacts.

Every wicked problem can be considered to be a symptom of another problem. Examples of wicked problems include climate change, poverty, and healthcare. Addressing these problems typically requires innovative thinking, collaboration across disciplines, and adaptive management strategies.

By such definitions, the challenge of ensuring perpetual access to licensed copies of electronic journals and books equivalent to, or better than, the opportunities that they have to do so for physical copies is, we believe, a wicked problem.

The authors believe it meets the criteria in the following ways:

- **Complexity and interconnectedness:** It involves factors like copyright law, economic models of publishers and libraries, technological change, and user needs – all interconnected and influencing each other.

- **No clear solution:** There's no one-size-fits-all answer. It requires a multi-pronged approach that satisfies various stakeholders.

- **Unique nature:** A holistic solution will require an unprecedented alignment of challenges and solutions that differ depending on the type of literature, publisher, format, licensing terms, library, geography, jurisdiction and, ultimately, the need of users.

- **Uncertainty and ambiguity:** The full impact of digital preservation and licensing on long-term access is still evolving.

- **Stakeholder conflicts:** Libraries, publishers, universities, and researchers and everyday readers all have different priorities and perspectives.

- **No stopping rule:** This is an ongoing challenge that requires constant adaptation as technology and user needs change.

- **Consequential impact:** Loss of access to electronic literature will hinder research and education.

In this Findings section of the report, we delve into the complexities that contribute to this specific 'wicked problem'. We explore the root causes of these disagreements amongst stakeholders who, undoubtedly, believe they are actively working towards solutions. However, a closer examination may reveal that current approaches may, in some respects, be unintentionally perpetuating the very problem they seek to solve.

2.1 Libraries vs. publishers

A symbiotic relationship between libraries and publishers underpins all forms of e-literature, particularly scholarly and academic works. However, this vital partnership appears to be under significant strain. They appear to be stuck in a dialogue of the deaf. A former publishing executive, who is in a position to know, candidly acknowledged in our interview the prevailing sentiment: "Libraries hate publishers and publishers hate libraries."

Our interviews with librarians revealed a near-unanimous dissatisfaction with the recent, rapid shift towards the exclusive licensing of e-publications. Cost concerns emerged as the central theme. Librarians perceive publishers' innovations in electronic publishing as primarily driven by profit maximization rather than the public good. They question why a mere access license for an e-publication can be more expensive than buying and permanently keeping a print copy. In their view, this translates to being charged a higher fee for reduced flexibility, "Paying more for less" as one interviewee phrased it, adding, "Publishers are charging us more for essentially the same product, but doing less". The traditional ownership model of print media stands in stark contrast to the ongoing expenditure necessary with e-publication, in which even an e-book can be licensed under a form of subscription leading librarians to feel as though they are "Not just paying more, but paying more over and over again... I don't have to pay for the same dress every time I wear it."

Another librarian, perhaps inadvertently capturing the academic community's metaphorical dependence on a steady stream of scholarly works, likened publishers' profit margins to "Drug dealer margins", highlighting the perceived exorbitant costs associated with access to e-literature.

This preoccupation with value for money has arguably become such a dominant concern in navigating the challenges of digital publishing that librarians may have limited bandwidth remaining to consider other critical aspects of the services they have traditionally taken for granted. As Alicia Wise of CLOCKSS pointedly observed when queried about the lack of collaboration between libraries and publishers to establish sustainable, long-term preservation solutions and guarantee perpetual access, "There's not a lot of leadership in this area."

The current dysfunctional dynamic between academia, libraries, and publishers suggests that these pressing issues are unlikely to be effectively addressed, adequately resourced, or definitively resolved through existing collaborative efforts.

2.2 Licences

The terms governing library access to e-literature are far from uniform. Significant inconsistencies exist depending on the type of publication and library in question. To illuminate this multifaceted landscape, we present our findings in three distinct sections. It is important to note that the analysis presented here, drawing upon a large-scale survey of thousands of licenses and librarian questionnaire responses, offers a broad overview rather than an exhaustive account.

2.3 Licences under which research libraries access scholarly and academic e-journals

E-journal subscription licences usually cover a fixed time period, typically between 1-5 years, with an automatic annual renewal at a pre-set rate, built into the contract (an example of typical terms of subscription in an e-journal licence is presented in Appendix B). Most offer a package of multiple e-journals (in some cases hundreds of titles), and there is also often a clause to allow for the addition of new journals and/or for subtracting any that the publisher either loses control or interest in, or which cease publishing.

Almost invariably the contracts for e-journals include an assurance of some kind of perpetual access for any edition of an e-journal issued during the subscription period, (again, see Appendix B for examples of this clause.). This is also sometimes described as 'Post-Cancellation Access'
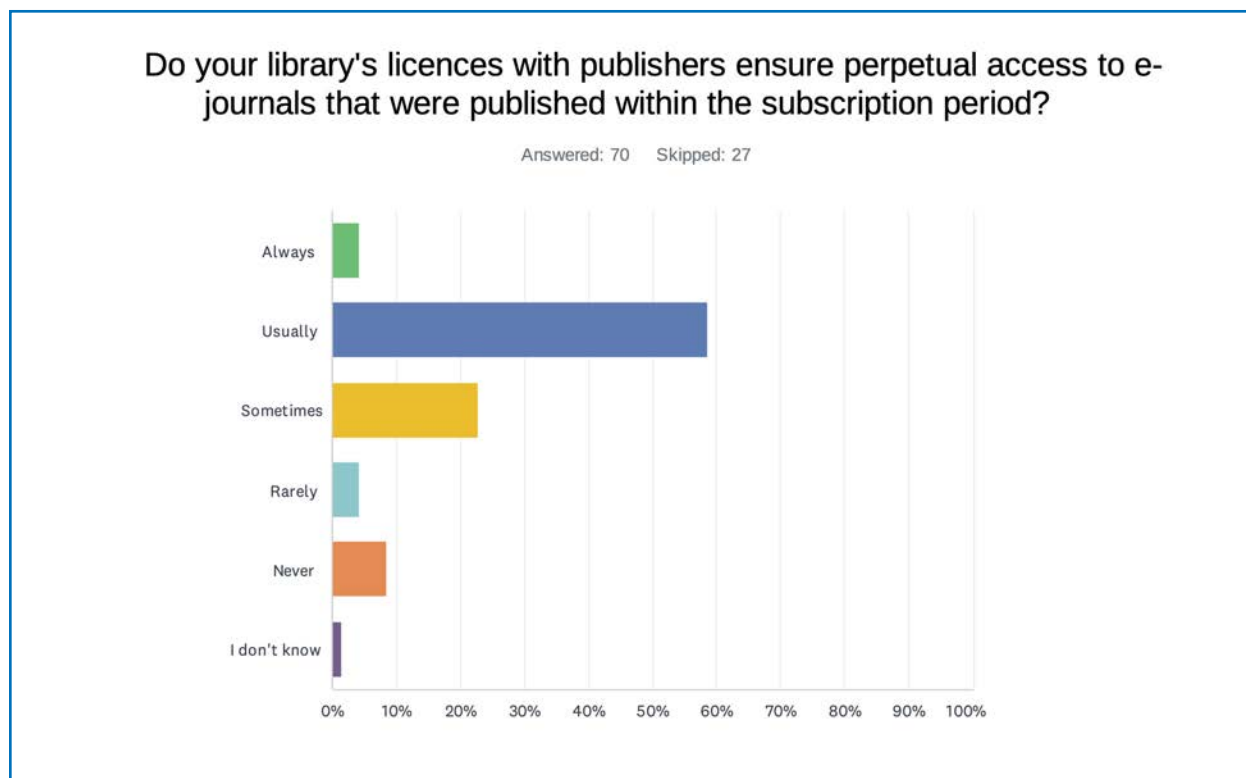

Table 1

(PCA). So, if a library stops subscribing to the e-journal, although it will no longer receive future editions, it will continue to have the right to access all the editions that were published within the subscription period in perpetuity. Our questionnaire mostly confirms this, as is revealed in Table 1. Almost 90% of librarians who responded to the questionnaire report that their contracts for e-journals either always, sometimes or usually promise perpetual access to such qualifying material.

It is apparent that licences for e-journals clearly and consciously attempt to replicate the experience of print journal subscriptions, but with a key difference. Although some older licences may have allowed libraries to download and retain an archival copy of an edition of an e-journal that it had subscribed to, this option is no longer permitted in most of the more recent licences we examined.

There are two further points of note in what might be referred to as the licence's 'small print'. Perpetual access can come with a caveat that it will be 'substantially equivalent to the means by which access is provided under this Agreement'. This could be interpreted as allowing a publisher the right to change the terms of access in the future.
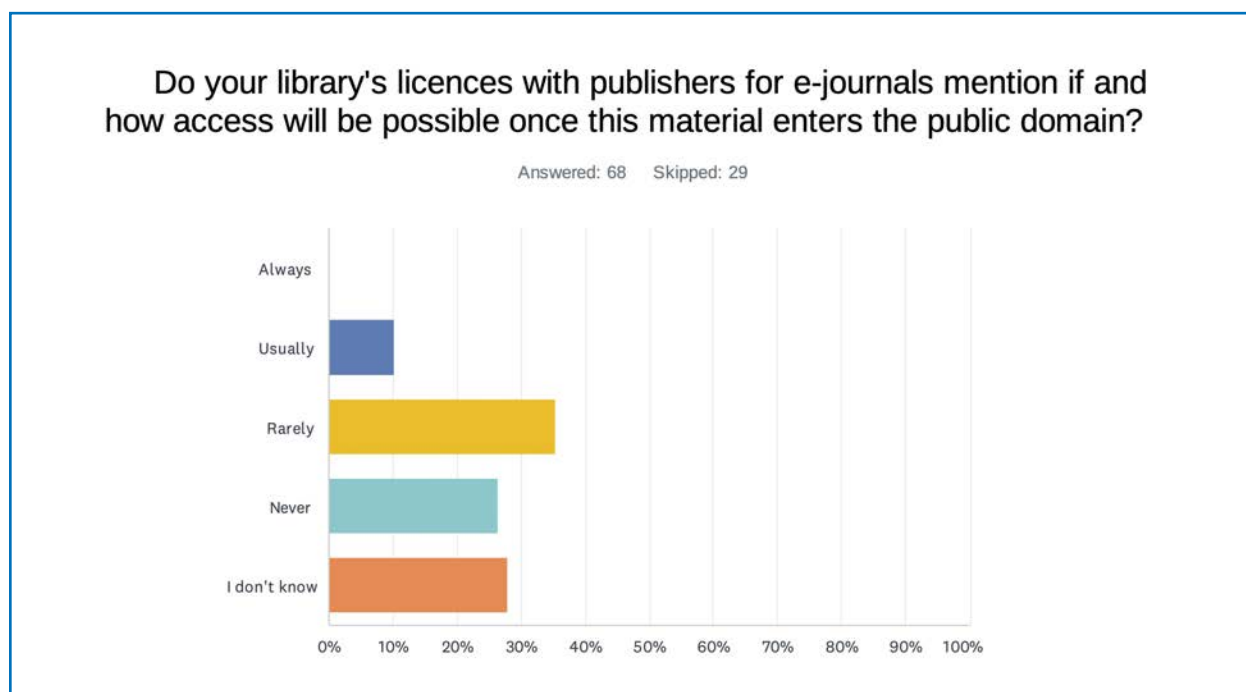

Do your library's licences with publishers for e-journals mention if and how access will be possible once this material enters the public domain?

Answered: 68   Skipped: 29

Table 2

A second point of note is that the licences do not provide for any possibility of a change in the terms when the copyright expires. The licences typically offer access 'in perpetuity' … 'in accordance with the provisions of the Agreement'. As can be seen in the excerpts from sample licence in Appendix C, those provisions are the usual legally binding restrictions for a copyrighted work. So, the contract appears to commit the library to perpetual copyright restrictions: there is no suggestion that these terms will ever be rescinded, even when a work enters the public domain.

Following this discovery, a question was included in the questionnaire to gauge current levels of understanding of this perpetual restriction. Table 2 shows that 65% of librarians report that this issue is rarely or never mentioned, and almost a quarter of all librarians surveyed do not know whether their licences for e-journals mention arrangements for access once the material enters the public domain.

For centuries libraries have owned and controlled their own collections and archives of printed works. With the introduction of e-journals, that is no longer the case and there are consequently numerous vulnerabilities in this altogether different arrangement. This report will go on to consider these vulnerabilities and assess the various solutions that have emerged so far in an attempt to protect libraries and, more importantly, researchers, scholars and ultimately the public, against these vulnerabilities.

2.4 Licences under which libraries access scholarly and academic e-books

Other than e-journals, we particularly wanted to find out to what extent libraries have perpetual access to scholarly monographs that they licence in digital form. Presuming that the e-books licensed by research libraries are almost exclusively scholarly monographs we separately analysed this subset of the respondents.

Table 3 indicates that more than 90% of all research librarians believe that they usually have perpetual access or Post Cancellation Access (PCA) included in the contracts for e-books that they licence. However, our research reveals that as many as 30% of contracts either do not offer perpetual access to scholarly monographs or significantly constrain access if the subscription is terminated. In one example the library is even contractually obliged to destroy any downloaded copies it may hold if it ever cancels the subscription. (See excerpts from sample licences in Appendix B.)
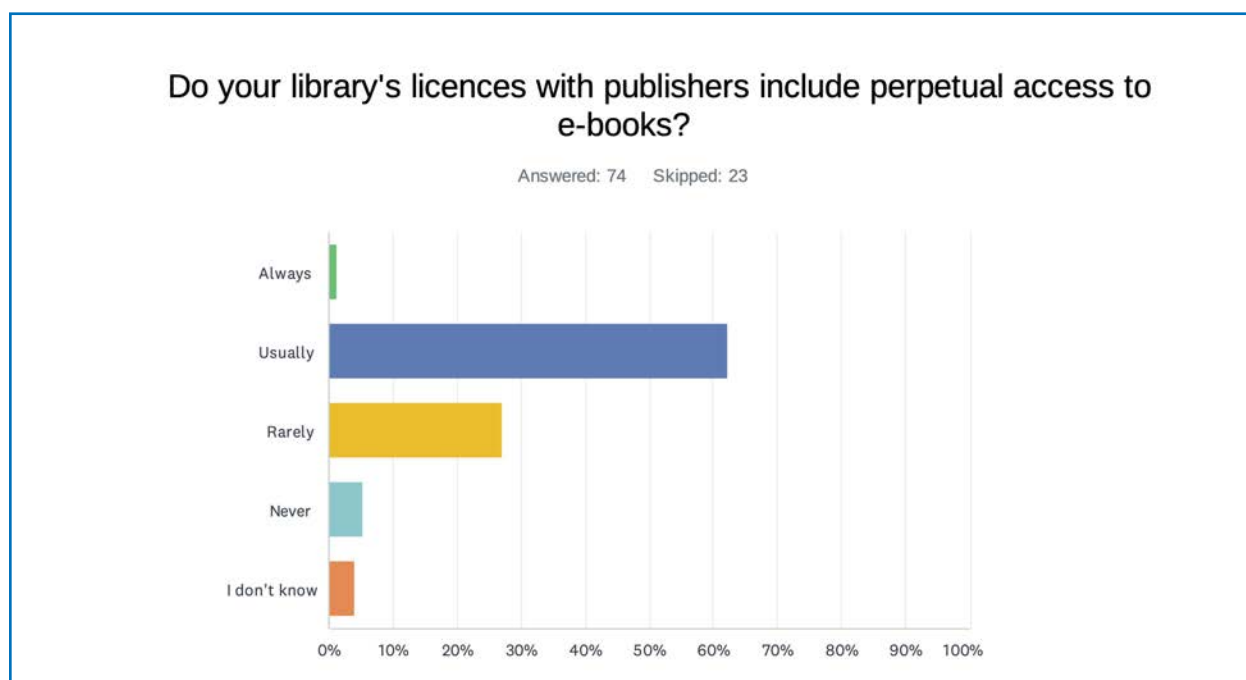

Table 3

Even when licences do offer PCA the contracts often pass some or all of the responsibility for archival preservation onto the libraries themselves. This means that many research librarians are mistaken in their belief that they have the contractual certainty of continuing access to the e-books that they have licenced, either because some contracts may not actually include it or because the preservation archiving necessary to make it a reality may never have taken place.

A further question on whether their licences for e-books mention the public domain (see Table 4) reveals that over 90% of respondents either think this is rarely or never mentioned or, again, simply don't know.

In our surveys of licences, we did not find a single instance in which the public domain is mentioned as a threshold, nor the inclusion of a 'sunset clause', after which time the licensing arrangements might change.
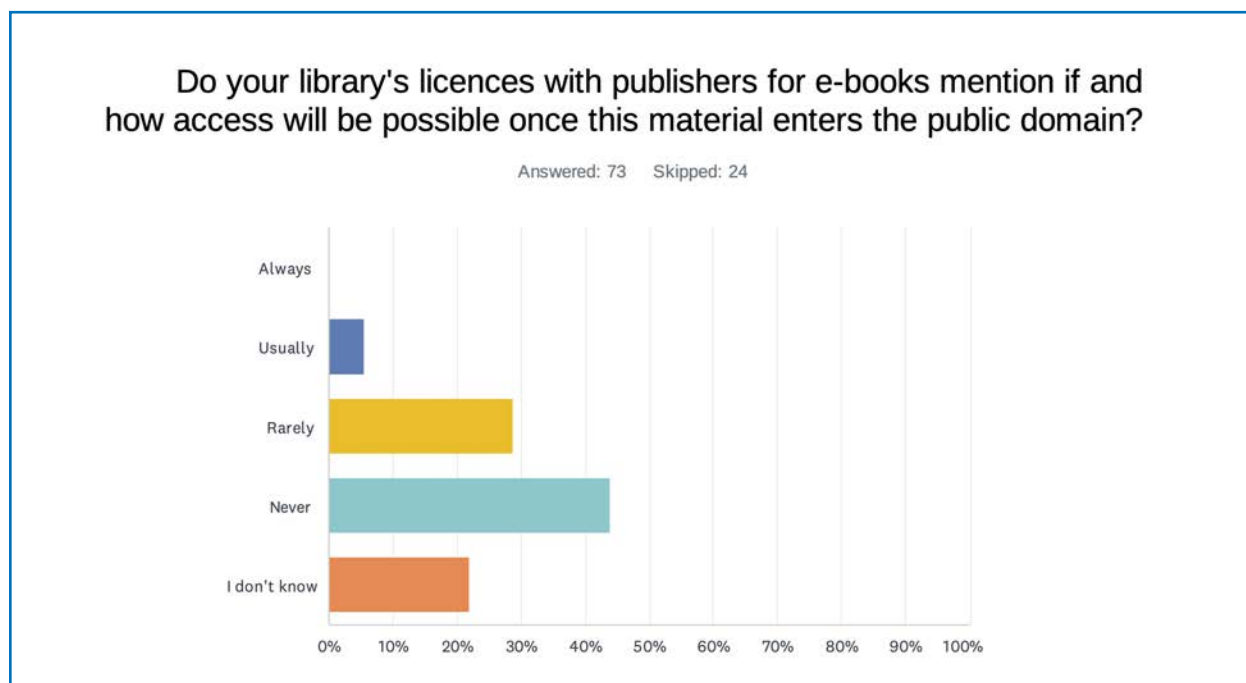

Table 4

2.5 Licences under which public (circulating) libraries access trade e-books

All but one of the circulating libraries which responded to our questionnaire said that they did licence e-books. This is supported by anecdotal evidence from our interviews which suggests that a growing share of libraries' collections and budget is dedicated to e-books. However, because librarians understand that it is a somewhat different customer base who accesses their digital collections from their print collections they still, at present, collect titles in print as well as digital editions if the budget allows it.

We also note that because the circulating libraries only account for a small share of trade books print and digital distribution, even if these libraries become exclusively digital there are still going to be many privately purchased print copies that may survive the passage of time.

Circulating libraries' licences for trade e-books are almost entirely metered by time and/or use: the library typically has access to the book for a set period of time (usually 12 to 24 months) or a fixed number of lends (often 26), whichever expires first. Public librarians are aware that this prevents them preserving this material but have not as of now seen this as their responsibility. They are also accustomed to a high level of churn, as print titles are constantly de-accessed because the limitations on available shelf space forces libraries to trim the size of their physical collections in order to be able update or refresh them with new titles.

We understand that the shift to digital further increases libraries' natural propensity for such churn. First because of the very short licence period under which trade books are made available, and because the cost of a licence is often three times the cost of a print book, libraries are likely to make licence non-renewal decisions sooner than they might otherwise.

Second, public (circulating) libraries do not traditionally archive books or hold significant collections of volumes that are not available to be borrowed from their buildings – they do not preserve and store books for posterity. In some larger libraries such as the New York Public Library, there is some 'incidental preservation' – where printed publications have been retained for decades simply by not being thrown away.

Even so, this is not something that these libraries undertake systematically, and they are currently neither skilled nor resourced to do the kind of intentional preservation at scale required for e-literature, without a significant change in their skills, remit and spending.

Self-evidently, time and use limits have had negative implications for public libraries' attitudes around responsibility for the long-term preservation of, and future access to, e-books.

2.6 Open Access

We asked many librarians how Open Access (OA) might affect their attitudes to preservation. We find that it is seen as likely to be a positive force, in that it makes more copies of e-literature more widely available. In theory this should also mean that there are a greater number of copies held in preservation archives. The hope – if not expectation – is that a significant number of libraries will come to adopt the LOCKSS system, open-source preservation software that is freely available, and dedicate a proportion of their storage capacity to the archiving of other published materials drawn from the wider community, thereby making a reality of the aspirational maxim that 'Lots of Copies Keeps Stuff Safe'.

One advantage of OA publications is that, unlike journals which are hidden behind paywalls, it has been possible to complete a more comprehensive survey of what OA journals have been lost. We refer to research by Michael Laakso, Lisa Matthias, and Najko Jahn, published in their paper 'Open is Not Forever'. Using the Internet Archive's Way Back Time Machine they identified 174 OA journals that had vanished from the web by 2020. They also suggest the true figure is much higher, because OA journals face unique challenges.

First, publishers have no financial incentive to invest in the preservation of content that is freely and widely available. Second, OA publishers are often small and, to save money, might use a less thorough approach to preservation e.g. invest in less redundancy. We therefore find there is a very long tail of OA publishers that are too fragile and underfunded to consider the kind of long-term planning necessary to establish systems for preservation.

The authors of that paper conclude with a call for action: "We close with a note on the urgent need for collaborative action in preserving digital resources and preventing the loss of more scholarly knowledge." Since the publication of the Lassko, Matthias and Jahn report, the preservation service JASPER has emerged to provide advice and support to OA publishers and authors on how to preserve their works.

JASPER seeks to prevent the disappearance of OA journals by creating a robust, sustainable preservation network. It does this, in part by partnering with digital preservation services like CLOCKSS, Portico, and the Internet Archive, to ensure that OA content is securely archived and accessible for future generations. We return to this in our recommendations.

2.7 Dark archives

Dark archives have emerged in response to the practice of issuing e-literature licences that do not permit libraries to hold archive copies of their own. Instead, a publisher places a preservation copy of its output in a dark archive, from where it will be made accessible to perpetual licensees, but only after an agreed trigger event – essentially if the publication has been unavailable for a continuous period of time. These dark archives (aka third-party archives) are all presently operated by non-profit organisations. The two largest are CLOCKSS and Portico.

There are key differences between these two organisations, in particular the manner in which they are funded, how they acquire e-literature to preserve, how they preserve it, how and where they store it and how, after an agreed trigger event, they release it. But it is beyond the scope of this report to compare these in depth, and we have found no significantly different vulnerabilities that arise as a result of their core operating models.

E-literature in dark archives is protected by seemingly robust preservation systems, with multiple levels of redundancy, and stored in multiple locations – in the case of CLOCKSS there are as many as 12 separate locations distributed across eight nations, (United States, Canada, United Kingdom, Germany, Japan, Australia, Hong Kong and South Africa). The content in storage is regularly checked for discrepancies which, once identified and where it is considered necessary, are diligently repaired.

The processes are well established and consistently adhered to and therefore, for as long as these organisations have the committed support of both publishers and libraries and the resources to continue operating, in the words of Kate Wittenberg, Managing Director of Portico, "There is no good reason for something to be lost". Although, after consideration, she added the proviso, "As long as everyone does what they are supposed to do".

Dark archives are most commonly used for academic e-journals, with CLOCKSS reporting 25,232 titles and Portico 34,660. (Both figures taken as of 13th May 2024 from Keeper's Registry.) However, according to research conducted by CLOCKSS themselves, only 29% of e-journals are currently preserved in dark archives, although the authors consider this figure to be somewhat misleading, as the five major publishers, (Elsevier, Springer Nature, Wiley, Taylor & Francis, and SAGE Publishing), representing an estimated 50-70% of all e-journals by volume, are all either utilising the services of CLOCKSS and/or Portico, or they are making similarly robust arrangements elsewhere. There is however a long tail of e-journals from relatively small publishers, with very small readerships, and this is mostly not represented in dark archives.

This information is supported by our research. Many of the contracts for e-journals that we read include a contractual commitment by the publisher to use at least one named dark archive – often more than one. (See Appendix C.) This is further supported by evidence from our questionnaire as revealed in Table 5 below. Almost seventy percent of the research and academic librarians who responded (we are excluding circulating libraries from this result) believed that the e-journals they subscribe to are preserved in a dark archive.

To date, the dark archives have been proven to work and, at the time of writing, more than 250 e-journal titles that might otherwise have been lost had been triggered and made re-accessible by one or other of the dark archives. However, it is largely in academic e-journals, where licences most typically promise perpetual access, that the use of dark archives is concentrated. Although some e-book publishers also use dark archives to preserve scholarly monographs, this is not usually as a result of a contractual commitment to libraries.

One university librarian interviewed said, "As most of our e-book packages have no PCA (post cancellation access), there is not much point in looking at preservation clauses, i.e., whether they use a dark archive such as CLOCKSS or Portico." This is confirmed in Table 6 below, which reveals that less than forty percent of research and academic librarians believed their e-books contracts with publishers usually included a commitment to use dark archives. The equivalent share for e-journals was almost seventy percent. The challenge of creating the behaviours and resources necessary to systematically deposit scholarly monographs and other e-books into dark archives is an issue we raise in our recommendations.

Generally, academic publishers see less long-term commercial value from e-books than they might from e-journals. For example, they can much more easily bundle and monetise the back catalogues of their e-journals, maintaining continuing demand and therefore a commercial market, while e-book demand rapidly dwindles or falls to such low levels that the cost of digital maintenance can easily outweigh the potential for revenue.
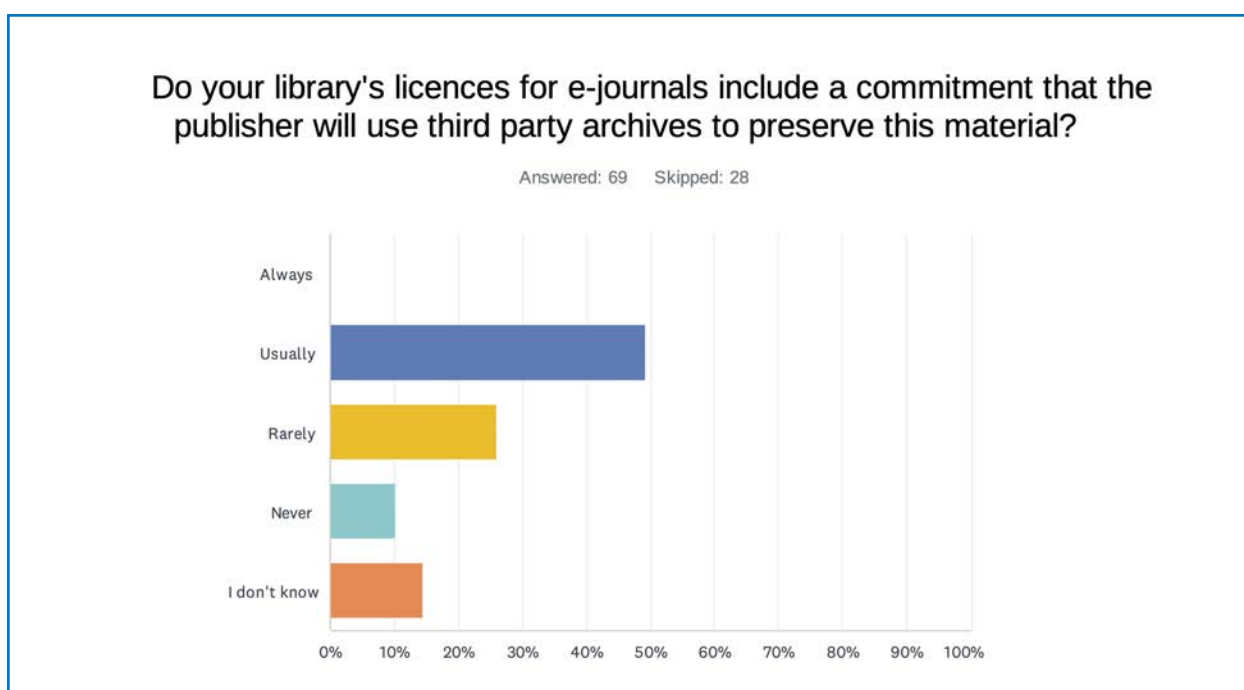

Table 5

One publisher explained that it is occasionally the case that once e-books are considered no longer profitable, they may be offered back to their authors (often with the provision that the author must pay up for the preservation file). If this offer, which generally includes a rights reversion, is not taken up, the publisher will consider deleting and ultimately destroying the preservation files, or at least no longer investing in the future maintenance of them which, over time, leads to much the same outcome.

In the trade books sector, we found no evidence of the use of dark archives.

It's important to state here that dark archives are not libraries. This is, in our opinion, the main reason why it has been possible to persuade publishers to supply pay-walled e-journals to these archives. On the other hand, making preserved e-literature accessible presents a far more complicated and multi-faceted legal challenge. As Kate Wittenberg of Portico succinctly states, "[Future] access is far more complex than preservation."

As a result, the fundamental problem for librarians is the inaccessibility of the dark archives. Those with concerns about issues such as format obsolescence, version control and link rot, are unable to assess for themselves whether material in the dark archives is secure. These risks increase significantly over time, particularly as many of these works may be sought out very rarely – perhaps once in a decade – and by then it will have been too late to take remedial action if a problem is discovered.

While we strongly commend the dark archives – which are in part governed and overseen by librarians or ex-librarians – reliance upon this system, and only on this system, seems to many a poor substitute for having control of the material on their own shelves. For Martin Eve, Professor of Literature, Technology and Publishing, at University of London's Birkbeck College, "The relationship between dark archives and publishers is a bit too close". And as one librarian said, "If something is preserved in a big black box, [and you cannot get to see inside it], how does that help to reassure anyone that it's all being truly safeguarded?"
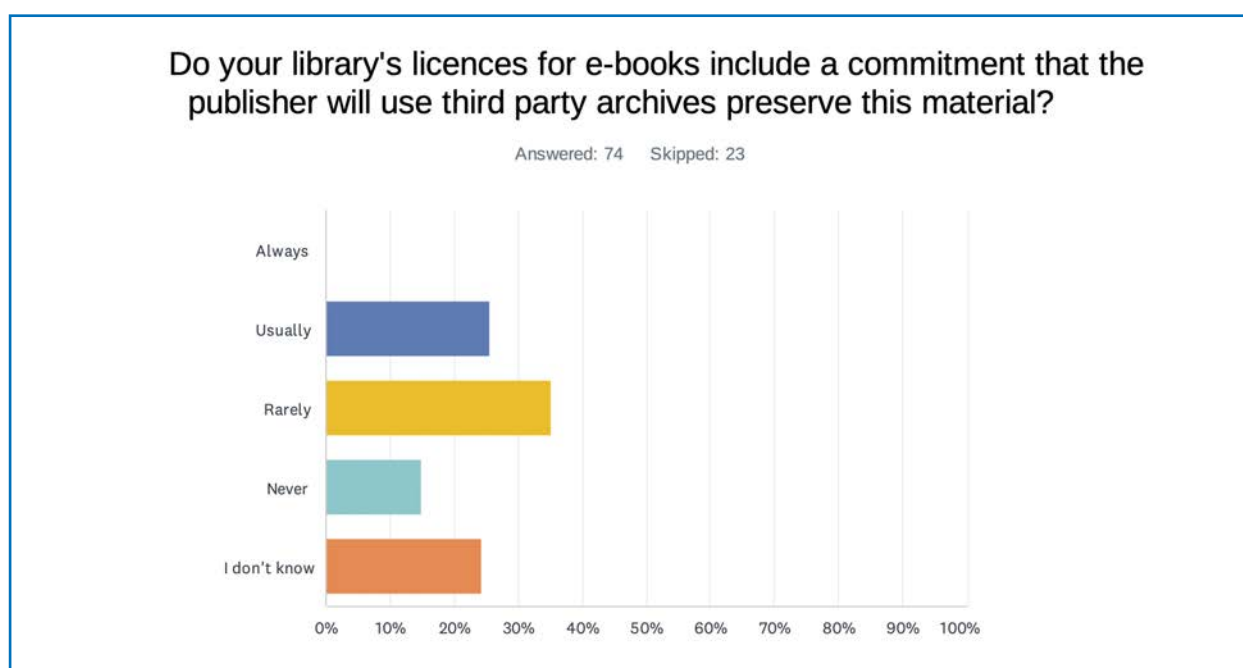

Table 6

This perspective invites us to reflect on whether Portico fully meets its intended purpose, for while the term 'portico' implies an open doorway – an entrance – as a dark archive, Portico seems to be a door that is almost permanently closed. Similarly, we must consider whether CLOCKSS aligns with the expectations implicit in its name. CLOCKSS preserves 12 copies of any given object, all managed within a single repository system.

Although this is substantial, it might not fully embody the idea of 'lots' of copies. Nonetheless, the fundamental goal of these dark archives is to ensure perpetual preservation of these materials with an expectation of future access.

We spoke with the senior management of both CLOCKSS and Portico, asking them to look ahead and speculate on future publisher behaviour. Specifically, we questioned whether publishers might increasingly offload their preservation responsibilities to the dark archives, using them as a convenient way to avoid the ongoing costs associated with archiving their out-of-commerce titles.

The responses were startling – this is already happening. Publishers are beginning to treat these dark archives not as a last resort in the event of catastrophic failure, but as their primary, and sometimes only, preservation strategy.

This revelation is dismaying. The dark archives were never intended to shoulder the full burden of preservation. They were designed as a safety net in exceptional circumstances, not as the standard fallback for publishers seeking to defer their responsibilities. It's akin to using an emergency exit as the main doorway or equipping passengers with parachutes instead of ensuring planes can land safely. This misapplication of the dark archives signals a significant departure from their intended role, and if left unchecked, could soon make this misuse their main function.

Yet, rather than viewing this development solely as a problem, we believe there may be an opportunity here—one that could transform a critical weakness into a significant strength. With appropriate support, investment, and collaboration from the sector, the dark archives could seize the initiative and redefine their purpose. Instead of being passive repositories of last resort, they might evolve into active, open, and globally accessible libraries. These institutions could offer enduring access to works that have been lost or are no longer commercially viable, providing a vital resource for scholars, researchers, and the public.

Such a transformation would not only enhance the accessibility of e-literature but could also pave the way for the dark archives to become commercially self-sufficient, not-for-profit organisations. By turning their passive role into an active mission, the dark archives have the potential to serve the global community in a far more dynamic and impactful way, ensuring that valuable works remain accessible for generations to come.

Finally – and crucially – it's not clear whether the public's right to access material contained within the dark archives will change when the material passes out of copyright and enters the public domain. An unanswered question remains: Is this when the responsibility for preservation (and storage?) of e-literature by the dark archives ends? And, if so, does that mean that all material they hold at that date is automatically triggered, either via an Open Access model or simply distributed to any and all who may want to possess a copy? Or is the end of copyright a concept that does not exist for the dark archives? Alicia Wise of CLOCKSS explained that they are committed to long term preservation and access but because the public domain is "a bridge we haven't yet crossed" … "We don't have a policy on this".

In conclusion, we find that the dark archives may offer the potential of an effective – albeit currently partial – solution to the preservation of e-literature. Among the many flaws in this system is one emerging concern: that keeping lots of digital copies of e-literature, while protecting knowledge, may not be environmentally sustainable because it requires running many more servers, requiring more power and, on today's technology, much more cooling.

Critically, we also find that the use of dark archives is contingent on contractual arrangements that may change or that may not be upheld. As a result, we cannot, at this stage, provide reassurance that the dark archives present an equivalent guarantee of long-term preservation and ongoing access as the print-based libraries historically have had, through owning and controlling the literature on their own shelves.

2.8 In-library preservation

Research, academic and public libraries have not, do not, and for the foreseeable future will not, preserve electronically published material. There are legal and technical reasons for this.

The legal reason is exposed by our questionnaire. 70% of librarians said they were rarely or never allowed to archive the e-literature they licence, with a further 20% admitting they did not know for sure whether this was or was not permitted. We return to this in our recommendations.

The core reason that many or most libraries do not preserve e-literature is that the resources, skills, and workflow processes required, for the long-term stable and secure preservation of digital files and formats, have not been developed within most libraries, not even in many of the very largest academic and research libraries. Despite the fact that LOCKSS preservation software is open source and freely available few, if any, academic libraries have adopted the levels of understanding and commitment to digital preservation that are equivalent to their capacity for physical services. Both CLOCKSS and Portico offer consultation on how to set up and manage a digital preservation environment, so in principle the establishment of an electronic archive should be achievable. Yet, to date, this is an area that remains, in the main, under-invested and under-funded.

2.9 The role of research and academic libraries

The concept of digital preservation as a core responsibility for research and academic libraries garnered mixed responses from interviewees for this report. While digital preservation is not yet universally embraced, it's undeniable that these institutions serve as repositories for a vast corpus of printed literature, in particular because today most scholarly e-books also have a print counterpart, thereby safeguarding a substantial body of knowledge that might otherwise have been already lost. In this regard they remain, at least for now, a cornerstone for scholarly access to historical information.
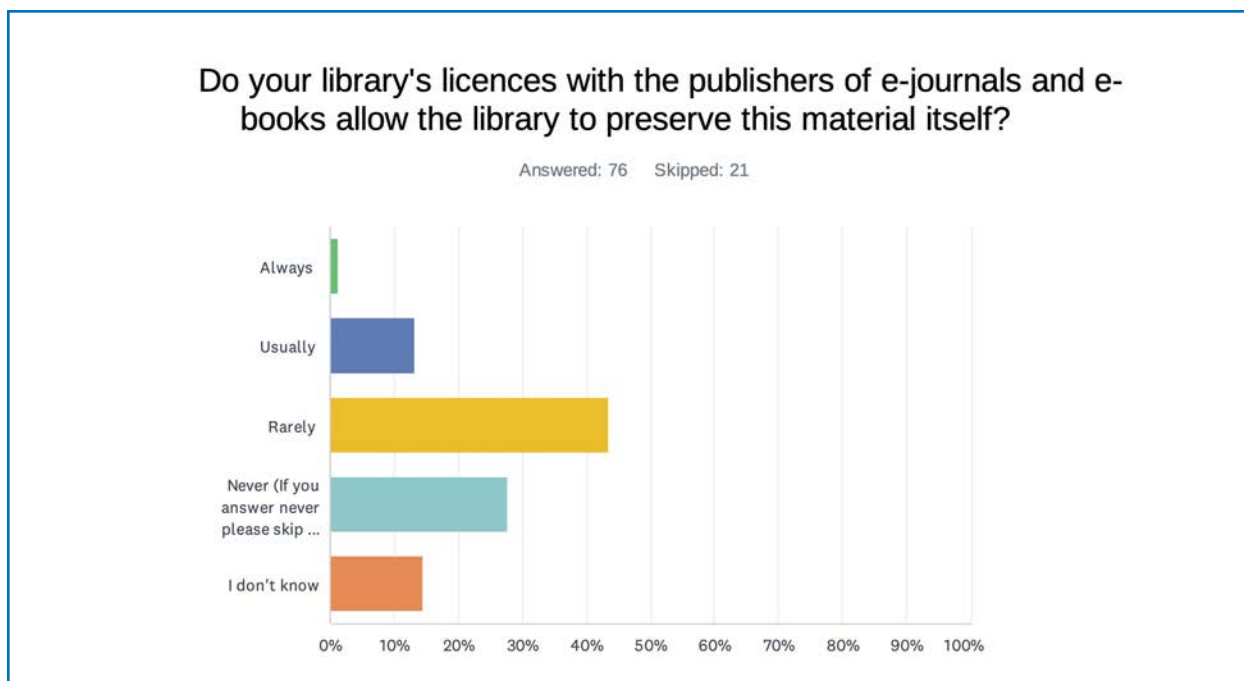

Table 7

However, a concerningly different picture emerges with regard to the preservation of e-literature. Many librarians we interviewed acknowledged limitations in fulfilling this duty for their digital collections. In fact, our librarian survey revealed that over half of respondents do not consider

preservation a key activity for their current digital collections at all. (See Table 8). Preserving e-literature simply isn't standard practice among libraries.

This doesn't negate their role in digital preservation and ongoing access entirely. Librarian advocacy has clearly influenced the publishers to commit to dark archives – a significant concession. As one publisher acknowledged, while an added cost for commercial entities, dark archives are seen as a necessary investment in customer satisfaction.

While some librarians perceive the digital shift as a threat to their core mission, we suggest that their role is now more crucial than ever. They can and should remain the guardians and educators of information literacy, while simultaneously advocating for – and enforcing, when they have the opportunity – sound preservation practices that are intended to guarantee future access.

Librarians may not possess the same ownership and control over e-literature that they enjoy with physical collections, but they are the most credible advocates to ensure its preservation and enduring accessibility.

2.10 The role of public and circulation libraries.

Traditionally, public libraries have not prioritised book preservation beyond local history materials or rare, low-printrun publications. This approach makes sense when you consider their mission: ensuring broad access to a diverse range of materials for their communities. Libraries prioritise acquiring new titles to meet evolving patron needs, and space limitations often necessitate weeding out less-used or out-of-date materials. Additionally, the inherent longevity of physical books, particularly those on acid-free paper, meant less urgency for active preservation measures. Libraries could rely on well-designed storage and judicious handling to maintain collections in good repair for extended periods.

Furthermore, preservation of historically significant materials was often considered the domain of national libraries and archives, or historical societies. Public libraries, with their focus on current and popular materials, simply weren't expected to shoulder this responsibility.
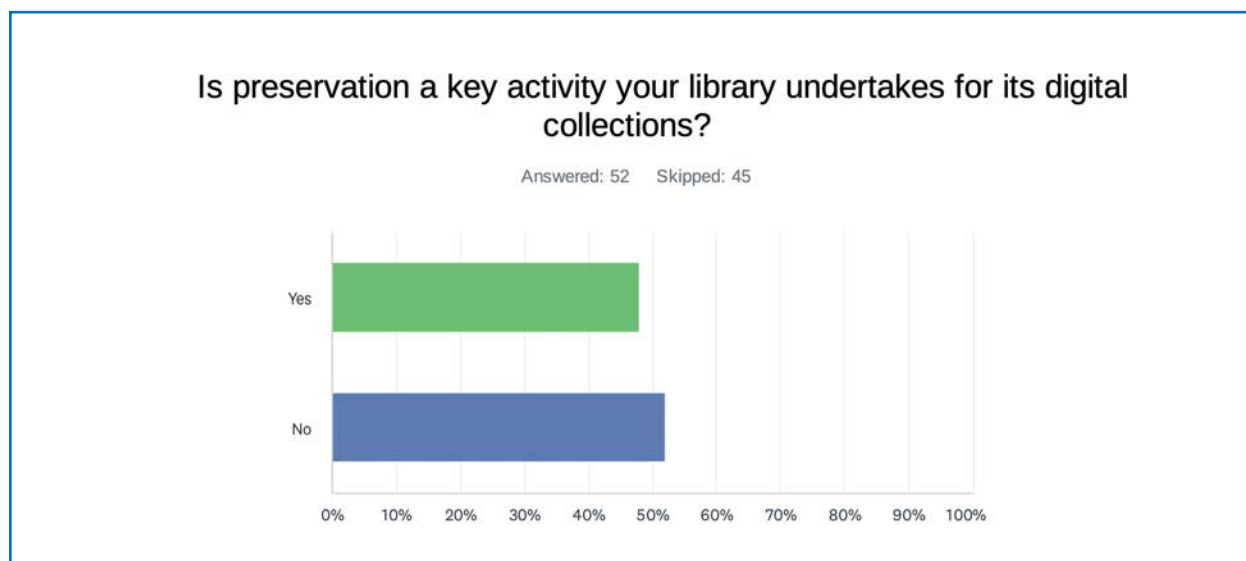

Is preservation a key activity your library undertakes for its digital collections?

Answered: 52   Skipped: 45

Table 8

However, the digital landscape presents a new set of challenges. E-literature, unlike physical books, is vulnerable to changes in access models, hardware incompatibility, and format obsolescence. The rapid churn of e-literature further complicates matters. Public libraries almost entirely rely on third-party vendors to provide end-to-end e-book lending services, meaning they never actually come into contact with the e-books that they 'lend' to patrons. If responsibility for preservation resides anywhere, it is with the vendor, not the public library. While some vendors may claim they systematically archive e-books – and there is wide scope for interpretation of what they may mean by such a claim – there is no guarantee of perpetual future access on behalf of the library or its patrons, especially if a title goes out-of-print or if the terms of licensing agreements change.

This raises concerns about the long-term preservation of trade e-literature within library collections. Public libraries, accustomed to a model of ownership and curation for their physical materials, at least for the duration of patron demand, lack anything like the same degree of control over digital resources. Furthermore, the sheer volume of e-literature production in this sector makes the concept of any individual library undertaking comprehensive in-house preservation a daunting task.

Should it ever become the case that a significant volume of the books lent out by public libraries is published exclusively in digital formats, with no print equivalent, then the indications are that the preservation of these works would be in an even more precarious state than their academic counterparts, due to the fact that few, if any, public libraries could successfully preserve these publications.

2.11 The role of national libraries

Most national libraries are given, or otherwise obtain, a copy of every work of e-literature published within their jurisdiction, which they are legally mandated to store and preserve in perpetuity. In this respect their role may be regarded as more that of an archive than a library.

Whereas a library collects and makes its resources available to the full membership of its community, an archive keeps objects primarily for posterity and for evidence. We found national libraries to be the only institutions where not just selective preservation, but the preservation of everything forever, is systemic. This is not contingent on intermittent practices or restricted by contractual arrangements: they are operating under the laws of legal deposit.

Historically the core purpose of legal deposit has been simply to retain a record of written culture and knowledge, offering some degree of limited access to the public, but not necessarily to provide for wide scale availability. Our research revealed that, in the majority of cases, the current practices of national libraries, compounded by the legal framework under which they receive the works, actively restrict access to e-publications to on-site only, and often via a limited number of library-owned devices.

Limiting access to e-literature is an understandable attempt to replicate the physical limitations of print material, as well as to prevent copyrighted work leaking out and spreading across the internet. However, it appears that these restrictions are often applied to electronic material that is published under Open Access and even when material eventually enters the public domain, meaning that despite having the potential to be the Library of Last Resort, there is no expectation on the part of many national librarians to replenish the scholarly publishing eco-system with missing or lost works, by default. A spokesperson for the British Library stated that, under present arrangements, all forms of legal deposit literature would always have very restricted access, even once it entered the public domain.

However, we encountered a range of views on this. Kate Zwaard, of the Library of Congress, said to us, "The purpose of collecting is so that people can use it. Otherwise, you're just a dragon hoarding gems. We're always thinking about access." Alison Stevenson of The National Library of Scotland said she expected legal deposit material should (eventually) be made widely available, not just when it enters the public domain, but ideally when it is of no further commercial value to the rights holder. "I think access is fundamental to the purpose of national libraries", she said, but also added, "Some publishers see legal deposit primarily as a preservation activity with access to both print and digital material necessarily limited to onsite at a legal deposit library.

The challenge for national libraries is to demonstrate the value to the public of wider online access to non-print legal deposit content, which is made freely available at the point of publication (e.g. OA e-books or web content) as well as that which is out-of-copyright."

Johannes Neuer, Director of the German National Library in Leipzig was less committed to providing unlimited global access. "Our current mandate is not to create a public domain". Although, as Clifford Lynch of the CNI says, "You'd think as cultural champions the national libraries would be promoting the journey of works into a meaningful public domain", yet: "The national libraries have been disgracefully silent about this".

However, at the far end of the spectrum, some national libraries are indeed beginning to see a future role in offering such a public service and are already developing systems and operational processes that are designed to retain the integrity of a published item, throughout its publishing lifecycle, with a vision of providing perpetual public access, albeit at some point in the distant future.

In any case, wherever they sit on this spectrum, national libraries are the ultimate failsafe in the preservation of knowledge, whether it is in print or, as is increasingly the case, when it is published exclusively as e-literature. As Johannes Neuer says: "We are the active memory of society". … whose job is to: "Keep [copies of deposited works] in original form and in perpetuity". While Leslie Weir, the Librarian and Archivist of Canada, reassuringly spoke of an obligation to preserve objects: "for hundreds of years".

2.12 Legal deposit preservation

All electronic literature is now typically placed in the archives of national libraries under legal e-deposit obligations. The years in which this began, and the specifics of the law, vary among jurisdictions, but it's now the accepted norm. However, we find that some national libraries are more confident in their long-term storage systems than others.

Despite the fact that it is now widely accepted that the preservation of e-literature requires an ongoing commitment to building and maintaining systems and processes entirely different from their print equivalents, this is not always how their digital archives have been established. In one national library (the librarian wished not to be identified), the move to digital archiving was conceived as a cost-saving measure and consequently the overall archiving budget was actually reduced.

Therefore, until there is a genuine change in this approach, we remain doubtful that the many national libraries holding electronic material on legal deposit are adequately resourced to ensure this material will be effectively preserved in perpetuity, with particular concerns for the US and the UK which house two of the most significant collections.

This said, a major cyber-attack on the British Library in late 2023 has at least had the positive side-effect of igniting a long and hard look at the systems with which electronic publications

around the world are being preserved. (At the time of writing, e-literature held on legal deposit by the British Library remains inaccessible, since the cyber-attack.) But, just as with other archives, the major vulnerability of these legal e-deposit systems is in their 'black box' inaccessibility. As George Wright, a library and metadata consultant in the Netherlands, observes, "The trouble with the legal deposit system is that it's inaccessible, so nobody knows if it's stable." Totally open accessibility (which risks a wholesale breach of copyright) may create vulnerabilities for sure. But so does total inaccessibility.

2.13 Format obsolescence

Long-term storage and retrieval of digital publications, considered over a timeframe of 20-50 years and longer, faces several challenges related to the obsolescence of the format within which it is contained. However, many of the interviewees felt confident that, with appropriate strategies and practices, the risk of loss of the raw text itself can be significantly reduced and, as a result, the general consensus is towards a feeling of confidence that the unformatted, raw text of e-literature, when considered as a separate entity in and of itself, is not vulnerable, per se, to format obsolescence.

One complicating factor that cannot be sidestepped is the risk of corruption to the physical storage media – e.g. hard drives, CDs, etc. on which the text files are stored. However, we find, and list here, strategies that are being used in what might be called model preservation systems, and which contribute to this sense that the text itself can be made fairly, or very, secure.

- Standardised formats: Using standardised and widely adopted file formats such as PDF/A (an ISO-standardised version of PDF) or plain text (TXT) helps to ensure long-term readability. These formats are considered unlikely to ever become obsolete and are supported by a great many software applications.

- Format migration: Regularly migrating data to newer formats helps mitigate the risk of loss. This, of course, requires diligence and, more importantly, continuous funding to ensure it is a process that is maintained indefinitely.

- Emulation environments: Emulation involves re-creating software that mimics older hardware or software tools or programmes, allowing 'obsolete' formats to be opened and accessed using modern day systems. Raw text in particular requires relatively simple tools to open, access and/or reformat files for on-screen reading.

- Metadata: Including comprehensive metadata alongside or within the digital files that contain the raw texts helps to maintain context, provenance, and key technical details, which will be essential for future retrieval and interpretation.

- Regular audits and checks: Periodic audits and refreshing of digital storage media help identify and remedy potential issues before they can turn into problems that result in data loss or other format obsolescence.

- Redundancy and geographic distribution: Storing multiple copies of digital text in geographically distributed locations (as with the CLOCKSS approach), reduces the risk of data loss that might be caused by localised disasters or digital degradation.

While these strategies significantly improve the long-term stability of digital text, adherence to standards, ongoing vigilance, and the rapid adaptation to newer technologies and formats are going to be essential to mitigate against future loss.

The nature of digital text also means it is possible to preserve very large quantities of material. At the time of writing the Keeper's Registry, which monitors the archival status of continuing resources, notes that 93,580 titles have been ingested into some form of archive. (Figure taken on 9th July 2024.) These numbers are dwarfed however by other electronic publishing initiatives. For example, OverDrive, which mostly distributes trade books to public libraries, claims to ingest about 50,000 books every month – 600,000 every year. (This figure was reported by Steve Rosato of OverDrive.) Google claims to have ingested around 40 million books, at first mainly as image files but increasingly these are being converted to text.

The formats in which these files have all been captured and stored are a 'trade secret' and it was not possible for the authors to determine whether these files are being preserved indefinitely with or without considerable, continuous investment. However, it is the opinion of the authors that Alphabet is unlikely to maintain the integrity of these files beyond their commercial use and it is therefore a major risk that, at some point in the next 50 years, the files will be degraded, lost, deleted or, potentially, transferred to another entity – who may not grant the same degree of free, public access as Google currently does.

2.14 The loss of functionality.

We have only so far been concerned with the preservation of the written word – essentially text. But we are warned that an increasing amount of scholarly output including scholarly monographs will soon make use of the greater potential of digital tools, with the inclusion of other objects such as video, embedded links, research software and even live data sets becoming ever more common. As already mentioned, we can find no data that quantifies this development either in e-journals or scholarly monographs, although anecdotal evidence from our interviews suggests this kind of multi-media work is on the increase despite there being no systems or established processes to ensure its preservation.

This, of course, presents a much greater challenge to those seeking to preserve it intact. In fact, most organisations that we interviewed did not have anything approaching the technical capacity to preserve such works with full functionality. Audio and video might occasionally be considered for preservation but usually captured at low resolution. Instead of an active or live data set, a snapshot of the data might be stored, representing a moment in time. We find this not only a concern but illuminating about how digital technology continues to evolve, increasing in its complexity, while outpacing our ability or technical capacity to preserve it.

2.15 The fragility of links

Even a simple text file has an inherent vulnerability: future readers may be unable to follow hyperlinks to the referenced works, not necessarily because the works themselves are lost, but because the links have decayed or become non-functional. This is much more than an inconvenience. Martin Eve, Professor of Literature, Technology and Publishing, University of London's Birkbeck College, was unequivocal on this point in particular: "Our contemporary epistemology or system of research rests upon footnotes and reference links – which are the only way to verify the statements of the past".

His fear is that a failure to preserve these links, or to make sure that the referential works cited are accessible to scholars and researchers, threatens the way that knowledge is created. As Isaac Newton wrote in 1676, in a letter to his fellow scientist Robert Hooke, "If I have seen further, it is by standing on the shoulders of giants".

The sudden and rapid switch to e-literature, especially in STEM subjects, introduces both a powerful new engine for knowledge and progress and, at the same time, a serious threat to the fundamental way human knowledge has been created and built upon since the discovery of the scientific method.

2.16 The need for vigilance

We find that too often the preservation and provision of ongoing access to e-literature aims to replicate the archiving conditions, practices, processes, expectations and outcomes of the print era. This does not appear to appreciate the substantial increases in resources, skills and continual reinvestment, necessary to maintain the integrity of digital archives, now and into the future.

Whereas print works can endure decades, centuries even, of 'benign neglect', allowing them to gather dust on the shelves of academic libraries simply by not being thrown away, a similar laissez faire approach will not serve to preserve e-literature. This requires much more than just keeping the servers on. It demands vigilant inspections of the material and active interventions. This approach is currently the attitude and approach of the dark archives, but it is implausible that equally effective efforts are being applied across all digital preservation initiatives taking place in libraries around the world. This is especially a concern as, whereas a print work can sustain some damage but remain usable, even the slightest damage can be catastrophic for the integrity and viability of digital files.

Such vulnerabilities are exponentially increased over time. Copyright often lasts a hundred years, and there are very few publications that remain commercially viable for that length of time. As Clifford Lynch of the Coalition for Networked Information Print said, "This is an enormous window of vulnerability".

2.17 The threat of hackers

The October 2023 attack on the British Library by the ransomware gang Rhysida, targeting the personal data of users, has had a disastrous, global impact on the scholarly communities' access to parts of their collections. That said, useful lessons can be learned, for which reason we include a brief analysis of this event, based on our interviews and the BL's own report and a further instance of analysis by librarian Simon Bowie of the LSE.

The Rhysida attack was what is referred to as a double ransom. The hackers threatened to both expose personal data of staff and users as well as corrupt the underlying systems. When the BL refused to pay, Rhysida followed through on their threat. According to the BL's own report, released in March almost five months after the attack, 'Key software systems, including the library management system, cannot be brought back in the form they existed before the attack.' At the time of writing this report, almost a year since the attack, although some of the services such as remote ordering and digitised materials are coming back on line, patrons are still unable to access the legal deposit e-materials even within the main British Library building.

Among the findings of the BL's own report are the startling claims that, "The Technology department was overstretched before the incident and had some staff shortages" and "The Library's vulnerability to this particular kind of attack has been exacerbated by our reliance on a significant number of ageing legacy application". As Simon Bowie points out in his illuminating piece these admissions hint at an institution turning a 'blind eye' to the continual changes in technology, culture and risk, perhaps to avoid costs or to make alternative trade-offs. The BL's own report exposes a fundamental flaw in the way the British Library cared for its digital

infrastructure and suggests: "'Senior management neglected the British Library's core library systems'".

We discussed the British Library cyber-attack with Marie-Louise Ayres, Director General, National Library of Australia. She noted that, when she took up her position, she inherited a team with only one person dedicated to cyber security. Today approximately 15% of her digital payroll is assigned – whether directly or indirectly – to cyber security. In other words, to operate at an appropriate level of safety, this is an expensive process. For all institutions, especially those with a public service remit, cyber security must be considered an integral element of its operations. As a rueful British Library now admits, there is: 'The need to embed security more deeply than ever into everything we do'.

2.18 The uncertainty of a meaningful public domain for e-literature

In a future where digital technology is set to become the primary medium for literary expression and consumption, establishing a 'meaningful public domain' for e-literature will be more vital than ever. Traditionally, print publications have benefited from well-established library practices that smoothly transition works from copyright protection to the public domain, despite a period of restricted availability. However, the digital landscape brings a new set of challenges that could disrupt this process.

E-literature faces technical hurdles far more complex than those encountered with traditional print media. While the preservation of print works often resulted from a fortuitous combination of neglect and enduring practices, digital works are far more intricate. They consist of code, metadata, and media files that depend on specific formats, hardware, and platforms—elements prone to rapid obsolescence. The ephemeral nature of digital platforms means that works published solely in electronic formats may become inaccessible long before their copyright terms expire, presenting a troubling scenario where legal access is not matched by practical availability.

Achieving a meaningful public domain for e-literature will require tackling these risks head-on. Addressing issues like format migration, platform stability, and long-term digital preservation is essential for ensuring that future generations can benefit from works that eventually leave copyright protection. Additionally, the landscape is further complicated by legal ambiguities, such as orphaned or abandoned works, and problematic industry practices, including 'recopyrighting' through minimal changes. The bundling of public domain works with copyrighted material, reliance on paywalls, and proprietary formats or non-expiring DRM locks also create obstacles to access.

While these challenges are significant, they also present opportunities for innovative solutions. A comprehensive strategy that integrates legal reforms, technological advancements, and archival efforts can turn these challenges into a pathway for a vibrant and enduring digital public domain. With focused action and collaboration, we can ensure that our digitally published cultural heritage is preserved and accessible, echoing the successes of traditional print media.

However, to consider these significant obstacles as opportunities, it is crucial to resolve many of the challenges that will hinder the emergence of a meaningful public domain for e-literature:

Key Challenges to the emergence of a public domain for e-literature

- Loss of digital works: Unlike physical books, digital works are vulnerable to loss due to format obsolescence, media degradation, and platform closures. A work may be legally public domain, but if the only existing copies are on defunct disk drives or were

deposited on a platform that no longer exists, then the work is effectively inaccessible.

• Orphaned or abandoned works: The duration of copyright creates a long period of neglect during which digital works may come to lack clear ownership information, making it difficult or impossible to identify rights holders and obtain permission for re-use, even for works that would reasonably appear to have passed into the public domain. This creates a legal grey area and discourages risk-averse users.

• 'Re-copyrighting' through minimal changes: Copyright law protects original works of authorship. In theory, a rights holder cannot simply extend copyright by making minor changes to a public domain work. However, some commercial publishers may create new 'derivative works' by introducing minimal changes (like a new introduction or by re-formatting). This not only strongarms libraries to re-licence even if they have a perpetual licence to the original work, but it effectively re-copyrights the entire package.

• Version control: Digital works can be easily modified and updated, making it challenging to determine the definitive and intended version for preservation or future access.

• Bundling with in-Copyright material: Publishers may bundle public domain works with in-copyright materials within a single digital package protected by Digital Rights Management (DRM) technology. This makes it difficult, illegal or otherwise impossible to extract and use the public domain portion without 'breaking' the DRM and/or infringing on the copyrighted material.

• 'Paywalls' and 'lock-in': Even when public domain works are digitised, they may be placed behind paywalls or locked into proprietary platforms, making them inaccessible to users who cannot afford subscriptions or lack access to the specific platform.

• Obscure formats and technological dependency: Digital works may be created in formats that become obsolete over time, requiring specialised software or hardware to access. If this technology is no longer readily available, or capable of being emulated, the public domain work becomes inaccessible despite its legal status.

• Embedded code: Increasingly e-literary works incorporate multimedia elements, interactive features, or rely on code to function. Preserving these elements requires not just saving the text, but also ensuring the underlying code remains functional.

These challenges highlight the need for innovative strategies to ensure that public domain works created in the digital era remain truly accessible and usable, particularly because further complications may arise in the future, due to the nature of the licensing arrangements that determine library access. Due to the lack of 'sunset clauses' within the e-licenses, it is not certain that, even with solutions to the issues above, a meaningful public domain will exist for e-literature. The following unanswered questions necessitate further exploration:

• Licensing vs. ownership: To what extent do current e-literature licensing arrangements function as de facto extensions of copyright, effectively restricting access and use even after the legal copyright term expires?

• The Impact of bundling: How does the practice of bundling public domain works with in-copyright materials within a single, DRM-protected package impede user

access to the public domain portion?

• Paywalls and platform lock-in: In what ways might paywalls and 'platform lock-in' practices for digitised public domain works permanently limit user access and hinder the public domain's core purpose of fostering widespread availability?

• Orphan formats and technological obsolescence: How can long-term access to public domain e-literature be ensured in the face of potential format obsolescence and technological dependency on outdated software or hardware, if there is no obligation on the part of rights owners to maintain accessible files or editions?

• The role of libraries and archiving institutions: What role can libraries and other archiving institutions play in advocating for user access to public domain e-literature and in developing strategies to mitigate the challenges outlined above?

• Beyond licensing models: Are there alternative legal or technological frameworks beyond traditional licensing models that could better support a robust public domain for e-literature in the digital age?

3. Commendations:

Our investigation into the long-term preservation of, and future access to, electronic literature reveals a critical juncture. Broad agreement exists among stakeholders regarding the seriousness of the issue, alongside an accord that the current approaches are unlikely to deliver a universally satisfactory solution. However, in this section, we present instances of good practice that are helping to preserve electronic literature, even if only partially.

3.1 Previous research

We are aware of the debt owed to those who have already researched this issue. We commend the many investigators who have gone before us, and who will likely continue their work into the future, despite many being in full time academic posts. While far from a comprehensive literature review, a selection of the key articles read, and sites visited for this report is in Appendix F.

3.2 Dark archives

We commend the dark archives and also the LOCKSS programme at Stanford Libraries. First, dark archives are doing vital preservation work. Already titles that might otherwise have been lost have been saved by these organisations. Second, the dark archives, and the LOCKSS software, have pioneered approaches to the technical challenge of preserving electronic literature. But this was more than a technical challenge, and we likewise commend their work in convincing sometimes reluctant publishers to use their systems. Third, the dark archives provide a positive instance of collaboration between publishers and libraries.

The LOCKSS program, built upon the principle of 'Lots of Copies Keeps Stuff Safe,' represents a significant step towards safeguarding digital content. Its core idea – widespread distribution across multiple locations – rightly acknowledges the inherent vulnerabilities of centralised repositories. Historically, this principle has proven its worth. The vast print runs of traditional journals, distributed across libraries worldwide, ensured a level of redundancy that offered some protection against loss.

However, we must also observe that a crucial distinction emerges when translating this principle to the real world. While LOCKSS embodies the aspiration of 'lots of copies', it lacks clear parameters defining 'lots'. There's currently no verifiable evidence that any e-journal, Open Access or otherwise, is being preserved within a network of independent LOCKSS instances that even remotely approximates the sheer number of physical copies distributed under a traditional print journal model (often numbering in the hundreds or even thousands).

Simply deploying the LOCKSS software is not enough to guarantee the widespread replication envisioned by its core principle. The reality may be a far cry from the ideal: while there is certainly a multitude of LOCKSS software installations around the globe, we find a considerably smaller number of these installations actually replicating the same e-journals across thousands, or even hundreds, of these instances.

The internet's reputation for fostering viral distribution – as exemplified by the phenomenon of 'viral memes' – can be misleading. Widespread distribution does not equate to uniform preservation across all digital content. The name 'LOCKSS' itself might inadvertently create a false sense of security, implying a robust safety net for e-books and e-journals based solely on the software's wide distribution.

We hedge this commendation with the caveat that some believe that the dark archives are too close to the publishers. We note the recent case of the science journal Heterocycles. This important resource was unavailable from the original publisher for over six months before it was eventually triggered by CLOCKSS in June 2024. Explanations have been offered, and accepted, for this delay, yet some harbour suspicions that a possible sale of this business was put ahead of the needs of scholars for prompt access to over 17,000 articles in Heterocycles.

3.3 Open Access

We have high regard for all those involved in the Open Access movement, which champions the free online availability of scholarly research, aiming to democratise knowledge and expedite scientific progress, by removing financial barriers to access and equity.

Although, by itself, this has no immediate direct bearing on preservation and perpetual access, the growth of Open Access demonstrates how – not without determination and resilience – established arrangements within the publishing industry can be challenged and superseded by new arrangements. Similar arguments to the case for Open Access can now be used to make the case for better access to electronic publications that are out of commerce and/or copyright.

The Open Access movement has certainly made tremendous strides in democratising access to scholarly research. Open Access journals have shattered paywalls, allowing a global audience to engage with vital knowledge, fostering collaboration and accelerating scientific progress. This revolution in scholarly communication is truly commendable. However, a crucial question emerges: is simply making research 'open', enough by itself to ensure its survival for future generations?

The recent study by Laakso and others, titled 'Open is Not Forever: a study of vanished Open Access journals,' raises a sobering concern. Their research identified 174 OA journals that disappeared from the web between 2000 and 2011, spanning a diverse range of disciplines and geographic locations.

These vanished journals, despite their initial OA status, became completely inaccessible, highlighting the vulnerability of digital scholarship in the absence of robust preservation measures. The report by Laakso and others emphasizes the critical importance of 'comprehensive and open archives' – without them, OA publications can become victims of digital amnesia, their valuable contributions lost to the scholarly record.

This is surely not an isolated incidence. Imagine a researcher meticulously building their case upon a foundation of prior studies, only to discover that a key OA journal referenced in their work has vanished without a trace. The knowledge contained within those vanished journals is no longer 'open' or accessible, jeopardising the integrity of the scholarly record and hindering future research endeavours.

The Open Access movement serves as a powerful example of the scholarly community's ability to challenge and reshape established practices within the publishing industry. The same spirit of innovation and collaboration needs to be harnessed to address the challenge of digital preservation. Just as Open Access democratised access to research, we now need a committed movement for 'Open Preservation' that ensures the long-term accessibility and survival of critical knowledge. We must learn from the cautionary tales documented by Laakso and others, and work together – publishers, librarians, and scholars alike – to build a robust digital infrastructure that safeguards the scholarly record for generations to come.

3.4 National libraries

National libraries around the globe deserve admiration for their unwavering commitment to public service and uninhibited access to information. They stand as champions of intellectual heritage, ensuring the preservation of cultural treasures for generations to come. This dedication extends beyond mere stewardship; it's a proactive commitment to the future, where knowledge is readily available for all.

Among these leading institutions, the National Library of Scotland stands out. As Alison Stevenson, their Director of Digital and Service Transformation, aptly stated, "Access is fundamental to the purpose of national libraries." This philosophy resonates deeply, reflecting a core value shared by many national libraries – that access to information is a public right, not a privilege.

The National Library of New Zealand echoes this sentiment. Tim Kong, their Director of Digital Experience, declared, "Citizens should be able to access this material, because it's their stuff." This embodies the principles and best practices of national libraries: recognising the inherent public ownership of cultural and scholarly resources.

It's important to acknowledge that the path to equitable and free access is not always straightforward. Challenges may be more pronounced in jurisdictions with larger publishing industries. However, the success stories of these visionary institutions serve as a beacon for others. They offer a blueprint for navigating these complexities and creating a more accessible future for e-literature.

The National Library of Australia exhibits such visionary leadership. Their National E-Deposit (NED) partnership program and the Trove platform through which the public accesses this material are testaments to a bold and successful approach. Here, future access to e-literature is prioritised from the outset. Authors and publishers are actively engaged in determining access levels and timing. This collaborative approach fosters a win-win situation, ensuring creators' rights are respected while maximising the potential for future public access. Furthermore, approximately 30% of deposited materials are immediately made freely available worldwide via Trove.

The National Library of Australia's commitment goes beyond immediate access. Remaining materials are either embargoed for a limited period or tracked for their eventual entry into the public domain. As Dr. Marie-Louise Ayres, Director General, states, "I feel extremely confident that we will be able to provide access long term."

This unwavering dedication to long-term accessibility sets a high bar for all institutions entrusted with safeguarding our cultural heritage in the digital age.

4. Recommendations – General

The world of print contained, by default if not design, the conditions necessary for a seamless convergence of the continual preservation of, and access to, knowledge and culture. Librarians, often within a single physical space, could efficiently manage both tasks. Archiving physical books requires relatively cheap ongoing maintenance, existing as they do in a stable format, ever accessible to future readers.

The digital world disrupts this model. E-literature requires discrete and costly storage solutions due to the near certainty of format obsolescence and legacy platform dependence. Access environments require continual attention to ensure compatibility with changing technologies and to recognise evolving user needs. Workflows must be regularly redesigned to accommodate ever-changing digital curation and migration activities.

Specialised skills are required to inspect and manage trusted digital repositories and to help users to navigate issues such as licensing and other access restrictions. Finally, continual, or at least regular up-grades are essential throughout, to safeguard against data corruption or loss, and to ensure for long-term accessibility.

These complexities demand the consideration of a more nuanced approach than simply attempting to replicate the objectives, conditions and activities required of the physical equivalent – a formal separation of the environments that house the core functions of preservation and those of access. This could even involve distinct institutions – e.g. controlled access archives specifically designed for long-term, secure preservation, and separated public access libraries optimised for user interaction and retrieval.

Implementations supporting this approach can be found in the growing adoption of the preservation storage repositories separated from public-facing service areas, now being designed and deployed by some national libraries. These repositories prioritise stability, long-term storage, and data integrity, employing specialist staff with expertise in digital curation and format migration.

This potential separation of preservation and access functions also embraces the challenges of cybersecurity. Digital preservation copies, to ensure their long-term viability and integrity, may necessitate far more secure storage within non-public-facing archives – environments significantly harder to penetrate. Conversely, accessible copies, by their very nature, need to be readily available to users. As many traditional print libraries currently operate as hybrids – part archive, part lending library – embodying a blend of both functions within a single environment, this has created an inherent tension and introduced risk.

However, because it is easily possible to differentiate long-term storage and near-term access services, there is an opportunity to build and operate separated and dedicated facilities, even with a single, overarching institution to manage both. Otherwise reconciling these distinct needs inside a single, user-centric infrastructure for digital materials, may prove particularly challenging for traditional libraries and potentially result in costly or worse outcomes.

The complexities of the preservation of digital literature are immense, and while this report outlines several key concerns, it is clear that a deeper, more thorough investigation is necessary to fully understand the scope and impact of these challenges

The research should also investigate emerging technologies such as cloud-based storage, blockchain for verification, and advanced redundancy models that may offer new preservation methodologies. By continuing to explore and refine best practices, this work will help ensure that

future generations have access to the vast and growing body of digital literature that is being created today. The findings from this deeper research can then inform and hopefully influence policy changes, standard-setting, and the allocation of resources to ensure that digital preservation receives the attention and support it critically needs.

4.1 Implementation of archival standards for electronic literature

We recommend the wholesale adoption of common international standards for preservation of electronic publications. Such standards should cover issues including, but not limited to, preservation software, storage systems and layers of redundancy. These criteria will assist those setting up archives and help libraries to select from among established archives.

Universal standards such as these already exist for the preservation of print literature, and best practices have evolved, over many decades, that are passed down in specific training programmes. Digital approaches need to follow this model in order to normalise a similar code of practice for the future of knowledge preservation.

We are aware of ongoing attempts to create and encourage the uptake of preservation standards and note among them the Digital Preservation Workshop. However, none has yet managed to introduce a user friendly and widely accepted set of standards that are being universally applied.

Electronic literature is in so many ways more vulnerable than print. Its preservation is clearly a more complex process. While ideally digital preservation should be done with the right expertise, we would not expect senior decision-makers within libraries' management to have acquired such expertise. Approved standards would help with the right choices. We particularly recommend that such standards be observed by national libraries, ensuring they fulfil their role as the final failsafe in the preservation process.

4.2 Development of affordable and user-friendly archival software

We recommend exploring the development of more affordable and user-friendly archival software systems, such as LOCKSS, to enable libraries to effectively archive digital literature and create collaborative redundancy. Such systems would allow libraries to preserve digital works independently and within the scope of their legal permissions, ensuring long-term access and safeguarding against potential loss.

To enhance the long-term preservation of digital literature, we advocate for continued development of affordable, user-friendly archival software solutions tailored to the needs of libraries and institutions with limited resources. While existing systems like LOCKSS have demonstrated success in preserving digital content, they can be technically complex and building the archival capacity and the associated retraining or further staffing makes preservation financially prohibitive for smaller institutions.

A key priority should be the creation of simplified platforms that streamline archival workflows, reduce operational overhead, and lower the barriers to participation.

Moreover, they should enable libraries to archive materials independently and collaboratively, ensuring that digital works are stored in geographically distributed locations, thereby reducing the risk of total loss due to local failures or legal issues. Collaborative redundancy – where multiple institutions share responsibility for preserving copies of the same digital work – would add an additional layer of security, safeguarding against data loss and ensuring the longevity of materials.

4.3 Support for and expansion of dark archives

We recommend further discussion, research and collaboration with the dark archives to discover new opportunities to grow sector-wide support for, and fund additional resources to build upon, their excellent preservation work.

Dark archives will play a pivotal role in the preservation of digital works, safeguarding materials that are not immediately accessible to the public but that remain crucial for the retention of long-term cultural, historical, and scholarly knowledge. These archives operate as a silent reserve, ensuring that irreplaceable e-literature content is not lost even if it is inaccessible. We encourage deeper engagement with dark archives, both in terms of research and resource allocation, to better understand how these archives can be further supported and expanded.

While universally considered essential, these dark archives will inevitably face challenges related to funding, reskilling, and the ongoing complexities of maintaining vast collections of digital materials over time. Collaborative initiatives between libraries, research institutions, and the private sector can help provide the necessary financial and technical resources to bolster their preservation work.

In particular, fostering partnerships with national libraries, academic institutions, and non-profit organisations could increase the capacity of dark archives, enabling them to store more materials and potentially support a broader range of digital content types, including multimedia works and complex interactive literature.

We also tentatively recommend engagement with a more radical proposal for the dark archives. Given our concern that the dark archives might become the default preservation system of publishers, perhaps they should embrace this future. What would the dark archive need to perform this important role?

4.4 Develop archival solutions specifically for Open Access materials

We recommend specific research into the possibility of developing affordable, easy-to-implement archival systems tailored to the particular needs of Open Access publications. Such solutions should include automatic archival redundancy to ensure that multiple copies are distributed across participating institutions by default. Following the approach set by initiatives such as JASPER, the project should also promote awareness within academia that publication alone does not guarantee the survival of knowledge. Long-term preservation and permanent accessibility must be integral to the publishing process.

While Open Access removes financial barriers to knowledge dissemination, it presents challenges for ensuring the long-term preservation of freely available content. Many researchers mistakenly believe that once their work is published under an Open Access licence, it is safe from loss or degradation. In reality, without dedicated archival measures, Open Access works are potentially more vulnerable to technological risks than their commercially published counterparts.

4.5 Development of copyright risk assessment for out-of-commerce works

We recommend the development of copyright risk assessment to encourage libraries to provide access to works that are out-of-commerce but not yet in the public domain.

To support an e-literature preservation system, but not contingent on it, we recommend research and development work on copyright risk assessment frameworks. A written work falls out of

copyright 75 years after the death of its author(s). By this time, finding the copyright holders may be next to impossible, especially so for works with multiple authors.

For this reason, libraries and others perform risk assessments: not to verify whether a work is out of copyright, but whether there is a reasonable likelihood that it is – in some cases whether it is demonstrably so comprehensively out of commerce that it could be treated as out of copyright. (While some marquee works retain commercial value up until their copyright ends, the vast majority of works lose commercial value much earlier.)

Presently, such assessments are done on a case-by-case basis. As such, it is often not viable to go through with this process or considered to be 'worth the effort'. However, if this process could be standardised, perhaps even programmed, libraries would be more ready to use it, thus enabling access to material that would otherwise be classed as not yet out of copyright, and therefore inaccessible. It could become standard practice for such assessments to be built into the life cycles of electronic literature, significantly increasing the volume of accessible works.

4.6 Preservation arrangements that reduce contingency

We recommend that all preservation arrangements lessen contingency so that works can survive shifts in the balance of power and/or withstand the passage of time until the emergence of a sustainable system that is fully resourced and enforced by both publishers and libraries.

We do not see an unfailing end-to-end system for preserving e-literature. There are certain contingent arrangements, most notably the dark archives, within certain sectors of the publishing industry and the library community, that serve this purpose. Should the balance of power shift, or economic models changes, such arrangements would likely be affected. We are not confident that such change will not happen (in the next hundred years or so) before these works enter the public domain.

It is also not clear what effect Open Access will have in the longer term. Therefore, we recommend that preservation shifts from being a clause in a contract that can be renegotiated, to being legally and culturally built into the systems of publishing and libraries.

4.7 Raising awareness among governments and librarians

We recommend raising awareness about this important issue so that governments, as well as librarians, are concerned about the loss of knowledge and will be more prepared to support plans to preserve it.

We propose a government awareness campaign to highlight the societal and cultural threats of knowledge being lost, and/or becoming less accessible, as a result of the shift to digital formats. Librarians in particular should be empowered to advocate for better preservation. Although this is not the case with e-journals, with other kinds of e-literature the lack of perpetual access has meant that libraries have no immediate incentive to fund or even demand preservation.

However, as the example of e-journals (the preservation of which is hugely thanks to advocacy from research libraries) shows, librarians can and do make a difference. Governments and decision-makers should be made aware of the danger of losing knowledge, and its impacts on their respective remits and agendas. National governments must be lobbied to help solve it. Given the other competing demands upon governments' purses, the need for such a campaign in order to gain their attention, becomes more urgent.

One proposal that might prompt governments to action is a documentary film or podcast that highlights these issues of preservation and ongoing access to e-literature. We admit to a bias, as one of the writers of this report is a documentary and podcast maker, but there is no doubt that the general public needs to be made aware that culture is in danger of being lost. We believe that citizens will be concerned about the vulnerability of the knowledge stored in electronic literature , and that this would come to matter to a significant proportion of them.

4.8 Granting libraries the right to download and/or print preservation copies

We recommend advocating for libraries to be granted the legal right to download and/or print preservation copies of digital literature for which they have already paid and secured perpetual access to, potentially through a collective legal challenge to current contractual limitations.

The preservation of digital literature is increasingly hampered by restrictive licensing agreements and contractual rules that prevent libraries from creating and keeping their own copies. To address this issue, we propose a concerted advocacy effort aimed at securing the right for libraries to download and print hard copies of digital materials for which they already hold perpetual access rights.

Current contractual arrangements often limit libraries to viewing or accessing digital publications on specific platforms, without allowing them to create either physical or digital backup copies. This restriction undermines their ability to ensure long-term preservation and access. Arguing for a change in these arrangements may require a collective legal challenge, bringing together libraries, professional associations, and advocacy groups to address and renegotiate these terms but the benefits of success would be transformative.

4.9 Reforming national libraries' approach to digital preservation

Our key general recommendation concerns national libraries, which we find are the only organisations systematically, via legal deposit, currently preserving electronic literature. However, this preservation is often undertaken in tandem with copyright registration, particularly in the United States, rather than as intentional cultural preservation for broader public access.

We recommend advocacy for a change in that approach as the first step. The next would be for national libraries to systematically provide public access to this material as it falls out of commerce and eventually out of copyright, so that national libraries are actively preserving knowledge, providing access and, in so doing, creating a meaningful public domain for e-literature.

National libraries, through legal deposit schemes, are uniquely positioned to undertake the preservation-for-access of electronic literature. However, in many cases, this preservation effort is intertwined with copyright registration processes, particularly in the United States. This approach often prioritises legal and administrative aspects over the broader goal of cultural preservation and public access.

To enhance the effectiveness of national libraries in preserving digital literature, we propose a two-step reform. First, we argue for a shift in the preservation approach from one primarily focused on copyright registration to one centred on intentional cultural preservation. This involves re-evaluating preservation strategies to prioritise the long-term access of digital works as a

primary objective, rather than as a by-product of copyright management.

Second, we recommend that national libraries systematically provide public access to digital materials as they fall out of commerce and eventually out of copyright. This proactive stance would ensure that digital works are not only preserved but also made available to the public in a meaningful way. By taking these steps, national libraries can play a crucial role in creating a robust public domain for e-literature, fulfilling their mandate to preserve knowledge and enhance public access.

5. Recommendations – Specific to Scholarly Monographs

The most obvious and fundamental difference between an e-book and an e-journal is word count, or length. A journal article is short enough to be read with ease on a computer screen, even perhaps a smartphone. While individual chapters of an e-book may similarly be accessed for quick reference, reading an entire monograph, which could require many hours over multiple sittings, typically necessitates dedicated hardware, such as an e-reader, or even a printed edition for comfort and usability.

This distinction may explain why scholarly monographs continue to be printed, despite being simultaneously published in digital formats. As a result, there is a reduced sense of urgency about the long-term digital preservation of these works. Outside of national libraries, e-books are frequently preserved within systems originally designed for e-journals, which may not fully accommodate the different, sometimes unique preservation needs of monographs.

This report highlights that if scholarly monographs shift to being exclusively published in digital formats, as has occurred with many e-journals, current preservation infrastructures beyond national libraries will likely prove inadequate. While the timing and inevitability of this transition are uncertain, the possibility cannot be ignored.

In anticipation of this potential shift, we propose seven specific recommendations to pave the way for the long-term preservation of digital scholarly monographs. These recommendations are designed not only to address the unique challenges of monographs but also to provide a framework that may also be applied to other types of e-books.

5.1 Assessing the prevalence and growth of digital-only scholarly monographs.

We recommend conducting comprehensive research to determine the current share of scholarly monographs published exclusively in digital or digitally enhanced formats, and to track the growth rate of this trend.

It is impossible to predict future trends accurately, especially in areas as complex as this, where societal values, scholarly practices, and even personal preferences are important factors. However, a lot could be learned from a comprehensive statistical analysis of current patterns, coupled with market research to gauge the attitudes of scholars to alternative format options, across different disciplines and regions.

This combined approach will provide a clearer understanding of ongoing shifts in publishing practices, allowing for a more informed response to the evolving landscape of scholarly monographs, particularly in relation to long-term preservation strategies. The research should also consider the appetite for emerging formats, including those that incorporate digital enhancements such as the inclusion of multimedia or interactive elements.

5.2 Identifying and evaluating already lost born-digital scholarly monographs

We recommend a very directed project to identify if any such born-digital scholarly monographs have already been lost and if so, how many to date.

Building on the detailed analysis by Laakso, Matthias and Jahn that identified lost Open Access e-journals and anecdotal evidence gathered during research, we propose addressing the issue of 'already lost' digital scholarly monographs through a dedicated project aimed at identifying

and quantifying any such missing works. This research should investigate which specific born-digital monographs have been lost to date and understand the reasons for their disappearance. The scope should include a comprehensive survey of available archive records, digital repositories, and bibliographic databases to identify any missing titles. Understanding the scale and causes of these losses will provide valuable insights into the future challenges to be faced by digital preservation efforts and inform strategies to reduce further losses.

5.3 Rethinking approaches to the preservation of digital scholarly monographs

We recommend a comprehensive re-evaluation of preservation strategies for digital scholarly monographs, including the feasibility, implications, and costs of developing capabilities to systematically acquire, print, and secure these works.

We propose a forward-thinking review of potential 'blue-sky' strategies for the preservation of digital scholarly monographs. This exploration should consider innovative approaches such as systematically obtaining, printing, and safeguarding these works for long-term preservation. Such a re-evaluation should weigh the pros and cons of establishing mechanisms to ensure that digital monographs, which may be at risk due to format obsolescence or technological changes, are preserved in simple formats including, but not limited to, print. This approach may act only as a stopgap to address the challenges of preserving born-digital academic literature to ensure its availability for future scholars, but due to the relatively small number of works currently published as digital-only, starting now would buy time until a more secure approach is developed or discovered.

5.4 Integrating scholarly monographs into dark archives

We recommend research into how to develop and set aside a capacity within the successful dark archives e-journal practices so that scholarly monographs start to be systematically preserved within these systems, in anticipation of a move to digital-only formats.

To further safeguard digital-only scholarly monographs, we suggest adapting the successful practices of dark archives currently used mainly for e-journals to include e-monographs. The goal is to establish a systematic approach for preserving these works, anticipating a future where digital formats become the norm. This integration will help ensure that all such scholarly publications are preserved consistently and effectively, retaining their accessibility and utility over coming decades.

5.5 Researching orphaned born-digital e-books and archival solutions

We recommend further research into the actual number of orphaned, born digital e-books and consideration of how to create an archive facility, alongside an accessible library, into which they can pass while simultaneously seeking the permission of the rights holder.

To address the issue of abandoned, or so called orphaned, e-books, we recommend conducting research to determine their number and assess the potential for specific solution for this category of publication. The research should focus on identifying which digital books have become abandoned or orphaned – where rights holders are unknown or unreachable – and evaluating the extent of this as an issue.

The findings should inform the feasibility for an archival system specifically designed to locate,

manage, preserve and provide access to these works. This approach would serve as both an archive and an accessible library, ensuring that orphaned e-books are systematically preserved while simultaneous efforts are made to locate and obtain permissions from rights holders. This dual approach will help safeguard valuable digital literature and enhance access by appeasing both rights bodies and scholarly communities.

5.6 Investigating preservation practices for other academic and trade e-books

We recommend similar research into other academic and trade e-books that are currently published exclusively in digital formats, including self-published works. The research should also assess whether there are established plans to identify and preserve these works as cultural artefacts. (Our understanding is that much of this e-literature is regarded only as lower-quality reading material and therefore not considered as cultural artefacts, unworthy of preservation.)

The area within the trade publishing industry that is most rapidly moving to digital formats is genre fiction. For example, during the COVID lockdowns of 2020-21 there was a significant growth in demand for e-books in the romance genre. In years to come these works might be of great interest to historians studying the social impact of the COVID lockdowns. But there are concerns that these works are not being properly preserved in order to be made available to future scholars.

We propose specific research into the preservation arrangements for other academic and trade books that are published exclusively in digital formats, including at least some of those that are 'self-published'. This research should also determine whether there are already any established frameworks for recognising and preserving any such works as cultural artefacts. Our current understanding suggests that much of this digital e-literature is often undervalued and perceived as ephemeral or lower-quality, which can lead to it being devalued and neglected. This study should aim to ensure that significant literature and cultural contributions that may not seem valuable today are identified and preserved for future generations.

5.7 Creating a centralised data source for e-book preservation status

We recommend investigation into the potential need for, and the feasibility of, a centralised data source on the preservation status of e-books – the equivalent of the Keeper's Registry for e-journals.

We propose an investigation into the creation of a centralised registry that would serve as a comprehensive data source for monitoring the preservation status of e-books. This registry would function similarly to the Keeper's Registry for e-journals, offering a transparent system for monitoring the preservation efforts surrounding digital books, particularly scholarly monographs. By consolidating data on the status of e-book preservation it would provide essential insights into what is being preserved, where, and how.

Such a resource would empower scholars, librarians, and the public to actively engage with the state of digital literature preservation, helping them locate vital works and ensuring these digital resources remain accessible for future generations. The creation of this registry would be a significant step towards safeguarding digital knowledge and fostering a collaborative, accountable approach to preservation across institutions.

6. Conclusion

Despite broadening our field of research somewhat to include other libraries and genres of publication, this inquiry primarily aims to determine whether research libraries could ensure permanent access to scholarly literature published exclusively in digital formats, comparable to or better than their current offerings of physical copies. Our findings have revealed numerous significant obstacles that hinder this specific objective. In particular, the fundamental activities of a research library – purchasing, retaining, and providing unimpeded access – are all compromised by the restrictive, unidirectional licensing terms that they must comply with today.

These terms systematically prevent libraries from acquiring true ownership of digital content, removing their freedom to retain and preserve materials as they have done with physical copies for centuries, which libraries have always purchased outright and owned. The present arrangements for e-literature place libraries at the mercy of publishers and licensors, who can alter or revoke future access – presumably at will.

As a direct result, many research libraries have failed to develop the capacity or operational practices necessary for digital archiving that align with their historic mission. This lack of capacity further undermines the long-term accessibility of digital scholarly works, ultimately jeopardising the future viability of research libraries themselves, as the custodians of academic knowledge.

The shift to Open Access (OA) publishing, while addressing near-term issues of accessibility, affordability, and equity, has paradoxically exacerbated the problem. OA journals, despite meeting their objective to provide free and open access, are particularly vulnerable to failures in coordinated preservation efforts. The decentralised nature of OA publishing often leads to inconsistencies in archiving practices, further complicating efforts to ensure long-term access.

Our investigation did not uncover a single example of a research library or academic institution that systematically acquires, stores, preserves, and plans with the objective of providing permanent free access to a comprehensive corpus of electronic literature, whether published under commercial or non-commercial terms. This lack raises genuine concerns about the future of research and scholarship.

Consequently, we currently see little hope of a planned and structured transition of these publications – produced almost entirely in the 21st century – into a future state where they meet the criteria of a 'meaningful public domain', in which such material would be released and made freely available, either to future scholars or the public in general, and be discovered, accessed, referenced, and utilised.

Although they possess the potential to be the source of a global solution, the national libraries seem to have fallen behind where they should be providing leadership. As electronic literature ages, its vulnerabilities become more pronounced. It is increasingly likely that

e-literature will only be comprehensively preserved within the legal deposit systems of various national libraries – currently an uncoordinated global effort – where access is at risk of remaining highly restricted.

Despite the enormous potential of digital material for accessibility, current arrangements suggest that the e-literature of today will likely be less accessible to future generations than print publications of the past are to us now. To prevent this, and also to reclaim their role of vital contribution to the creation of new knowledge, research libraries must address the challenges posed by restrictive licensing terms and develop robust digital archiving practices to re-discover their mission of preserving scholarly knowledge for future generations. Collaborative efforts,

strategic planning, and enhanced preservation initiatives are essential to ensure that the wealth of 21st-century digital literature remains accessible and beneficial for years to come.

Our purpose here is not to assign blame, but to make the point that while there exists a consensus across publishing, academic, legal, governmental, and library sectors regarding the moral imperative of perpetual preservation for secure future access, no single stakeholder within this ecosystem either has assumed, nor can assume, responsibility for ensuring this outcome. The prevailing business and operational frameworks of all stakeholders are so misaligned that identifying an empowered and adequately resourced candidate from the list that is capable of addressing, financing and meeting this challenge remains a distant prospect.

Indeed, in our opinion, the ultimate power to act resides with those who are focused on stabilising their short-term operational models and revenue streams, leaving them with no remaining bandwidth to contemplate or invest in mitigating the long-term repercussions that arise from prioritising their own objectives over the enduring needs and rights of future generations.

Final word

Marie-Louise Ayres of the National Library of Australia has thought about the issues covered in this a great deal, and she has acted decisively and, in our opinion, effectively, but she admitted: "There's no doubt that across the world digital collections will be lost". The final word we give to Clifford Lynch of the CNI, who told us. "The problem will be persistence. For almost all published works there is a long 'dead period' and, for born-digital publications, the likelihood of survival throughout such a period of neglect will be significantly diminished. The fear is the digital public domain will be less rich than print".

This is not, however, a cause for despair. While some e-literature will undoubtedly be lost, and some already has been, this technology is still in its relative infancy. We believe there is still time to act – but the window is closing. Now is the critical juncture to create systems that not only prevent e-literature from disappearing, but that actively preserve it and make it accessible in ways never before possible with print. Through these efforts, we can unleash and secure for future generations the full democratising potential of electronic publishing, thereby setting the scene for a rich and enduring digital public domain.

Appendices

APPENDIX A - INTERVIEWEES

Organisation type	No. of Interviewees
Research and Public Libraries	13
National Libraries	9
Archives	6
Publishers	3
Lawyers	2
Others	10
Total	**43**

Research and Public libraries

Helen Cargill, Head of Open Research, Kings College London.

Rob Cartolano, Associate Vice President for Digital Programs and Technology Services, Columbia University Libraries.

Mikel Christofferson, Director of European Bureau of Library, Information and Documentation Associations (EBLIDA) – Denmark.

Cathal Macauley, University Librarian, Maynooth University.

Pilar Martinez, CEO, Edmonton Public Library.

Lauren Moore, Assistant Commissioner for Libraries and State Librarian, New York State.

Andrew K. Pace, Executive Director, Association of Research Libraries.

Mila Pollock, Executive Director Library and Archives, Cold Spring Harbor Laboratory.

Susan Reilly, Director of Irish Research e-Library, Maynooth University.

Kathleen Riegelhaupt, Director of eReading, New York Public Library.

Judy Ruttenberg, Senior Director of Scholarship and Policy, Association of Research Libraries.

Christoph Schmidt-Supprian, Head of Content Management, Trinity College Dublin Library.

Vicky Varga, Executive Director of Collections, Marketing and Technology, Edmonton Public Library

National libraries

Marie-Louise Ayres, Director General, National Library of Australia.

Tim Kong, Director Digital Experience, National Library of New Zealand.

Sally Halper, Director of Research Services, British Library.

Eoin McCarney, Head of Published Collections, National Library of Ireland.

Johannes Neuer, Director, German National Library, Leipzig.

Helen Shenton, Librarian and College Archivist, Trinity College Dublin.

Alison Stevenson, Director of Collections, Access and Research, National Library of Scotland.

Leslie Weir, Librarian and Archivist of Canada.

Kate Zwaard, Associate Librarian for Discovery & Preservation Services, Library of Congress.

Archives

Ben Bunnell, former Manager of Library Partnerships, Google.

William Kilbride, Executive Director at Digital Preservation Coalition.

Paul Wheatley, Head of Research and Practice at Digital Preservation Coalition.

Cathy Williams, Head of Archives, Kings College London.

Alicia Wise, Executive Director, CLOCKSS Archive.

Kate Wittenberg, Managing Director. Portico.

Publishers

Duncan Campbell, Executive Director of Sales & Account Management at Copyright Clearance Center.

John Sargent, former CEO, Macmillan Publishers USA.

Brian McLendon, former SVP, Hachette Book Group.

Lawyers

Martin Bradley, Knowledge Rights 21.

Eoin O'Dell, Associate Professor of Law, Trinity College Dublin.

Others

Femi Adelakun, Director of Research and Data, Urban Libraries Council

Martin Eve, Professor of Literature, Technology and Publishing, University of London's Birkbeck College.

Stuart Hamilton, Head of Libraries Development, Local Government Management Agency (LGMA) in Ireland.

Alan Inouye, Director of Public Policy, American Library Association.

Rice Majors, Associate University Librarian for Scholarly Resources, UC Davis Library.

Ross Mounce, Director of Open Access Programmes, Arcadia Fund.

Brooks Rainwater, President and CEO, Urban Libraries Council.

Steve Rosato, Manager of Professional Business Unit, OverDrive.

Ann Rossiter, Executive Director, Sconul.

George Wright, Library and metadata consultant

APPENDIX B - LICENCES

(For the purposes of this report, all licences have been redacted and anonymised.)

Scholarly and academic e-Journals subscription

Below are redacted excerpts from an e-journal licence that reveal the clauses concerning the subscribed materials.

Example A

SECTION 1. SUBSCRIPTION.

1.1   *Subscribed Products.*
██████ hereby grants to the Subscriber the non-exclusive, non-transferable right to access and use the products and services identified in Schedule 1 ("Subscribed Products") and provide the Subscribed Products to its Authorized Users (as defined herein) which rights are worldwide with respect to off-campus remote access by Authorized Users and perpetual as described in detail on Schedule 1 regarding access to formerly subscribed titles and archival print copies (which rights survive the termination of this Agreement), all subject to the terms and conditions of this Agreement.

1.2   *Authorized Users/Sites.*
Authorized Users are the full-time and part-time students, faculty, staff and researchers of the Subscriber and individuals who are independent contractors or are employed by independent contractors of the Subscriber affiliated with the Subscriber's locations listed on Schedule 2 (the "Sites") and individuals using computer terminals within facilities at the Sites permitted by the Subscriber to access the Subscribed Products for purposes of personal research, education or other non-corporate use.

1.3   *Authorized Uses.*
Each Authorized User may:

- access, search. browse and view the Subscribed Products;

- print, download and store a reasonable portion of individual items from the Subscribed Products for the exclusive use of such Authorized User:

- incorporate links to the Subscribed Products on the Subscriber's intranet and internet websites and in electronic coursepacks, reserves and course management systems and instructor websites, provided that the appearance of such links and/or statements accompanying such links will be changed as reasonably requested by Elsevier;

- provide print or electronic copies of individual items from the Subscribed Products to other Authorized Users and to third-party colleagues for their scholarly or research use; and

Upon notification by the Subscriber that it wishes to initiate the service as described below (the "Transactional Service') access, search, browse, view, print, make electronic copies and store for the exclusive use of such Authorized User or, if the Authorized User is a librarian/information specialist, for the exclusive use of another Authorized User certain journal articles and book chapters from the ███████████ online service that are not subscribed to as part of the Subscribed Products, with each twenty-four (24) hour access period for a selected article or chapter, a "Transaction.

Example B

1.4   *Restrictions on Use of Subscribed Products.*
Except as expressly stated in this Agreement or otherwise permitted in writing by ▮▮▮▮, the Subscriber and its Authorized Users may not:

- abridge, modify, translate or create any derivative work based on the Subscribed Products, except to the extent necessary to make them perceptible on a computer screen to Authorized Users;

- remove, obscure or modify in any way any copyright notices, other notices or disclaimers as they appear in the Subscribed Products;

- use any robots, spiders, crawlers or other automated downloading programs, algorithms or devices to continuously and automatically search, scrape, extract, deep link, index or disrupt the working of the Subscribed Products without express written permission of ▮▮▮▮, which shall not be unreasonably withheld provided Authorized Users are engaging in activities for academic research or other educational purposes which do not affect overall platform stability, availability and performance; or

- substantially or systematically reproduce, retain or redistribute the Subscribed Products.

2

Example C

Authorized Users who are individuals who are independent contractors or are employed by independent contractors may use the Subscribed Products only for the purposes of the contracted research work for the Subscriber.

1.5   *Intellectual Property Ownership.*
The Subscriber acknowledges that all right, title and interest in and to the Subscribed Products remain with ▮▮▮▮ and its suppliers, except as expressly set forth in this Agreement, and that the unauthorized redistribution of the Subscribed Products could materially harm ▮▮▮▮ and its suppliers. Notwithstanding anything to the contrary contained in this Agreement, more extensive usage terms might be permitted for open access content in the Subscribed Products as identified in the individual journal article as stated in the applicable user (e.g. CC) license.

Scholarly and academic e-Journals perpetual access

Our research reveals that academic and research libraries usually have perpetual rights to scholarly e-journals published within a subscription period. Below are two examples of such clauses - both are from contracts between academic publishers and a university library. Our research suggests they are representative.

Example D

<div>

XII. PERPETUAL LICENSE

Notwithstanding anything else in this Agreement, Licensor hereby grants to Licensee a nonexclusive, royalty-free, perpetual license to use any Licensed Materials that were accessible during the term of this Agreement. Such use shall be in accordance with the provisions of this Agreement, which provisions shall survive any termination of this Agreement. The means by which Licensee shall have access to such Licensed Materials shall be in a manner and form substantially equivalent to the means by which access is provided under this Agreement.

</div>

Example E

<div>

5.9   On expiry of the Institution's Subscription for the Journal(s), the Institution and each Authorized User shall be entitled to continue to exercise at no charge the non-exclusive rights granted hereunder (subject to the terms and conditions of this Site License) but only in respect of Material published and paid for by the Institution during the period of the Licensee's Subscription.

</div>

Scholarly and academic e-Journals use of third-party dark archives

Scholarly and academic publishers often commit to using specific third-party archives in e-journal licences. Below are two examples of such clauses. Again, both are from contracts between academic publishers and a university library.

Example F

Archival/Backup Copy. The Licensor has made provision for an archive for the long term preservation of the Licensed Materials for the provision of access by the Licensee to the Licensed Materials (the "Archive"), to be provided by a third party provider, which is currently Portico, the electronic-archiving initiative of ITHAKA. Access to the Archive is contingent on the occurrence of specified trigger events which would prevent the Licensor from providing access to the Licensed Materials. In order to benefit from the service offered by Portico, the Licensee must become a contracted customer of Portico. The Licensor reserves the right, at its sole discretion, to change the third party provider of its archive at any time. The Licensee must ensure it and any Authorized User continues to comply at all times with the permitted use and restrictions set out herein. For the avoidance of doubt, the Licensor shall have no obligation to provide access to the Archive where this License has been terminated as a result of the Licensee's breach. For the avoidance of doubt, access to the Archive is not provided in relation to any materials licensed to the Licensee by the Licensor on a subscription basis

Example G

4.5 Licensor cooperates with, and participates in a number of   digital preservation services provided by e.g. CLOCKS, LOCKSS and Portico for the preservation of certain online products of Licensor. In the case of a triggering event set forth in Licensor's agreements with the digital preservation services, Licensee shall be entitled to access the Content pursuant to such agreements.

Licensor and Licensee agree that the terms of Licensor's

Scholarly and academic e-books perpetual access

Although many university libraries have e-book contracts that replicate e-journal contracts and offer perpetual access or Post Cancellation Access (PCA) to works published during a subscription period, not all e-books are licenced this way.

We have also found contracts in which all access ends when the licencing term to the package(s) ends. Below are two clauses – provided by a university library – that illustrate this.

Example H

8. Term & Termination

Client's use of the Products is subject to the Term Start and Term End Date listed within each applicable Order Form. Thereafter, the

Products will automatically renew for up to three (3) successive one (1) year terms, subject to the appropriate adjustments to Fees,

unless either party gives the other written notice of termination at least ninety (90) days prior to expiration of the current term. Either

party may terminate the Agreement if the other commits a material breach and fails to cure within thirty (30) days after receipt of

written notice. Upon termination, all rights granted herein will immediately cease, and Client will return or destroy all ▮▮▮▮▮ Products.

Example J

10.3 **On termination or expiry of a Subscription the Licensee shall:**

10.3.1 immediately delete any downloaded copies of the Publications or part thereof

made by the Licensee and if requested by ▮▮▮▮▮ shall certify such deletion; and

10.3.2 use all reasonable endeavours to ensure that all Authorised Users delete any

downloaded copies of the Publications or part thereof made by them.

Scholarly and academic e-books use of third-party dark archives

We did not find it typical that contracts for e-book licences commit the publisher to the use of third-party archives. In some contracts, this is because the licence does not offer PCA, in which case there is no incentive to archive the material. But even in licences that do offer PCA the contract often transfers the burden (and the cost) of archiving e-books onto the library, as can be seen in examples K and L below.

Example K

> ███████ **acknowledges that Licensee may engage the services of third-party trusted archives and/or participate in collaborative archiving endeavors to exercise Licensee's rights under section XII, 'Perpetual License', of this Agreement. Licensee agrees to cooperate with such archiving entities and/or initiatives as reasonably necessary to make the Licensed Materials available for archiving purposes. Licensee may perpetually use the third-party trusted system to access or store the Licensed Materials, so long as Licensee's use is otherwise consistent with this Agreement. Licensor further acknowledges and agrees that, in using the third-party archival system, Licensed Materials may be made available to other system participants who indicate a right to those Licensed Materials.**

Example L

> (1) Optional Offline Copy Policy. ███████ acknowledges that Licensee may engage the services of third-party trusted archives and/or participate in collaborative archiving endeavors to exercise Licensee's rights under Section 3(a), 'License Grant', of this Agreement. ███████ agrees to cooperate with such archiving entities and/or initiatives as reasonably necessary to make the Licensed Products available for archiving purposes. Licensee may perpetually use the third-party trusted system to access or
>
> 2
>
> store the Licensed Products, so long as Licensee's use is otherwise consistent with this Agreement. ███████ further acknowledges and agrees that, in using the third-party archival system, Licensed Products may be made available to Authorized Users only.

# APPENDIX C – SURVEY RESULTS

Librarian questionnaire

To broaden our analysis and understanding of the current challenges and concerns surrounding the preservation of knowledge in digital formats, we engaged with librarians across the United States, the United Kingdom, and the European Union. These professionals were invited to participate in a comprehensive questionnaire designed to capture their perspectives on existing practices within the sector and to assess the extent of their concerns regarding the long-term accessibility of digital materials.

A total of 97 responses were gathered from a diverse array of librarians representing research and academic institutions, public and circulating libraries, school libraries, national libraries and archives. This varied respondent pool provided valuable insights into the differing approaches and challenges faced by libraries of all types and sizes.

The findings from this survey are compiled and presented in full in the following section, offering an in-depth view of the current landscape and the collective outlook on the preservation of digital knowledge across the library community.

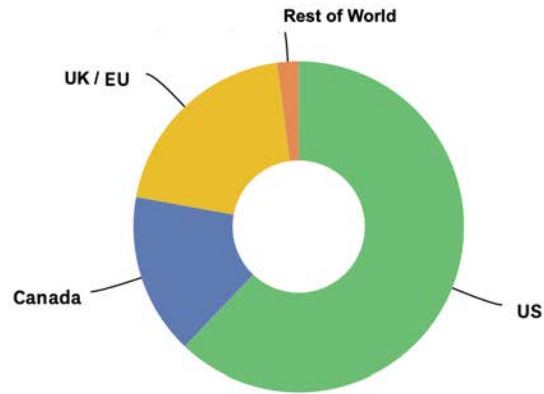## What kind of organisation do you work in? (You may need to tick more than one box.)

Answered: 95   Skipped: 2
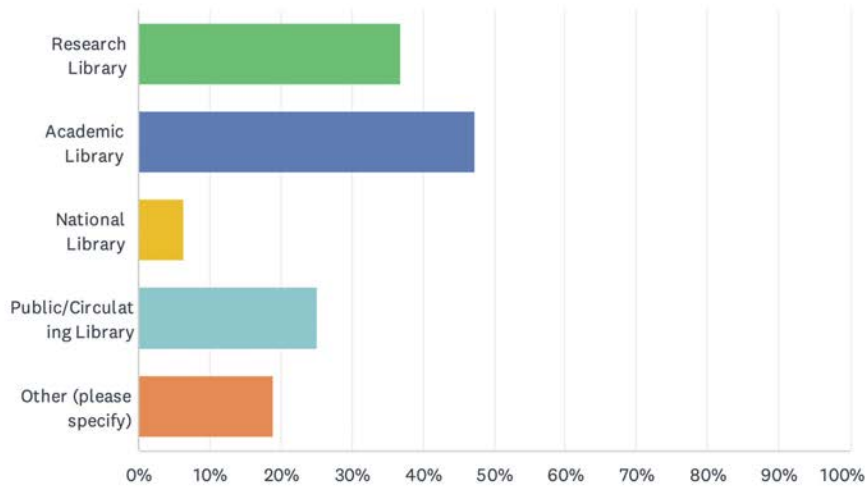
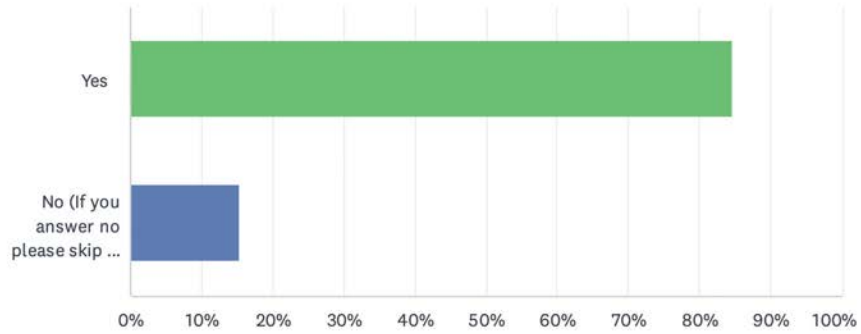| ANSWER CHOICES	RESPONSES	
Research Library	36.84%	35
Academic Library	47.37%	45
National Library	6.32%	6
Public/Circulating Library	25.26%	24
Other (please specify)	18.95%	18
Total Respondents: 95		

In which of these regions is your organisation based?

Answered: 95   Skipped: 2
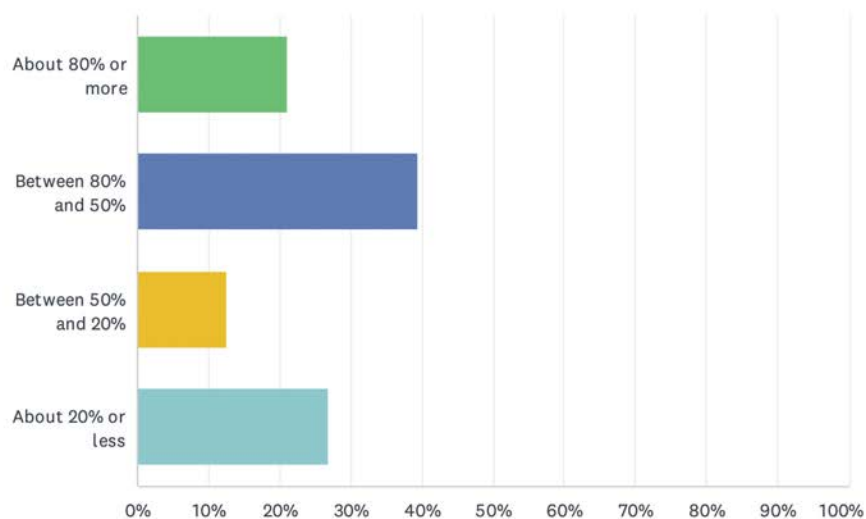
Does your library subscribe to any e-journals?

Answered: 85   Skipped: 12

ANSWER CHOICES	RESPONSES	
Yes	84.71%	72
No (If you answer no please skip to the next page and continue from question 11.)	15.29%	13
TOTAL		85

## Approximately what share of your library's annual resources budget is spent on e-journals?

Answered: 71   Skipped: 26

ANSWER CHOICES	RESPONSES	
About 80% or more	21.13%	15
Between 80% and 50%	39.44%	28
Between 50% and 20%	12.68%	9
About 20% or less	26.76%	19
TOTAL		71

Do your library's licences with publishers ensure perpetual access to e-journals that were published within the subscription period?
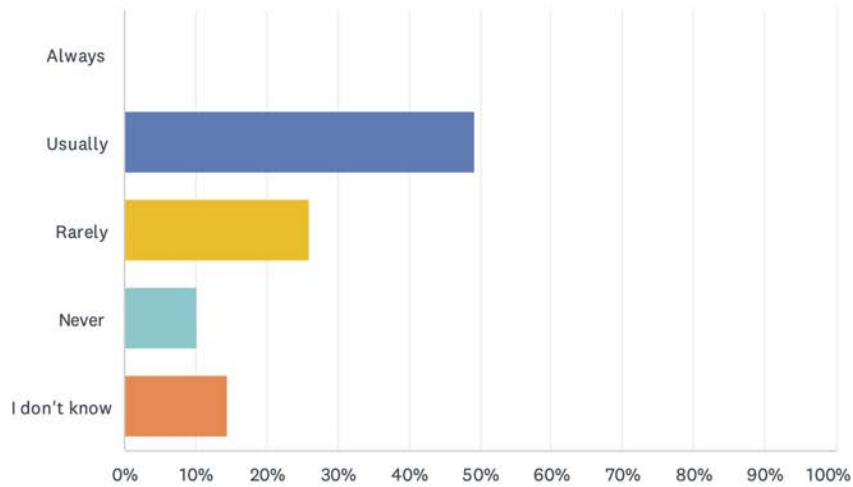
Answered: 70   Skipped: 27
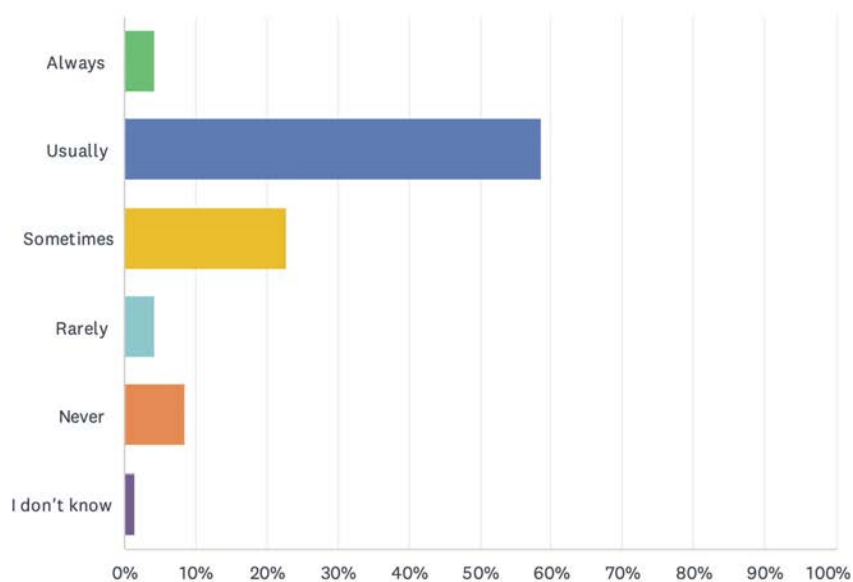
ANSWER CHOICES	RESPONSES	
Always	4.29%	3
Usually	58.57%	41
Sometimes	22.86%	16
Rarely	4.29%	3
Never	8.57%	6
I don't know	1.43%	1
TOTAL		70

Do your library's licences for e-journals include a commitment that the publisher will use third party archives to preserve this material?
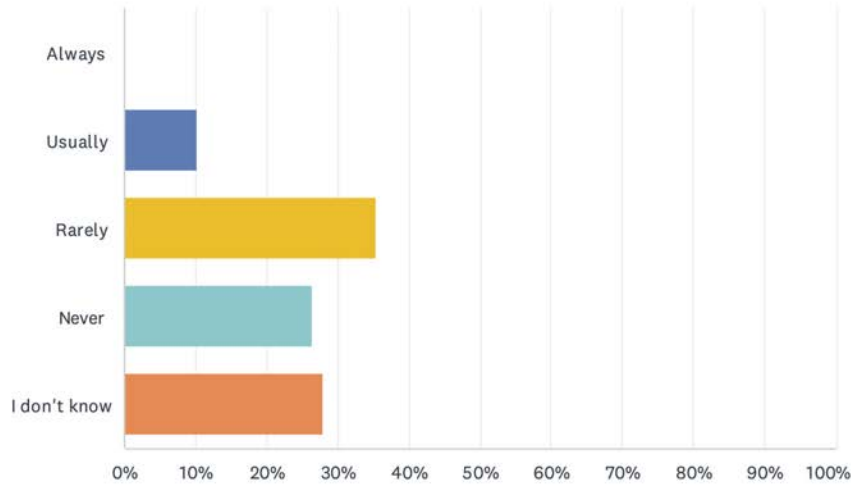
Answered: 69   Skipped: 28

ANSWER CHOICES	RESPONSES	
Always	0.00%	0
Usually	49.28%	34
Rarely	26.09%	18
Never	10.14%	7
I don't know	14.49%	10
TOTAL		69

Do your library's licences with publishers for e-journals mention if and how access will be possible once this material enters the public domain?

Answered: 68   Skipped: 29

ANSWER CHOICES	RESPONSES	
Always	0.00%	0
Usually	10.29%	7
Rarely	35.29%	24
Never	26.47%	18
I don't know	27.94%	19
TOTAL		68

Any other issues or recommendations concerning e-journals that you wish to bring attention to.

Answered: 21   Skipped: 76

#	RESPONSES	DATE
1	I'm only involved in licensing for niche music publications and my responses don't reflect the broader trends beyond my library at my University.	7/31/2024 6:13 PM
2	Information about what happens if/when the ejournal becomes open access has not been mentioned in any of our recent agreements, and we are not aware that it has ever appeared in our older agreements.	7/16/2024 1:33 PM
3	We are unsure that we understand question #8. In what way should a license agreement between a library and a publisher address the issue of access once the content enters the public domain?	7/10/2024 9:38 PM
4	A lot of time is spent tracking publishers and title changes, license updates that go along with that, and finding exact dates of perpetual access versus non-perpetual.	7/9/2024 12:46 PM
5	Journals published in languages other than English or from non-Western countries pose unique challenges. These may follow different standards and expectations and translation may be challenging.	7/8/2024 5:54 PM
6	NLM has conducted two studies in the past of license terms related to post cancellation access and publisher participation in 3rd party repositories. For question 7 above, we responded usually but I do not have current data on that and I don't know what percentage is being used to define "usually". A more applicable response in this case would be "sometimes."	7/8/2024 4:28 PM
7	RE question #7: Licenses that mention a third party archive usually only include it as an option for us, rather than a publisher making a commitment to using a third party.	7/5/2024 12:20 AM
8	Some of our licenses that contain perpetual access for materials have clauses that are variable - for instance they'll say that the publisher will endeavour to provide access on their site, but that if they are unable to/if that's not feasible, that they'll provide access to us directly for us to host. This happens very rarely - but we are largely not set up to host and maintain access this way. We do have server space for this kind of content, but we couldn't handle large amounts of content.	7/4/2024 4:16 PM
9	Most of what we license will not enter the public domain for decades.	7/1/2024 11:56 PM
10	For these and the ebooks categories of questions about licenses, in order to gauge a response, I assumed we were talking about the # of licenses, not the % of the budget that they represent. In other words, a large publisher license like Elsevier and Springer Nature are very large in terms of the % of our budget but represent one license each, whereas we have many, many smaller licenses that represent a much smaller % of the budget. I counted # of licenses.	7/1/2024 6:37 PM
11	Some publishers have proposed a moving wall for access to subscribed materials (previous 10 years, for example). Would prefer perpetual access to all years that we have paid for.	7/1/2024 4:18 PM
12	Transfers often result in us using our Portico access because the old publisher has removed the material and the new publisher won't grant us access unless we start a new subscription with them. In this case, preservation of access is also a consideration on top of preservation of the information itself.	6/26/2024 9:57 PM
13	The concern around what happens to content where the library is not allowed to archive the materials, but the publisher hasn't made a commitment to preserving it themselves. There is also the long term concern about digital storage for this content and where the professional conversation is around "shared electronic" agreements similar to shared print agreements.	6/26/2024 8:32 PM
14	The issue of the public domain does not come up in our license agreements. If it did, I would not see that it would have a significant impact in our decision to preserve the journal content -	6/26/2024 7:47 PM

	preservation of public domain content is equally important to that of licensed, restricted content.	
15	accessibility, translation, data tracking	6/26/2024 5:57 PM
16	many different types of e-journals represented in different ways, e.g. one-time purchases usually are in perpetuity whereas subscriptions in most cases aren't. Something to be considered.	6/24/2024 5:51 PM
17	The terms and conditions are more and more mentionned in the license agreement as part of the signed agreement, but the terms and conditions are published online and can change anytime.	5/30/2024 8:51 AM
18	A print copy is easy to "move around", it can have several readers. You also pay for it once and can keep it forever. An e-journal is more expensive since you have to pay a license for each reader and you must pay for it over and over again each year to keep access to the archive. E-journals are expensive.	5/28/2024 9:53 AM
19	Issue: The occasional report that some publishers attempt to bypass the Library at renewal time and attempt to licence content directly with the institution's Legal Affairs Office. By doing so, they attempt to remove some existing conditions in order to tie the institution to more restrictive clauses than originally negotiated by the Library.	5/28/2024 8:24 AM
20	A large part of the e-journals we subscribe to are made available through aggregated databases and when a subscription to a database ends, access to the included materials ends	5/23/2024 12:28 PM
21	We subscribed a licence with Media Library Online - https://www.medialibrary.it/home/index.aspx	5/23/2024 10:06 AM

Does your library license any e-books?
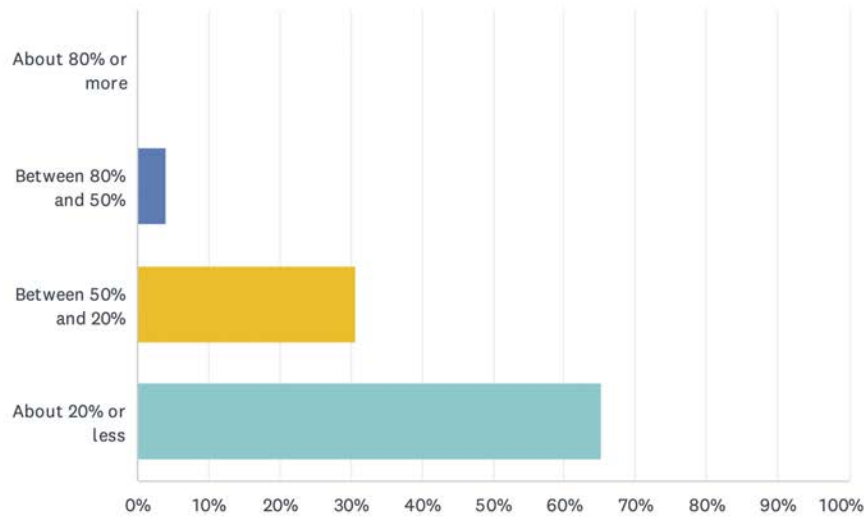
Answered: 80   Skipped: 17
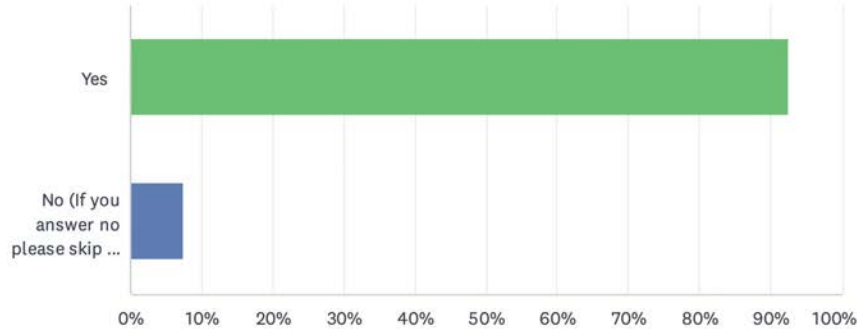
ANSWER CHOICES	RESPONSES	
Yes	92.50%	74
No (If you answer no please skip to the next page and continue from question 17.)	7.50%	6
TOTAL		80

Approximately what share of your library's annual resources budget is spent on e-books?
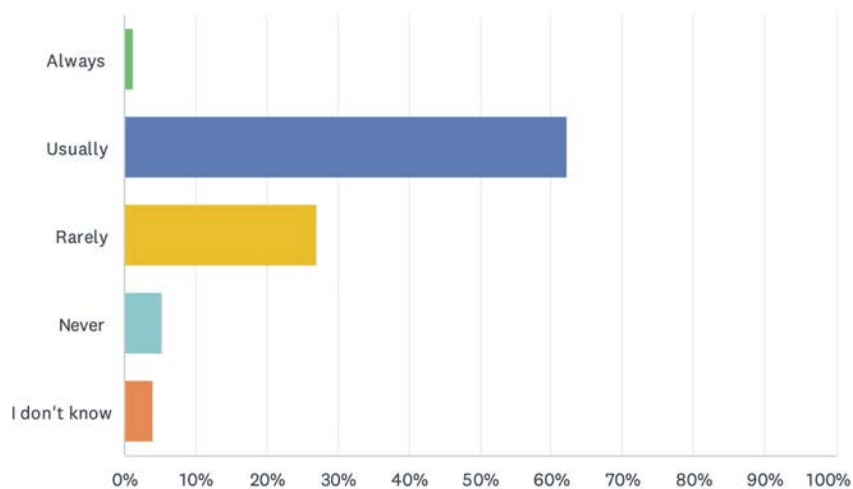
Answered: 75   Skipped: 22

ANSWER CHOICES	RESPONSES	
About 80% or more	0.00%	0
Between 80% and 50%	4.00%	3
Between 50% and 20%	30.67%	23
About 20% or less	65.33%	49
TOTAL		75

Do your library's licences with publishers include perpetual access to e-books?
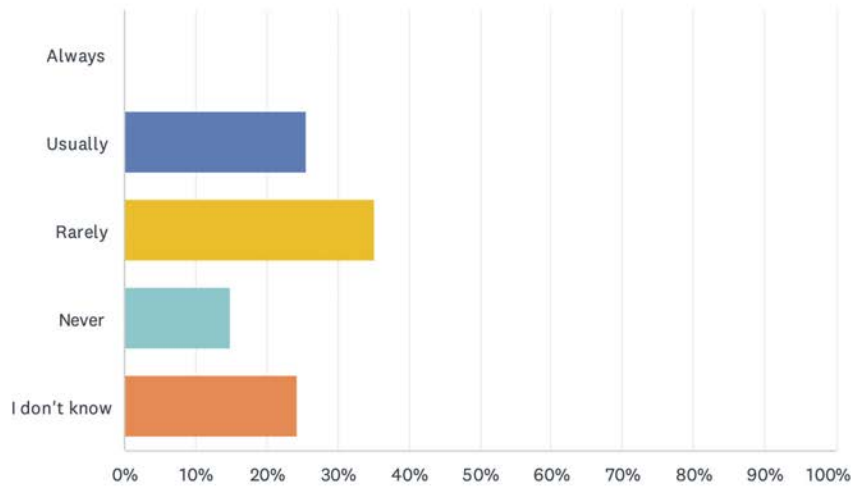
Answered: 74   Skipped: 23

ANSWER CHOICES	RESPONSES	
Always	1.35%	1
Usually	62.16%	46
Rarely	27.03%	20
Never	5.41%	4
I don't know	4.05%	3
TOTAL		74

Do your library's licences for e-books include a commitment that the publisher will use third party archives preserve this material?
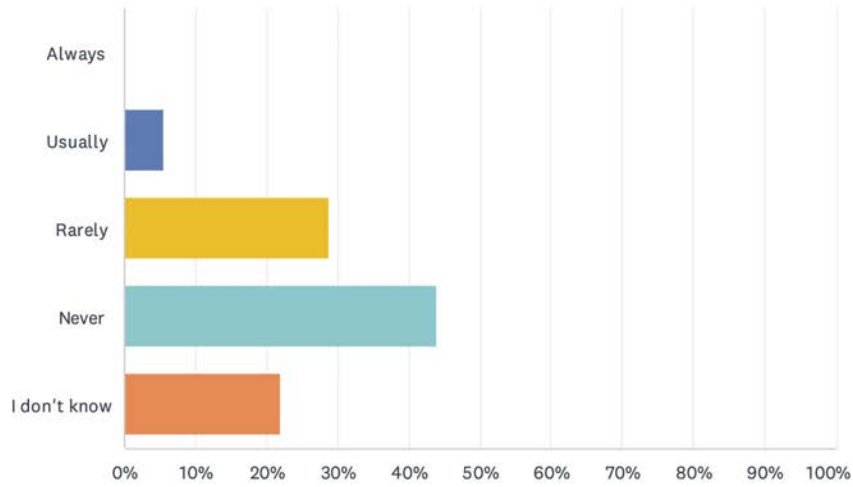
Answered: 74   Skipped: 23

ANSWER CHOICES	RESPONSES	
Always	0.00%	0
Usually	25.68%	19
Rarely	35.14%	26
Never	14.86%	11
I don't know	24.32%	18
TOTAL		74

Do your library's licences with publishers for e-books mention if and how access will be possible once this material enters the public domain?

Answered: 73   Skipped: 24

ANSWER CHOICES	RESPONSES	
Always	0.00%	0
Usually	5.48%	4
Rarely	28.77%	21
Never	43.84%	32
I don't know	21.92%	16
TOTAL		73

Any other issues or recommendations concerning e-books that you wish to bring attention to.

Answered: 19   Skipped: 78

#	RESPONSES	DATE
1	There is a diminishing amount of content available with permanent licensing rights. Most now come with metered access.	7/31/2024 6:36 PM
2	On Question # 12, some medium to smaller publishers offer perpetual access and, rarely, even ownership, but most of the the Big 5 only do for digital audio. Even then, it is questionable if "perpetual access" is likely. Files would regularly have to be saved and updated to new versions to even have a hope of access in many years.	7/31/2024 5:43 PM
3	We casnnot recall seeing any mention of what would happen to our ebook access if the material enteres the public domian in any of our ebook licenses.	7/16/2024 1:35 PM
4	We are unsure that we understand Question #14. See comment in Question #9	7/10/2024 9:39 PM
5	We select perpetual access e-book licenses whenever available, but many publishers do not offer this, only subscription access. Perpetual access licenses are more likely to include information on long term preservation, but subscription content often does not include these guarantees.	7/8/2024 5:58 PM
6	We buy a large number of ebooks through aggregators (ProQuest eBook Central, EBSCOhost). We have perpetual access to those books but those contracts do not provide for third-party archiving of content. Similar to journals, some contracts stipulate they'll providing ongoing access via FTP or hard drive transfer to us - like journals, we have some capacity to host this content, but we couldn't handle much more than ad hoc, and we don't have capacity for any DRM/simultaneous user limits.	7/4/2024 4:17 PM
7	I can't think of a license that has addressed public domain. I may be forgetting one but it is definitely not something on our radar during negotiations.	7/1/2024 6:39 PM
8	No.	7/1/2024 4:18 PM
9	The issue of the public domain does not come up in our license agreements. If it did, I would not see that it would have a significant impact in our decision to preserve the e-book content - preservation of public domain content is equally important to that of licensed, restricted content.	6/26/2024 7:48 PM
10	data tracking, cost, accessibility, ownership, license models	6/26/2024 5:58 PM
11	These are very real challenges regarding preservation. I think there's been more thought about the issue on the e-journal side than the e-book side.	6/24/2024 5:57 PM
12	The survey may receive better questions if different formats for ebooks are included. Eaudio and ebooks are often spoken of in the same terms.	6/5/2024 3:05 PM
13	We are part of a national tender for e-books - the e-book platform is contract managed by the Local Government Management Agency and they set the requirements. I don't know if public domain and archiving has even been considered in this contract.	5/29/2024 10:10 AM
14	Most of them are packages within larger providers, e.g. EBSCO, OVID and it's not clear from licenses,	5/28/2024 12:10 PM
15	Issue: Some publishers charging ongoing hosting fees for e-books you have purchased outright (supposedly perpetual access but only if you continue to pay the hosting fee).	5/28/2024 8:25 AM
16	We mostly subscribe to encyclopedias and similar works, and in most cases access is terminated when a license expires - much like databases	5/23/2024 12:37 PM
17	We subscribed a licence with Media Library Online - https://www.medialibrary.it/home/index.aspx	5/23/2024 10:06 AM

18	lack of interlibrary loan	5/21/2024 2:00 PM
19	We purchase ebooks through Overdrive and make them available to users through the Sora platform. If we ever end our relationship with Overdrive, we will lose access to all of our ebooks, even the ones we imported from other vendors or publishers. We knew this going in, but there aren't any other options available that offer such comprehensive collections, easy-to-use administrative and purchasing tools, AND an intuitive interface for patrons. Our system, and certainly not our small school libraries (often with a staff of 1), cannot manage the organizational and hosting requirements of a large ebook collection without vendors like Overdrive, giving them a great deal of power in the licensing relationship, since NOT offering ebooks is not an option these days.	5/20/2024 8:52 PM

Do your library's licences with the publishers of e-journals and e-books allow the library to preserve this material itself?

Answered: 76   Skipped: 21
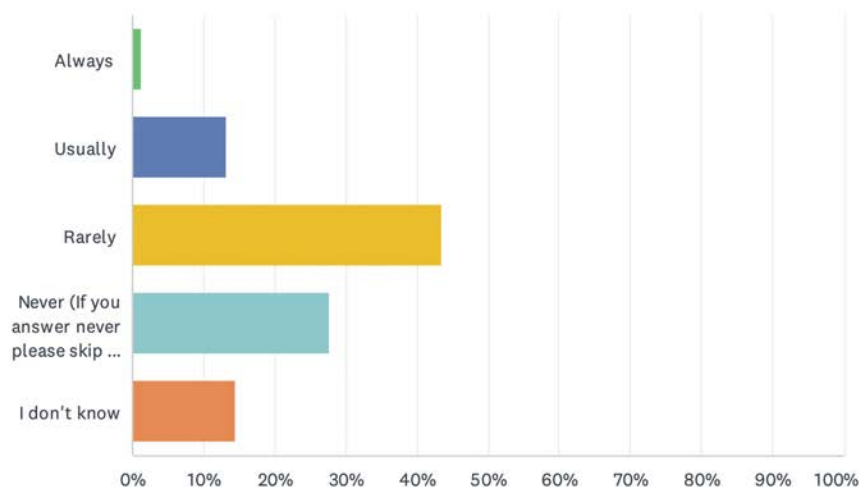
ANSWER CHOICES	RESPONSES	
Always	1.32%	1
Usually	13.16%	10
Rarely	43.42%	33
Never (If you answer never please skip to the next page and continue from question 22.)	27.63%	21
I don't know	14.47%	11
TOTAL		76

If your library preserves any e-publications, how is this done? (You may tick more than box.)

Answered: 36   Skipped: 61
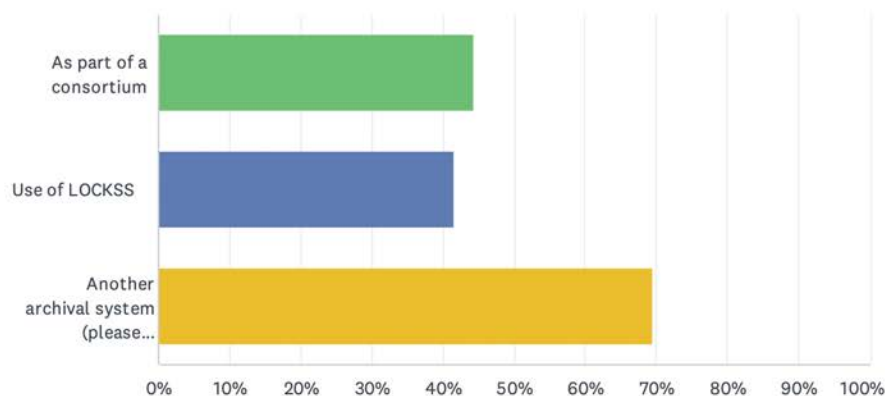
ANSWER CHOICES	RESPONSES	
As part of a consortium	44.44%	16
Use of LOCKSS	41.67%	15
Another archival system (please specify)	69.44%	25
Total Respondents: 36		

| #	ANOTHER ARCHIVAL SYSTEM (PLEASE SPECIFY)	DATE
1	Portico, CLOCKSS	8/6/2024 7:09 PM
2	DSpace	7/31/2024 7:43 PM
3	Internal	7/31/2024 7:04 PM
4	Local archival preservation system	7/31/2024 6:16 PM
5	Portico	7/12/2024 5:59 PM
6	Portico	7/12/2024 12:47 AM
7	local preservation system	7/10/2024 9:39 PM
8	CLOCKSS, Portico	7/9/2024 11:25 PM
9	Portico, Janeway, Exploro	7/9/2024 12:47 PM
10	CLOCKSS, Portico, Scholars Portal	7/4/2024 6:41 PM
11	We store some material on a server for regular access but I wouldn't call this "preserved" in the traditional sense	7/4/2024 4:17 PM
12	We only preserve content that we publish locally or content that is archival. We use ArchiveSpace among other things	7/2/2024 12:01 AM
13	on our own servers	7/1/2024 6:40 PM

14	Internal server	6/28/2024 2:26 PM
15	Portico	6/27/2024 8:45 PM
16	CLOCKSS	6/27/2024 6:44 PM
17	Portico	6/27/2024 3:03 PM
18	Portico	6/26/2024 8:32 PM
19	Portico, LOCKSS, and CLOCKSS	6/26/2024 7:52 PM
20	local institutional repository; CLOCKSS; Portico	6/25/2024 8:54 PM
21	Locally produced e-publications are preserved through local repository.	6/25/2024 5:56 PM
22	We subscribe to LOCKSS, so I assume we could do it. We also rely on Portico.	6/24/2024 6:04 PM
23	The french national consortium Couperin has their own national archiving platform PANIST : https://www.panist.fr/	6/4/2024 6:00 PM
24	I am not sure. Local Studies publications by our local studies are preserved, but I do not know how.	5/29/2024 10:12 AM
25	CSC – Science Information Technology Center, a non-profit limited company owned by the Finnish government and universities, managed by the Ministry of Education and Culture	5/24/2024 9:22 AM

If your library preserves any e-publications, where is this material stored? (You may tick more than box.)

Answered: 37   Skipped: 60
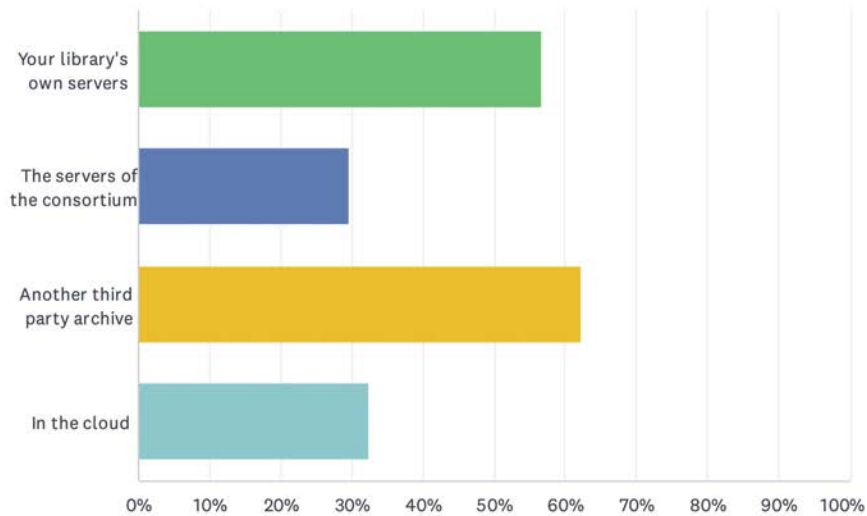
ANSWER CHOICES	RESPONSES	
Your library's own servers	56.76%	21
The servers of the consortium	29.73%	11
Another third party archive	62.16%	23
In the cloud	32.43%	12
Total Respondents: 37		

If your library preserves any e-publications when does this responsibility end?

Answered: 36   Skipped: 61
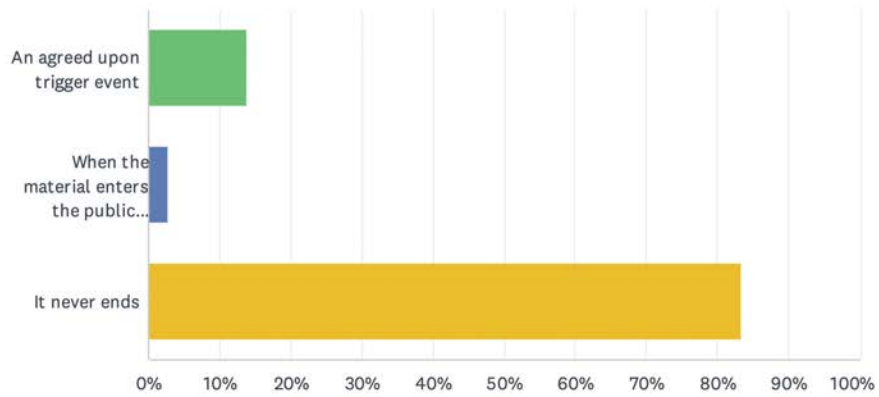
ANSWER CHOICES	RESPONSES	
An agreed upon trigger event	13.89%	5
When the material enters the public domain	2.78%	1
It never ends	83.33%	30
TOTAL		36

Is preservation a key activity your library undertakes for its digital collections?

Answered: 52   Skipped: 45
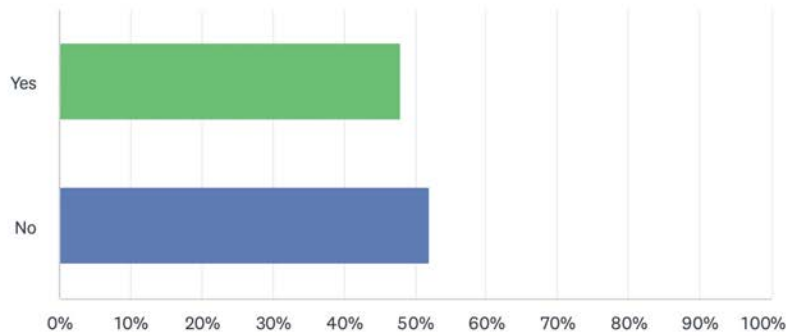
ANSWER CHOICES	RESPONSES	
Yes	48.08%	25
No	51.92%	27
TOTAL		52

Any other issues or recommendations concerning library based perservation.

Answered: 27   Skipped: 70

#	RESPONSES	DATE
1	This activity refers specifically to the escores that the library acquires for perpetual access. I went a little outside the question....ejournals and books will follow these steps if they are acquired for the library to host. Vendor hosted content is generally not preserved.	7/31/2024 6:16 PM
2	Our resources are limited. We struggle to meet current demand. Ensuring preservation is beyond our means. We do limited preservation of titles important for local access.	7/31/2024 5:46 PM
3	The Libraries have focused on preserving locally created or digitized special collections. We have not attempted to preserve licensed content.	7/16/2024 1:38 PM
4	Streaming media is an emerging area with similar preservation needs as e-publications.	7/9/2024 11:26 PM
5	Interested to know what is being preserved across our consortia. What kind of overlap is there?	7/9/2024 12:47 PM
6	Preservation is a key priority for digital special and archival collections, but not for general collections.	7/8/2024 6:03 PM
7	Our licenses generally have terms that allow the library to receive a copy of the content as a method of post cancellation access. However, NLM has not pursued this option with any publishers.	7/8/2024 4:34 PM
8	Increasing use of cloud, 3rd party, and/or trustworthy network arrangements. Less on-prem activity.	7/4/2024 6:41 PM
9	In Ontario, most preservation of electronic content happens through Scholars Portal a service offered through the Ontario Council of University Libraries (OCUL). As a result the majority of our licenses through the OCUL consortium have good preservation clauses and this is expanding to our national consortium the Canadian Research Knowledge Network (CRKN). In terms of our local licenses (those negotiated by our institution) we often prioritize perpetual access and third party archiving. In practice, we rely on publishers to provide this access on their websites and if asked we would be hard pressed to quantify exactly what we had perpetual access to (it would require a great deal of investigation of purchase/subscription history). We are also not members/participants of Portico, CLOCKSS or LOCKSS, so even where our licenses provide for access via those services/methods we would not be able to take advantage.	7/4/2024 4:17 PM
10	We do not actively preserve e-journal or e-book content acquired perpetually from publishers. We have had internal discussions about resources needed for preservation of this content, but it hasn't been a priority, mostly due to not having had a situation arise that would require us to have a solution in place. We realize this is a bit short-sighted, but it just hasn't been a top priority.	7/2/2024 8:00 PM
11	I answered yes to the previous question, because preservation is certainly a key activity for our digital collections, which we understand to mean our own institutional content for which we hold the copyright. Preservation is NOT a key activity for our electronic resources, which seems to be the focus of this survey.	7/2/2024 4:35 PM
12	Only for things we publish or that we hold in Special Collections (in electronic format). We also have committed to preserving some materials in our institutional respository. We are not focused on preservation of materials to which we subscribe or have purchased electronically (other than supporting things like LOCKSS, CLOCKSS, Portico, etc.)	7/2/2024 12:01 AM
13	We answered questions about library based preservation in relation to licensed content, not content locally digitized by the library. In regards to question 19: While most of our self-hosted content is for perpetual use, we do host at least one collection that will only be made available to our users if a trigger event occurs.	6/28/2024 2:26 PM

14	We focus on preserving our own created collections and not third party databases.	6/26/2024 9:58 PM
15	Libraries have systems and mechanisms for thinking about and preserving ebook and ejournal content. That conversation seems to just be at the beginning stages for audio-visual materials and those formats are going to have significant challenges on their own.	6/26/2024 8:35 PM
16	I am not sure if question 20 refers to all digital collections or just e-journals and e-books. I answered for these since our university owned digital collections is another unit,	6/26/2024 8:32 PM
17	As participants in LOCKSS, we come under regular pressure to reconsider participation due to the local costs of labor and server space that are not inherent in centrally managed programs like Portico and CLOCKSS. Resources are finite, so analysis of which program(s) can preserve the most critical part of a library's digital content (however that is determined) needs to be carried out. The data provided by common third party preservation services is very difficult to compare.	6/26/2024 7:52 PM
18	even if we did preserve how would we provide access legally based on what is happening at IA?	6/26/2024 5:59 PM
19	As a matter of cost containment, we may sometimes choose to avoid securing local copies of licensed digital content, even when allowed by an agreement with a publisher.	6/25/2024 8:54 PM
20	We read this question as applying specifically to instances when we need to take direct action in the preservation, not, for example, Portico, where there is a more passive support for preservation of e-publications.	6/25/2024 5:56 PM
21	There is a greater effort made in the preservation of e-journals than e-books, and I'm not sure why. Makes me wonder why. I'm not sure what type of protections would be available if a publisher went out of business, especially if they host the e-books.	6/24/2024 6:04 PM
22	Q20 : the consortium undertakes preservation as a key activity on behalf of their members. Preservation on our national archive platform is a request in the Couperin consortium's negociation guidelines.	6/4/2024 6:00 PM
23	We have some e-publications preserved on physical discs rather than any servers.	5/31/2024 5:52 PM
24	Q 20: In our situation, there is no general answer to this question. We are in contact with institutions such as Portico, for example, to ensure precisely that e-books from our collection are protected in accordance with their editions. We go to great lengths to ensure that existing holdings here are actually backed up in digital form for the long term before we discard the last copies, for example. Preserving and securing permanent access is therefore an important task with regard to the development of our digital collections.	5/30/2024 3:47 PM
25	Clarity on what you mean by library based preservation. We preserve journal articles/institutional publications in an open access institutional repository - some of which will move to permanent preservation. Long term organisational records and cultural heritage are preserved using a commercial provider, Preservica currently, and we place high priority on this type of material - unique to the organisation. Arguably journal articles are preserved elsewhere and not so high a priority for preservation by us.	5/28/2024 12:16 PM
26	Preservation of e-content is certainly a key activity for the National Library of Sweden, but with a focus on digitized collections, which are stored on the library's own servers	5/23/2024 12:44 PM
27	School libraries do not have the staff expertise, time, or resources to preserve e-publications. Even if licenses allow it, we can't do it.	5/20/2024 8:55 PM

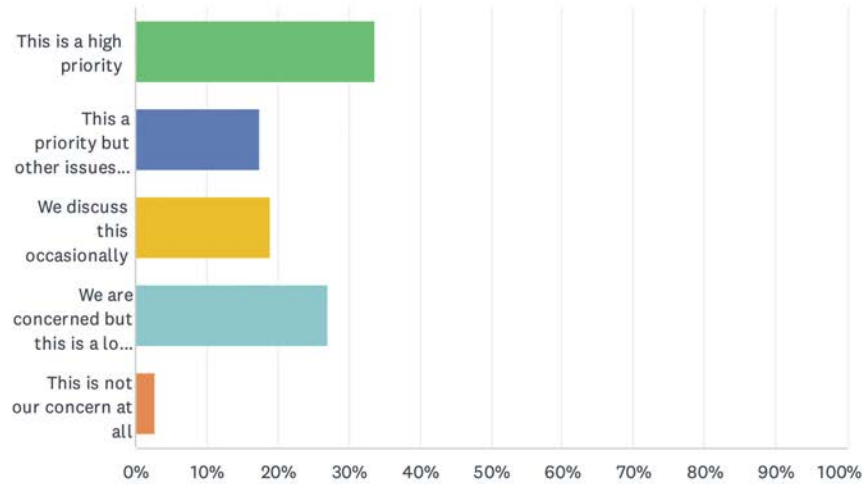## Have you been unable to access an e-publication because it has not been preserved?

Answered: 67   Skipped: 30
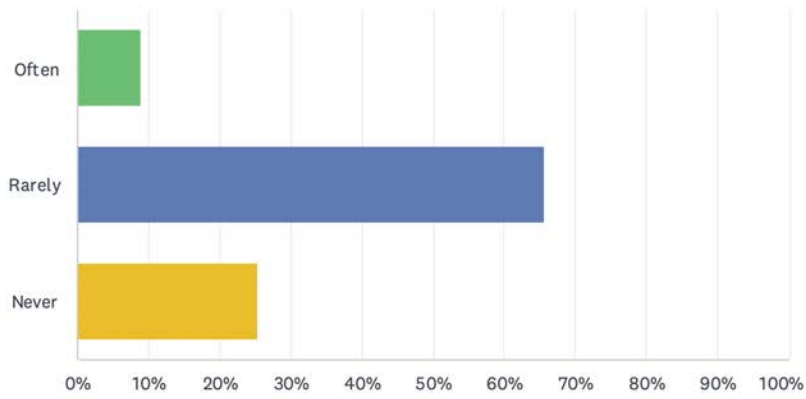
ANSWER CHOICES	RESPONSES	
Often	8.96%	6
Rarely	65.67%	44
Never	25.37%	17
TOTAL		67

To what extent is digital preservation a priority for your organisation?

Answered: 74   Skipped: 23

ANSWER CHOICES	RESPONSES	
This is a high priority	33.78%	25
This a priority but other issues prevent us addressing it	17.57%	13
We discuss this occasionally	18.92%	14
We are concerned but this is a low priority	27.03%	20
This is not our concern at all	2.70%	2
TOTAL		74

APPENDIX D – TRADE AND OTHER ACADEMIC E-BOOKS

The overlooked risks

The neglect of trade books published exclusively in digital formats – spanning genre fiction, popular non-fiction, and other academic e-books – poses an overlooked threat to future knowledge. Without deliberate preservation, vast portions of cultural and literary output risk being irretrievably lost, creating significant gaps that future generations would otherwise seek to understand and study. This issue stems from a complex mix of institutional disinterest, as well as technical, legal, and financial challenges. No single body – whether academic libraries, public libraries, or dark archives – currently assumes responsibility for safeguarding these materials, which may lead to their gradual disappearance.

Trade books, particularly in genre fiction and popular non-fiction, are often dismissed as ephemeral or of lesser value compared to scholarly works. Yet these books are cultural artifacts, offering crucial insights into the social and cultural context of their time. If left unpreserved, future historians and scholars will lose the opportunity to study societal trends, values, and the everyday concerns of the early 21st century. For example, genre fiction like romance novels, which surged in popularity during the COVID lockdowns, could offer valuable insights into the pandemic's cultural impact – yet as things stand, almost none of these works are likely to be preserved for future study.

Digital-only publishing and self-publishing

The sheer number of digital-only trade books, including self-published works, already overwhelms current preservation systems. With hundreds of thousands released annually, there is no infrastructure in place to curate or protect these books. Many are simply ignored, particularly because a significant portion are published through platforms like Amazon, which does not allow libraries to purchase or preserve its e-books. The lack of mechanisms for long-term preservation at scale means that an enormous amount of cultural output risks fading into obscurity.

Distribution and licensing challenges

Distributors such as OverDrive, which supply digital books to public libraries, operate primarily with commercial interests in mind, not preservation mandates. Licensing agreements mean that libraries lease rather than own e-books, and when these agreements expire, the books often no longer appear in their catalogues. There is no guarantee that these materials will be preserved for future generations and even when publishers revert the rights back to authors, they often require the authors to pay for preservation files, creating a further barrier to long-term access.

Academic libraries and public libraries

Academic libraries often disregard trade books, focusing instead on scholarly works that align with their primary research objectives. As a result, they have little incentive to preserve general interest, digital-only trade e-books, particularly those seen as "popular" or genre-based, which are often dismissed as having less academic value. Meanwhile, public libraries, although they frequently offer access to digital trade books through platforms like OverDrive, do not view themselves as preservation institutions. Public libraries do not systematically archive their

collections, especially digital works, and once these books lose their commercial value, they may simply disappear, unnoticed and unpreserved.

Dark archives

Dark archives, which specialise in preserving digital content for the long term, are primarily focused on scholarly journals and academic publications, and largely disinterested in trade books. Legal complexities around intellectual property and the sheer volume of trade books make it unlikely that these archives will, unprompted, take on the task of preservation. Moreover, the restrictive nature of dark archives, which limits access to preserved content, further reduces their potential to safeguard these works for future public use.

National libraries

While national libraries collect some digital trade books through legal deposit mandates, their efforts, limited in scope and intent are often inaccessible to the public. Their collections are preserved in closed systems, making them unavailable for general use. Moreover, the sheer number of trade e-books published each year far exceeds the current capacity of national libraries to collect and store them all.

The long-term consequences: A cultural and academic void

The most alarming consequence of this neglect is the permanent loss of a vast swath of literary and cultural history. Popular non-fiction, like unofficial biographies or social commentaries, often reflect the values and controversies of their time, while academic e-books that serve as learning resources today may become unavailable to future students and researchers. If libraries fail to preserve these digital materials, critical educational content will be lost, leaving a gap in the understanding of evolving educational practices and cultural narratives.

Additionally, genre fiction, often dismissed today as lowbrow or ephemeral, could later prove invaluable for understanding societal norms, fantasies, and anxieties of the early 21st century. By neglecting these works, future generations may be deprived of a rich source of insight into popular culture and everyday life.

The need for action

The neglect of digital-only trade books in genre fiction, popular non-fiction, and other academic e-books represents a looming crisis in the preservation of digital culture. National libraries, while making valuable efforts, cannot address this issue alone. Academic and public libraries, constrained by limited resources and cultural priorities, are not in a position to preserve these works. Dark archives remain focused on academic journals, while commercial distributors prioritise short-term profit over the long-term costs of perpetual preservation. Without coordinated action, much of the literary and cultural output from the early 21st century will be irretrievably lost, creating a lasting void in cultural and academic history.

A collaborative approach involving national libraries, preservation institutions, and publishers is crucial to ensure that future generations benefit from the same rights to learn from, and build upon, the culture and knowledge of the past, as our generation and those before us have enjoyed.

APPENDIX E – TRAFFIC LIGHTS

R/A/G assessment and summary

The following tables are intended as a visual indicator of the current digital preservation efforts across the sector. While each category, when examined in isolation, reveals only partial solutions in operation, with significant vulnerabilities in key areas, there is cause for optimism when considering their collective potential. When viewed together, these efforts offer a promising foundation on which to build, in order to preserve and maintain continuing access to many scholarly digital resources for a century or more.

However, there are currently no comprehensive plans to make the majority of these files accessible to the public, once other means to purchase or otherwise obtain them come to an end, even after they enter the public domain. Still, we commend and encourage all preservation initiatives, as the more digital work is preserved, and the more opportunities for access points are be created, then the greater the chances of long-term success.

The tables also reinforce the core message of this report: at present, national libraries are currently the only institutions systematically preserving e-literature. We emphasise that the key to safeguarding knowledge lies in not only empowering and resourcing national libraries to be able to develop further capacity – not only in storage and preservation but also in the provision of general access – but also in creating shared resources, and building scalable models, to facilitate collaboration across the whole sector.

Academic and scholarly e-journals.

STEM subjects often benefit from some of the most robust arrangements for long-term preservation. The Dark Archives are particularly well-equipped to handle the publishing schedules in these fields, routinely collecting and storing the majority of journals published by major academic publishers. This is largely due to the well-established nature of STEM publishing, which enjoys significant institutional support and a clear framework for digital preservation.

As Clifford Lynch from the Coalition for Networked Information pointed out, "This situation [here] is mostly not terrible". The structured and predictable nature of STEM publishing, combined with the significant resources allocated to these fields, ensures that many STEM e-journals are well-preserved and should remain accessible for the long term.

In the Arts, Humanities, and Social Sciences, the situation is less reassuring due to a more diverse publishing landscape, often involving smaller, niche publishers with limited resources for long-term preservation. Additionally, funding and institutional support are less substantial, and varied publishing formats, technical barriers, and complex legal issues further complicate preservation efforts.

Digital-only scholarly monographs:

The preservation of scholarly monographs presents both challenges and opportunities for libraries and archivists as digital formats become increasingly prevalent. While electronic versions offer ease of access, searchability, and reduced physical storage requirements, long-term trends remain uncertain and print editions are still the preferred choice of librarian and readers. Unlike e-journals, typically available via subscription models, e-monographs are

often sold as one-time purchases or licensed for limited durations.

This creates concerns about future access, especially if a publisher or distributor ceases operations or alters licensing agreements. Securing perpetual copies to digital monographs for the sole purpose of perpetual preservation would benefit both libraries and publishers, and many publishers appear open to this idea, providing clear terms and boundaries can be guaranteed. However, libraries face significant obstacles, as preserving digital materials requires specialised infrastructure, expertise, and other resources that many libraries lack.

National libraries, which have a legal mandate to preserve published works, will continue to be critical in ensuring the long-term preservation of digital-only monographs. Dark archives, designed to store digital content for extended periods, offer another potential solution, although their effectiveness in preserving monographs specifically for access remains uncertain. As digital monographs continue to grow in popularity, collaboration between libraries, publishers, and preservation institutions will be essential. To overcome licensing issues, technical challenges, and resource limitations, all stakeholders must work together to develop sustainable systems for the long-term preservation of scholarly monographs.

Trade and other academic e-books

We have found scant preservation arrangements for these works, which presents a growing concern, particularly as a growing number of titles each year are being published exclusively in digital formats or offered through print-on-demand services. This shift towards digital publishing creates substantial risks for the long-term preservation and accessibility of these publications. Our findings indicate that archiving efforts for digital trade e-books in libraries are minimal at best, leaving a considerable portion of contemporary literature and pedagogic resources vulnerable to loss.

We found no evidence that any substantial proportion of these publications is being systematically placed into long-term preservation systems and, as a result, we conclude the threat level is already serious and likely to become worse. We discuss this in greater depth in Appendix E.

Traffic light assessment

The tables are colour-coded to indicate the current status of digital preservation activities across eighteen fields, organised as two, 3x3 matrices. Each colour represents the level of progress being made toward long-term preservation in academic and public libraries, dark archives, and national libraries, highlighting both strengths and areas requiring further attention.

GREEN	Reflects well-established preservation efforts.
AMBER	Suggests moderate progress with some vulnerabilities.
RED	Indicates significant gaps and high-risk areas.

NB. Areas marked with lighter Green indicate good or very good progress, although we feel there is still a requirement for additional support or resources to reach full assurance levels.

STEM SUBJECTS	ACADEMIC AND PUBLIC LIBRARIES	DARK ARCHIVES	NATIONAL LIBRARIES
COMMERCIAL AND OPEN ACCESS E-JOURNALS	Limited by licensing restrictions, lack of infrastructure, and inadequate resources. Even with licences for perpetual access, they are generally unable to archive or ensure long-term preservation effectively. May be missing opportunity to preserve Open Access e-journals for the long term.	Excel in preserving STEM e-journals through established systems and partnerships with major publishers and benefit from predictable publishing schedules and robust infrastructure, though smaller, less well-resourced journals may still face risks. Overall, they offer a reliable approach for long-term preservation.	Effective in preserving all types of e-journals due to statutory obligations like legal deposit. Maintain comprehensive collections and ensure redundancy across multiple sites. However, their focus has risk of being skewed towards commercial and larger open-access providers leaving smaller e-journals at risk.
SCHOLARLY MONOGRAPHS IN DIGITAL FORMATS	Prefer print editions as preservation systems are robust. Reticent to preserve digital scholarly monographs due to restrictive licensing agreements and inadequate resources. Lack of perpetual access rights and technical infrastructure for digital materials hinders long-term preservation efforts. Potential publisher willingness to cooperate.	Primarily focused on e-journals, and therefore less equipped for preserving digital monographs. While their digital preservation capabilities are advanced, cost, complexity, and licensing restrictions limit their ability to plan for monographs effectively, especially in addressing the eventual public domain availability of these works.	Well-suited to preserving digital monographs, benefiting from legal deposit obligations, and established archival systems. Their focus on comprehensive records and capacity for preservation ensures the long-term safeguarding of digital-only monographs, though balancing accessibility and preservation remains a challenge
TRADE AND OTHER ACADEMIC E-BOOKS	Academic libraries focus on scholarly works, neglecting STEM-related digital trade books due to perceived low relevance. Public libraries, although providing access via platforms like OverDrive, lack formal preservation strategies for long-term storage. Neither institution prioritises preserving these STEM e-books for future scholarly use.	Focussed on scholarly e-journals and do not usually preserve STEM-related trade e-books. Legal complexities and volume constraints make them otherwise unsuitable for this purpose. Additionally, their restricted access models limit the potential for making these materials freely available for future researchers.	National libraries collect some STEM digital trade e-books through legal deposit, but efforts are limited. Preference for print editions. Rarely planning to ensure long-term public access once out of copyright, and the sheer volume of e-books likely exceeds their current storage and preservation capabilities.

AHSS SUBJECTS	ACADEMIC AND PUBLIC LIBRARIES	DARK ARCHIVES	NATIONAL LIBRARIES
COMMERCIAL AND OPEN ACCESS E-JOURNALS	Face significant challenges preserving AHSS e-journals due to licensing restrictions, inconsistent publishing patterns, and limited resources. Ability to archive and ensure long-term access often hindered by such obstacles, impacting the safeguarding of many diverse and irregularly published journals.	Perceived reduced effectiveness for locating more obscure AHSS e-journals. May be challenged by irregular publishing patterns and/or smaller publishers in AHSS fields. Their focus on major publishers means these less conventional journals may not be adequately preserved, posing a risk of loss.	More effective at preserving obscure or harder to locate AHSS e-journals due to their statutory deposit obligations. However, their focus on legal deposit may not fully address the preservation needs of smaller, niche, or non-English language publications, leaving gaps in coverage.
SCHOLARLY MONOGRAPHS IN DIGITAL FORMATS	Prefer print editions as preservation systems are robust. Reticent to preserve digital scholarly monographs due to restrictive licensing agreements and inadequate resources. Lack of perpetual access rights and technical infrastructure for digital materials hinders long-term preservation efforts. Potential publisher willingness to cooperate.	Primarily focused on e-journals, and therefore less equipped for preserving digital monographs. While their digital preservation capabilities are advanced, cost, complexity, and licensing restrictions limit their ability to plan for monographs effectively, especially in addressing the eventual public domain availability of these works.	Ideally positioned for long-term preservation through legal deposit and archival mandates. However, funding constraints and limited expectation for substantial growth in digital-only AHSS e-journals may pose short-term challenges. No plans to offer public access when other availability disappears.
TRADE AND OTHER ACADEMIC E-BOOKS	Academic libraries focus on scholarly content, neglecting AHSS digital trade e-books like popular non-fiction and textbooks. Public libraries offer access but lack systematic preservation, leading to the risk of losing these valuable resources as they degrade. Both types of institutions fail to ensure long-term preservation of AHSS digital works.	Dark archives specialise in scholarly and academic content, not in AHSS trade e-books. Their focus on complex legal and access issues, combined with the sheer volume of digital trade books, makes them ineffective for preserving AHSS digital works. This limits their role in long-term preservation.	National libraries collect some AHSS digital trade e-books via legal deposit but provide limited access. Their efforts are insufficient to cover the vast number of e-books published annually. The preservation practices are often restricted, failing to address the comprehensive preservation needs for AHSS-related digital material.

APPENDIX F – SELECTED REFERENCES

Selected reading

'Learning Lessons from the Cyber Attack: British Library Cyber Incident Review.'
https://www.bl.uk/home/british-library-cyber-incident-review-8-march-2024.pdf

Bergstrom T, Rieger OA, Schonfeld RC. (2024) 'The Second Digital Transformation of Scholarly Publishing Strategic Context and Shared Infrastructure'
https://doi.org/10.18665/sr.320210

Bowie S. (2024) The British Library is a Warning for all Academic Libraries.
https://blogs.lse.ac.uk/impactofsocialsciences/2024/03/19/the-british-library-hack-is-a-warning-for-all-academic-libraries

Dressler VA. (2017) 'The state of affairs with digital preservation at ARL member libraries: A survey and analysis of policy'. Digital Library Perspectives, vol. 33 No. 2, pp. 137-155.
https://doi.org/10.1108/DLP-08-2016-0030

Eve MP. (2024). 'Digital Scholarly Journals Are Poorly Preserved: A Study of 7 Million Articles.' Journal of Librarianship and Scholarly Communication, 12(1), eP16288.
https://doi.org/10.31274/jlsc.16288

Greenberg J, Hanson K, Verhoff D. (2021) 'Guidelines for Preserving New Forms of Scholarship.'
https://doi.org/10.33682/221c-b2xj. ↑

Laakso M, Matthias L, Jahn N. (2021) 'Open is not forever: A study of vanished open access journals.' Journal of the Assoc Information Science Technology, 72:1099–1112.
https://doi.org/10.1002/asi.24460

Meddings C. (2011) 'Digital Preservation: The Library Perspective.' The Serials Librarian, 60:1-4, 55-60.
https://doi.org/10.1080/0361526X.2011.556437

Moody V, Hamilton M. (2024) Mapping Federation Journeys For Optimising the UK Research Infrastructure.' JISC
https://www.jisc.ac.uk/reports/mapping-federation-journeys-for-optimising-the-uk-digital-research-infrastructure

Salo S. (202) 'Is There a Text in These Data? The Digital Humanities and Preserving the Evidence.'
https://direct.mit.edu/books/oa-edited-volume/4933/chapter-standard/625170/Is-There-a-Text-in-These-Data-The-Digital

Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access.
https://doi.org/10.7551/mitpress/11885.001.0001

Sang Y. (2017) 'The politics of e-books', International Journal of Media & Cultural Politics, 13:3, pp. 211–28.
https://doi.org/10.1386/macp.13.3.211_1

Sharp PA et al. (2023) 'Access to Science and Scholarship: Key Questions about the Future of Research Publishing'.
https://assets.pubpub.org/d535ifal/Access%20to%20science%20and%20scholarship%20-%20MIT%20report%20v1.4-41701631814319.pdf. ↑

Warso Z. (2023) 'Digital Public Space.' #Digital Rights Revisited.
https://openfuture.eu/wp-content/uploads/2023/10/231017Digital-rights-revisited.pdf

Wittenberg K, Glasser S, Kirchhoff A, Morrissey S, Orphan, S. (2018). 'Challenges and opportunities in the evolving digital preservation landscape: Reflections from Portico.' Insights: the UKSG journal, 31(0), 28.
https://doi.org/10.1629/uksg.421

Ithaka report into the Need for Shared Infrastructure:
https://sr.ithaka.org/publications/the-second-digital-transformation-of-scholarly-publishing/#concluding-remarks

The 'Nelson Memo' - updating the requirement for access to federal documents:
https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf

William Kilbride (Digital Preservation Coalition) Bit List:
https://www.dpconline.org/digipres/champion-digital-preservation/bit-list

Richard Poynder: Everything you ever wanted to know about OA, but were smart enough not to ask!
https://richardpoynder.co.uk/Jaws.pdf

Mikael Laakso - Open Access Books
https://zenodo.org/records/7305490

Selected books

The Public Domain **Enclosing the Commons of the Mind**	James Boyle	2008
What if we could reimagine Copyright?	Rebecca Giblin & Kimberlee Weatherall	2017
The Case for Books	Robert Darnton	2009
BiblioTECH	John Palfrey	2015
The Wealth of the Commons	David Bollier & Silke Helfrich	2012
Media and Memory	Joanna Garde-Hansen	2011
The Structural Transformation **Of the Public Sphere**	Jürgen Habermas	1991
Athena Unbound	Peter Baldwin	2023
Copyright Wars **Three Centuries of Trans-Atlantic Battle**	Peter Baldwin	2014
Palaces for the People	Eric Klinenberg	2018
Ground Control	Anna Minton	2009
Infrastructure **The Social Value of Shared resources**	Brett M. Frischmann	2012
Public Innovation **Intellectual Property in a digital age**	William Davies and Kay Withers (IPPR)	2006
The Gutenberg Parenthesis	Jeff Jarvis	2023
The Future of the Museum **28 Dialogues**	András Szántó	2020
Book Wars	John B. Thompson	2021

Selected online resources

Keepers Registry:
https://keepers.issn.org

Communia:
https://communia-association.org

Crossref:
https://www.crossref.org

Digital Preservation Coalition:
https://www.dpconline.org

Directory of Open Access Journals (DOAJ):
https://doaj.org

Digital Preservation Management:
https://dpworkshop.org

Internet Archive:
https://archive.org

Open Future:
https://openfuture.eu

Coalition for Networked Information:
https://www.cni.org/

Datacite:
https://datacite.org/

APPENDIX G: BIOGRAPHIES OF THE WRITERS

Tony Ageh, OBE

Tony Ageh is a renowned digital innovator and media strategist, celebrated for his transformative contributions to public service media. Before leading the New York Public Library's digital efforts, he held key roles at the BBC, where he was instrumental in the creation of the BBC iPlayer and overseeing the BBC Archive, revolutionising digital content accessibility. His leadership in digital archiving and preservation has significantly influenced policies aimed at safeguarding digital media for future generations.

Michael Bayler

Michael Bayler is a pioneering business strategist, published author and speaker, specialising in transformation and innovation in sectors including technology, telecoms, banking and financial services, media and entertainment, consumer goods, and life sciences. He has advised many of the world's leading brands, media and technology companies and international stars, including King.com, NatWest, Coca-Cola, Unilever, Zurich Insurance, Diageo, Telefonica, BBC, Discovery, FremantleMedia, Bacardi Global Brands, Warner Bros., BSkyB, BT, SABMiller, Mundipharma, Cognizant, HCL, Nokia, Ogilvy, Robbie Williams and Simon Cowell.

Michael has written two acclaimed business strategy books, "Promiscuous Customers: Invisible Brands – Delivering Value in Digital Markets" (Capstone, Oxford, 2002) and "The Liquid Enterprise – How The Network is Changing Value, What It Means for Business, and What Leadership Needs to Do About It" (Infinite Ideas, Oxford, 2016).

Chris Durlacher

Chris Durlacher is an award-winning documentary-maker. His George Orwell - A Life In Pictures won an EMMY and numerous other awards. One reviewer said, "It changed the rules for bio-pics". He worked on the BBC series, The Secret History of our Streets, which won a Royal Television Society award, and helped reinvigorate the genre of televisual social history. As well as being responsible for over 50 hours of documentary films, he now teaches his craft and works as a consultant in the wider knowledge economy.

Julian Turner

Julian Turner is a publishing board executive who provides advisory, consultancy and leadership support services through Cofunction Ltd. After spells at The Guardian and Euromoney Institutional Investor plc, he founded and was CEO of AIM-listed Electric Word plc for 17 years before its sale in 2017. Julian's current directorships include The Day News & Media Ltd, Kademy Ltd, World Textile Information Network Ltd, Veterinary Business Development Ltd, and Camden Learning Ltd. He has a strong interest in mental health and is an honorary member of the Department of Psychiatry at Oxford University where he is Director of Education at Brainwaves, the adolescent mental health research project

AGEH consulting