

British Library response to the Government's consultation on the Collection and Preservation of UK Offline and Microform Publications and UK Online Publications (Available free of charge and without access restrictions)

The free web contains a rich variety of information and cultural material; a comprehensive archive of the UK web space would be a unique resource of immense value for research of all kinds, and of benefit to a much wider potential audience interested in the UK's national heritage. This is important and urgent. Much material is already being lost and a "digital black hole" is developing, which future generations will find incomprehensible. The British Library believes that Option 2, Legal Deposit Libraries' Regulation-based Harvesting and Archiving, offers by far the best solution to the digital heritage challenges of the free web, for publishers, libraries and the spectrum of users of the envisaged web archive.

INTRODUCTION

1. The British Library (BL) welcomes publication of the Government's consultation on the Collection and Preservation of UK Offline and Microform Publications and UK Online Publications (Available free of charge and without access restrictions). Given the importance of this consultation to the future of education and research and to the sustainable access to quality content, in our response below we have identified four major areas we believe should be addressed. We look forward to engaging actively in this consultation and playing our full part in the policy development process and delivery of major developments.

2. The British Library was established by statute in 1972 as the national library of the United Kingdom. It is one of the world's greatest research libraries – it benefits from legal deposit and is the main custodian of the nation's written cultural heritage. The purpose of legal deposit is to ensure that the nation's published output (and thereby its intellectual record and future published heritage) is collected systematically and as comprehensively as possible, both in order to make it available to current researchers within the libraries of the legal deposit system and to preserve the material for the use of future generations of researchers in higher education, business and personal research. The strength of the BL's current collections and the value we create for the UK is due in no small part to the fact that we have been a beneficiary of legal deposit of printed publications. The legal deposit of e-publications will be equally critical to the Library's strength going forward.

3 The Library's incomparable collections have developed over 250 years; they cover three millennia of recorded knowledge, represent every known written language, every aspect of human thought and a considerable sound, music and recordings archive. Sir Isaac Newton said: "If I have seen further it is by standing on the shoulders of giants". This is what the BL seeks to assist its users to do. In 2007/08, more than 8.2 million

British Library collection items were consulted by, or loaned to, academic researchers, business researchers, and private individuals. An independent economic impact study commissioned by the British Library suggests that the total value added to the UK economy by the Library each year is £363m, or £4.40 for every £1 of public funding.

4. The British Library plays a vital function in the life of the nation as a cultural heritage resource by managing, preserving, and ensuring access in perpetuity to the UK's national published archive and the national repository of sound. The Library is an integral component of both the national research infrastructure and the UK Science Base, and it plays a correspondingly significant role in ensuring the research excellence of the UK. The Library contains a vast array of inspirational material and expertise that supports the creative industries and, through the services of our Business & Intellectual Property Centre, we support entrepreneurs and SMEs in developing, protecting and exploiting their ideas. We operate at the fulcrum of the creative economy and recognise that the ongoing digital revolution in production and distribution technologies is causing fundamental shifts across industry business models and consumer patterns, and is raising broader questions about the traditional balance of rights in intellectual property, between the rights holder and the public good.

5. In 2003 the UNESCO *Charter on the Preservation of Digital Heritage* stated that access to a nation's "digital heritage" will "offer broadened opportunities for creation, communication and sharing of knowledge among all peoples," while warning that it was then "at risk of being lost" and that "its preservation for the benefit of present and future generations is an urgent issue of worldwide concern."¹ The British Library believes the situation is even more urgent today.

6. We believe that the Library can help deliver a truly digital future for Britain by growing faster its role as custodian of Britain's collective digital memory – acquisition of digital content, ensuring its sustainability and its continuing access and long-term preservation – a critical public service that acts as a springboard for research and education, for new forms of creativity, and for knowledge creation. Without such effort and investment, future researchers and citizens will find a black hole in the knowledge base of the 21st century, and without such guaranteed long-term commitment to preservation, our content and creative industries will be inhibited in their global market success.

¹ UNESCO, *Charter on the Preservation of Digital Heritage*, 15 October 2003. Viewed from http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html on 29/12/09

7. There are four areas we would like to highlight in advance of responding to the specific questions posed in the consultation document:

Content – underpinning research. The development of the role of public custodian of quality content in the digital world; the preservation of digital material on a mass scale as a key infrastructure for education, creativity and competitiveness.

Content – underpinning democracy. The proper representation of diverse voices and the safeguarding of public accountability that is only possible through legal deposit.

Enhancing access. The imperative to enhance access to digital material in the UK and beyond to ensure inclusiveness and optimal exploitation of new opportunities.

Urgency. The vulnerability of freely available UK online publications and the urgent need for their preservation.

MAIN BODY

Content – underpinning research. *The development of the role of public custodian of quality content in the digital world; the preserving of digital material on a mass scale as a key infrastructure for education, creativity and competitiveness.*

8. As the national library for the United Kingdom, the British Library has played a major role in upholding legal deposit in the pre-digital world, preserving centuries of printed expression for future generations. In the digital age its custodianship of the archive of UK free web has the potential to bring dividends under legal deposit greater than those for traditional materials – notably powerful in-depth and cross-searching of immense benefit to education, business and personal research.

9. One of the most important aspects of the UK free web is its combination of breadth and depth of content. Any given subject will be represented by news, comment, technical information, guidelines, research, policy documents, blogs and other forms of information.

10. Furthermore, the geographical, social and ethnic diversity of web publishing constitutes one of the most important transformational aspects of the free web. If comprehensive archiving of the free web is enabled through regulation, this diversity will be recorded, preserved and made accessible for generations to come. Without regulation this will be lost both as a resource for current engagement and for future researchers.

11. The web is about connectivity, with the relationships between sites part of knowledge itself. Only a comprehensive web archive can preserve this qualitative aspect to any meaningful scale, facilitating ease of research and remaining faithful to the spirit of the live web and its structure of knowledge.

12. The web responds quickly to national and world events in ways that traditional agencies and conventional broadcast media cannot always achieve. The Library's topical special collections of the UK Web devoted to the 2004 Tsunami, the 2005 London Terrorist Bombings and the 2012 Olympics demonstrate the web's ability to

provide information more quickly, in more depth and in greater variety, than comparable old analogue forms. Despite best efforts, however, these are partial collections because of the limitations of permissions-based archiving. Regulation would radically improve the quality of content across the digital collection.

13. Since 2004, the British Library has archived almost 4,000 individual websites on a permissions-cleared basis. In the Library's extensive experience the very small number of website owners who refuse to have their sites archived overwhelmingly cite the burden of having to clear third party rights. This is compelling evidence that a permissions-based process is particularly discriminating against collaborative and group websites, so excluding the very sites most likely to use the potential of the web most innovatively.

14. If the British Library does not archive the UK free web it is likely that private non-UK organisations will. This is a strategic weakness in "Digital Britain's" infrastructure. The US-based Internet Archive, for example, has been archiving UK sites on a large-scale basis without seeking individual permissions. While the Library applauds the ambition of this project, there are no guarantees that such material has been archived to an acceptable preservation standard, that the works have been archived legally, or that archived sites of UK national interest will remain accessible while in private hands. The Internet Archive would charge, and charge considerably, for a UK public body to recover the UK material the Internet Archive has already archived.

15. Legal deposit, delivering a comprehensive web archive, will make sure that the content of the UK web has a permanent legally-protected place in the nation's collections.

Content – underpinning democracy. *The proper representation of diverse voices and the safeguarding of public accountability that is only possible through legal deposit.*

16. Legal deposit of the UK free web would secure the breadth of publication for current and future research. However, because maintaining the provision of public information is a democratic necessity it would also affirm and underpin the democratic values of the UK.

17. The full democratic range is represented within the UK free web: local, regional and national government sites exist alongside those of political parties, pressure groups, third sector bodies, and advocacy organisations. Many disadvantaged groups and individuals have for the first time found powerful visibility through the free web. They are joined by community sites and a myriad of blogging, interactive and collaborative sites who report and comment on the issues of the day.

18. The web fosters empowerment in society. This should not be negatively impacted by the exclusion of websites from the historical record..

19. The British Library's experience is that, despite best efforts, it is impossible to represent properly the range of voices in any one area of public life through permissions-based selective archiving. Five years of experience suggests, despite apparently general goodwill, that 75% of all sites contacted simply do not reply²: any collection produced through permissions-based archiving is bound to be incomplete and unintentionally biased. Only a legal deposit archive will reflect the full spectrum of activity on the free web and preserve it for analysis by current enquirers and future researchers. The UK's democratic vitality deserves the comprehensive witness only a legal deposit web archive will guarantee.

20. Important publications, from research and practice reports to advice and policy documents, are increasingly only available on the free web, and may be unilaterally taken down at any time. The accountability of all participants in public debate, constituting a central tenet of our democracy, is therefore increasingly vulnerable. This applies to local, regional and national governments, public services, and to the private and voluntary sectors operating in the public domain. A comprehensive web archive, renewed at appropriate frequencies of collection, would therefore not only preserve the fruits of research but underpin the democratic process.

² The likely reasons for non-reply include the administrative burden of correspondence, advertised addresses being out of date, firewalls excluding email requests, and a reluctance to be involved in clearing third party rights. All the evidence from those that do reply is that there is overwhelming support for web archiving.

Enhancing access. *The imperative of enhanced access to digital material in the UK and beyond to ensure inclusiveness and optimal exploitation of new opportunities.*

21. In the pre-digital environment the British Library has kept to the spirit and the letter of the Legal Deposit Libraries Act and its predecessors over several centuries, protecting the reference-only status of legal deposit works. The Library supports enabling access in as far as it does not conflict with legitimate commercial interests in the equitable functioning of legal deposit in the world of print and for paid-for digital content. Paid-for digital content is proposed as the subject of a separate future consultation but the Library wishes to emphasise at this point that, while it argues strongly for the widest possible access to archived websites under Legal Deposit, it does so in contrast to its belief that appropriate access restrictions should apply to paid-for digital content.

22. However, the Library believes that this aspect of the Legal Deposit Libraries Act 2003 is not appropriate to the UK free web. This is because, unlike printed or paid-for digital content, the publications of the live free web are distributed as widely as it is possible to distribute a published work. To corral such websites into a reference-only environment centred on British Library premises is entirely inappropriate to their original form, function and distribution.

23. There is no preservation gain in limiting access to a UK web archive because full access does not increase wear and tear on the digital original.

24. There should be no loss of income to free web publishers by continuing to give access to the original content via the web itself.

25. Giving “24/7” “without walls” access to the web archive would allow the Library to meet its UK-wide responsibilities, saving travel and other costs for millions, and broadening geographical access to information. It would enhance the accessibility of the nation’s creative and cultural resources for new groups of users and allow disabled members of the public easy access to one of the most significant national heritage collections there will be. The current proposal also appears to exclude the making of temporary electronic copies for visually impaired readers, so being more restrictive than the Copyright Designs and Patents Act 1988 as amended by the Copyright (Visually Impaired Persons) Act 2002. The Library seeks an explicit exception for the creation of temporary electronic copies in an appropriate form for visually impaired readers.

26. The internet has created a widespread expectation for information to be delivered via the web. Pioneering web archives, including the permissions-cleared UK Web Archive provided by the British Library, have contributed to this; providing a reduced access web archive would be perceived as retrograde and contrary to the intentions of the original website publishers.

27. The research, creative, and economic opportunities of a UK web archive will be severely curtailed if access is as restricted as it is under the current Legal Deposit Libraries Act. A UK-wide web archive delivered at only a few physical locations in the UK would be wholly out of keeping with the “Digital Britain” approach in which “Britain can sustain its position as a leading digital economy and society.”³

³ DCMS/BIS, *Digital Britain: Final Report*, 2009, p.8. Viewed from <http://www.culture.gov.uk/images/publications/digitalbritain-finalreport-jun09.pdf>

28. *For these reasons, and while of course respecting the law as it stands, The British Library, urges the Secretary of State to seek a legislative amendment which will allow a UK web archive to be delivered via the web to all. However such an amendment should not delay the regulatory process or impact on our ability to start collecting immediately.*

Urgency. *The vulnerability of freely available UK online publications and the urgent need for their preservation.*

29. It must be noted that material published on the web is not preserved automatically, despite a perhaps common assumption that it is. UK web pages are constantly revised or deleted; whole websites disappear every day..

30. One recent study of research papers published on the web over a ten year period showed that nearly one third of the citation-links in research documents no longer linked to the original sources, most of which simply registered as “page not found”.⁴ This is only one of a number of studies over the years which have documented the large scale and growing number of missing links and pages.

31. It is not only current research that is in danger: the historic primary sources for future research are vulnerable, too. Sites devoted to important public events are particularly at risk: more than 150 websites for the 2000 Sydney Olympics are now only available within the National Library of Australia’s archive and would not exist today without the archive.

32. The British Library is now working towards building a web legacy collection for emerging websites associated with the London 2012 Olympics on a permission-cleared basis. However, having been at the forefront of Legal Deposit, the UK is now lagging behind internationally. Many other countries, like Australia, already have the legal authority to archive comprehensively: including France, Germany, Canada, Denmark, Finland, New Zealand and Norway. We urge the government to recognise the importance of legal deposit of the free web to the UK’s competitiveness.

33. In the UK Web Archive classic examples of historic and creative significance that no longer exist in the live environment include the website of the late former Foreign Secretary, Rt Hon. Robin Cook MP and the site of the renowned British composer Peter Maxwell Davies. Both are preserved in the archive for posterity but untold others will have perished before the Library was able to identify and request them. Using the current permissions-requesting system it is impossible to act quickly enough: there is an urgency for whole-domain harvesting.

34. The Library’s experience over five years of web archiving suggests that only 25-30% of contacted web owners actually respond to a request for their site to be archived. On this basis, and despite the apparent goodwill of most website owners, at least 70-75% of free UK websites will be lost forever unless regulation for domain-scale archiving is expedited.

⁴ Mohammad Hanief Bhat , “Missing Web References — A Case Study of Five Scholarly Journals”, In *Liber Quarterly: The Journal of European Research Libraries*, 19 (2009): 2. Viewed from <http://liber.library.uu.nl/> on 29/12/09

35. The permissions approach is very expensive. As the Legal Deposit Advisory Panel figures quoted in the DCMS consultation suggest, the permissions process requires a hefty £6,476 per year per terabyte (Tb). This is to say nothing of the cost and administrative burden on website publishers of the permissions process: the permissions based approach is a “lose-lose” situation for publisher and library alike.

36. Under regulation for domain-scale archiving, costs would be radically reduced to an estimated £215 per Tb. At the same time this would achieve at least seventy times the coverage.

37. The British Library believes therefore that the only economically sensible approach is regulation for a legal deposit archive of freely available UK websites delivered via the web to all.

Responses to Specific Consultation Questions

The British Library's response to the consultation is largely set out above. Below are the Library's answers to the specific questions raised within the DCMS consultation questionnaire.

Question 1: What are your views on the options considered for this content?

The British Library believes that Option 2, Legal Deposit Libraries Regulation-based Harvesting and Archiving, offers by far the best solution to the digital heritage challenges of the free web, for publishers, libraries and the spectrum of users of the envisaged web archive.

Question 2: Are there any other options that should have been considered? If so what are they?

No. However, the British Library believes that the nation's web archive should be freely available via the web without access restrictions, with legal deposit protections. While realising that the restriction on access is required by the 2003 Act, and while, as a matter of course respecting the law as it stands, the British Library urges the Secretary of State to seek a legislative amendment which will allow an archive of UK free web material to be delivered via the web to all. As set out above, the Library urges this in contrast to its continuing belief that appropriate access restrictions should apply to paid-for digital content, the proposed subject of a future consultation. Please see paragraphs 21-28.

Question 3: Do you agree with the analysis of these options? Explain why.

Yes. In the Library's view, Option Two offers the best solution because it is the only one that safeguards UK content in perpetuity under the auspices of a publicly accountable body. It is the only one that will deliver a comprehensive archive rather than a fragmented and biased one. It is the only option that has no cost to the publisher while representing by far the best value for money for the legal deposit libraries and their users, including higher education, business, the creative industries and members of the British (and global) public.

Question 4: Do you agree that harvesting provides the most efficient and timely solution for deposit of publications in this category? Explain why.

See Question 5.

Question 5: Do you agree regulation is the most cost-efficient method of collection for the Libraries and imposes no direct financial or administrative burden upon the publishers? Explain why.

Yes to both questions 4 and 5.

Under Option One (Permissions-Based Archiving) there would be substantial costs on both publishers and libraries. There would be vast inefficiencies and delays because of the bespoke nature of selection and the understandably low rate of publisher response. The goal of creating a comprehensive web archive would not be achieved under this option.

The British Library's experience is that, despite best efforts, it is impossible to represent properly the range of voices in any one area of public life through permissions-based selective archiving. Five years of experience suggests, despite apparently general goodwill, that 75% of all sites contacted simply do not reply⁵. Any collection produced through permissions-based archiving is bound to be incomplete and unintentionally biased. Only a legal deposit archive will reflect the full spectrum of activity on the free web and preserve it for analysis by current enquirers and future researchers. The UK's democratic vitality deserves the comprehensive witness only a legal deposit web archive will guarantee.

Under Option Three (Archiving Left to the Market) there would be no safeguarding of research results or creative materials published only on the web or the range of voices and information presented there. In the British Library's view, it is likely that richer publishers would be more likely to maintain self-archiving but poorer ones would not; there would be no guarantee of public access to the sites that were archived. The only archived material that would be accessible would be there for short-term commercial reasons, with charging likely and a high risk of such limited access being withdrawn should commercial utility be judged to have failed. Websites published by individuals, advocacy and pressure groups, and community organisations would simply disappear.

Option Two (Regulation-Based Harvesting and Archiving) has no cost to the publisher because it is an automatic process initiated by the Legal Deposit Libraries. As the Legal Deposit Advisory Panel figures illustrate, it is by far the most efficient method for the libraries and it guarantees an archive that is large enough and maintained frequently enough to bear proper witness to the remarkable UK free web.

⁵ The likely reasons for non-reply include the administrative burden of correspondence, advertised addresses being out of date, firewalls excluding email requests, and a reluctance to be involved in clearing third party rights. All the evidence from those that do reply is that there is overwhelming support for web archiving.

Question 6: Do you agree that this is an appropriate definition for the type of publications that should be included in scope for regulations? Explain why. Is there anything else that should be included in this definition? Is there anything that should be excluded from this definition?

Yes, this is largely an appropriate definition and one that has clear parallels with traditional legal deposit. As in the past, the non-UK and the private or internal are not eligible or appropriate for legal deposit. With digital, however, UK chargeable content, whose analogue counterpart would have been liable for legal deposit (and will remain so if produced solely in physical form), is the subject of a later, separate, consultation and so it is again appropriate for this category to be excluded from the present consultation on the free web.

The Library notes the exclusion stated in the 2003 Act of “recorded sound and film, where such works comprise the sole or main purpose of the content, or where other material is incidental.” It understands, albeit with regret, that the purpose of this was to exclude cinema films, music records and TV and radio broadcasts from legal deposit. The Library regards this as a structural weakness in the creative and knowledge economy of the United Kingdom and suggests a change in primary legislation would safeguard the collecting of film and sound material under legal deposit. For the purpose of this consultation, however, the Library wishes to emphasise the importance of the legal deposit of sites that use mixed content combining text with sound or film (or both). Such a multimedia approach, for example the use of a podcast within an online article, is a communicative strength of the free web and should be recognised through legal deposit. The Library also notes that, while it is possible to exclude certain websites from the harvest which distribute cinema films, music recordings or TV and radio programmes, it is technically difficult and disproportionately costly to drill down to a nuanced filtering of sites which have mixed content.

The Library believes that the term “Sites with technical barriers” (5.2) is potentially ambiguous. Some sites have certain technological barriers which are not intended to exclude access (for legitimate reasons of privacy) or to generate income but to present content or guide users through a complex site in the most appropriate way, for example database-driven sites. We do not believe that the regulation should exclude such sites from harvesting. Although some technologically advanced elements in some websites may present a challenge to the current generation of web harvesters, they should not be excluded from regulation simply because they cannot yet be harvested. This, the Library therefore suggests, should be rephrased in any regulation.

The Library does not believe anything else should be included in this definition, and does not believe anything has been excluded.

Question 7: Do you agree with the territorial definition of the UK web? Explain why. Is there anything else that should be included in this definition? Is there anything that should be excluded from this definition?

This is broadly an appropriate definition of the UK web, though the Library wishes to seriously question one element.

For UK territoriality, it makes sense that the publishers “should be based in the UK or have a UK address (physical or electronic)”. It appears that there should be the word “and” after the initial defining clause, i.e. “and... Publications should be lawfully published or made available by or on behalf of that publisher from a UK address”.

It is also right that the UK web is about publications “made available to the public”, reinforcing the earlier mentioned exclusion of private materials such as intranets, which are inappropriate for legal deposit.

However, the Library very much agrees with the reported LDAP questioning of the phrase “with no connection to the UK” which the Library finds to be an unnecessary and seriously mistaken exception. It is unnecessary because the totality of the other exclusions are sufficiently narrowing and, unlike this exclusion, they continue in the spirit of the Legal Deposit Libraries Act from which the proposed regulation stems.

It is mistaken because there will be many publications which are fully in scope in terms of UK territory but which will be neither by British authors (if such identity could be ascertained practically at the envisaged scale, which it cannot be) nor on ‘British subjects’ (again, an extremely difficult judgement to make and one that is not practical to do so within domain harvesting). If, hypothetically, it were possible to apply such exclusions this would have a catastrophic effect on the content of the archive. Take for example an open access research paper published on a UK website written by three non-UK scientists on the arguably non-UK subject of changes in the temperature of the Indian Ocean: this would be excluded on both nationality and subject relevance, and the UK research community, and the wider nation, would be the more impoverished for it.

The experience of traditional legal deposit is that authors from all over the world are published in the UK and they deal with all manner of subjects, not constrained by ‘Britishness’. This is a very important UK strength. The Library believes that this continues in the live web environment and legal deposit should continue to record it: indeed, one of the defining elements of ‘Britishness’ may be its lack of parochialism, its hospitality to other voices and expertise and its openness to the full spectrum of human knowledge and creativity. The Library therefore strongly urges that this exclusion be deleted.

The Library would prefer a UK-based internet protocol address to be one element of the UK territoriality definition. However, the following statement in Annex E does present a workable way forward for pre-scoping the 50,000 or so sites affected: “Internet Protocol (IP) addresses may indicate the location of a computer or server connected to the

internet and can therefore be a useful way of identifying whether a publication might potentially be within scope. However the IP address on its own is not sufficient; the publication must either be within a UK domain as above or the publisher must have a physical address within the UK.”

The Library welcomes the statement that “a number of agencies with helpful practices might also aid the Legal Deposit Libraries in identifying publications” and would be keen to work closer with Nominet, the top level UK domain registrar, and other registrars to help identify appropriate sites.

Question 8: Do you agree with this analysis of the UK Web Domain? Explain why. What do you think the impact of your analysis would be?

Yes. It makes sense to use the figures supplied by organisations such as Nominet, the internet registry for .uk domains, and to draw on information supplied by the Legal Deposit Advisory Panel. The British Library intends to use de-duplication methods for static content in the web-archive for efficiency.

Question 9: How do you see a Deposit Library driven system of web harvesting interfacing with a publisher driven duty to deposit under the 2003 Act?

In a sense the publishers are ‘delivering’ their publications to a Deposit Library by publishing on the web: any Deposit Library open to the web will ‘receive’ them. The harvesting process, as set out in the Consultation’s Annex F, secures such deposit. This is wholly in keeping with the Legal Deposit Libraries Act, with the added advantage of placing no burden on the publishers.

Question 10: How could Deposit Libraries most efficaciously ensure a comprehensive body of eligible content is deposited?

In the Library’s view, harvesting the eligible UK domain at appropriate frequencies under the proposed regulation is the most efficient and effective way of guaranteeing a comprehensive body of in-scope content. As suggested above, access to the information held by Nominet and other registrars would be of great help in this process.

Question 11: Do you agree with this costing model? Explain why. Are there costs that need to be factored in or excluded?

Yes, because these figures are based on the real experience of the Legal Deposit Libraries. Harvesting at scale is by far the best economic option, for publishers, libraries and users of the web archive alike since it spares producers the costs of preserving their own materials and it guards against the possibility of valuable research, information, and creativity being lost.

Although they are difficult to quantify, what is missing from the costing model are the opportunity cost losses to future research, to the creative and other industries, to digital

empowerment and to political accountability if legal deposit is not extended to the free web.

Question 12: Do these assumptions adequately reflect the financial burden of publishers? Is there anything that needs to be included or excluded?

As the majority of surveyed publishers suggested, the costs to the publisher of the harvesting under regulation are likely to be minimal or in fact zero. In addition, an archive stores publications that may perish on the publisher's own site, thus preserving the publisher's own content in years to come. This represents a potential saving for the publisher.

The British Library does recognise that making a work freely available on the web does not mean that it is not subject to copyright and the Library very much recognises the concerns of publishers about the potential risks of unauthorised use or copying. As for any analogue material, the British Library will of course respect and uphold the intellectual property and copyright legislation.

Legal deposit also gives protection to the publisher for defamatory material held within the archive.

The Library recognises that time-limited material on a website may warrant special treatment within the archive. For example in a very small number of cases it may not be appropriate to give access to an extract from a novel which appeared on a website for a deliberately brief time (in order to encourage sales of the novel and more users to the site) but which was collected for the archive within a domain-scale harvest. The Library has longstanding experience of embargoing selective examples of physical legal deposit material to protect publishers' overwhelming commercial interests and would apply this principle as appropriate to examples in the web archive.

Question 13: Do you agree with the analysis of these options? Explain why.

Yes. For the reasons set out in this response, the British Library believes that Option Two, Regulation-based Harvesting and Archiving, is by far the best option for a responsible approach to archiving the UK live web.

Question 14: Do you agree with the analysis on making content available to the Deposit Libraries? Explain why. What else needs to be taken into consideration?

The British Library substantially agrees with this analysis of making content available and strongly urges the Option Two regulation. However, for the reasons set out above, the Library proposes a further amendment to the legislation which would see the web archive made available in the medium it was first seen, i.e. via the web itself, with appropriate legal deposit protections for libraries and publishers in such cases as defamation. The Library places this in contrast to its continuing belief that appropriate access restrictions should apply to paid-for digital content, the proposed subject of a future consultation. The Library already operates for both print and digital content an appropriately responsive takedown procedure in such cases as defamation, and this would also operate for archived websites.

Question 15: Do you agree with this costing model? Explain why. What else needs to be taken into consideration?

See Question 16.

Question 16: Do you agree with the analysis of the costs and the impacts of each option? Explain why. What else needs to be taken into consideration?

Yes to both questions: these figures are based on real experience and a detailed consultation to which the British Library has substantially contributed. There is an overwhelming case for regulation-based harvesting on economic reasons alone, before the immense cultural benefits are considered.

Question 17: Do you agree with risks identified here? Explain why. Are there other risks that have not been considered? What would their impact be? Are some of these risks actually not really risks? Why?

Yes: the Library has contributed to the consultation process thus far and believes that these are the key risks. The Library will be able to remove from public access quickly such material identified as illegal in the web archive (and will seek in any case to avoid archiving such material) and, exceptionally, it will be able to embargo or entirely remove from public access material identified as time-limited commercially valuable publications.

Question 18: Do you agree with LDAP's recommendation to regulate for this content? If not, what should be done instead?

Yes, this is the essence of the British Library's response. The Library believes there is no other practical responsible option.

Question 19: Do you agree with LDAP's proposed method for depositing of content? If not how else could this be done?

Yes, as set out in the Library's response to Question 9, the proposed method is by far the best option for publisher, library and nation alike.

Question 20: Do you agree with LDAP's analysis of access provisions? Explain why. What other options are there?

The British Library believes that the LDAP option does not fully recognise the reasonable public expectation that a collection of freely available websites should be delivered freely by the web itself, nor the immense benefits of such access. The Library maintains, however, its continuing belief that appropriate access restrictions should apply to paid-for digital content, the proposed subject of a future consultation.

The Library also believes that, subject to existing copyright law, readers should be allowed to copy and paste text electronically and make permanent electronic copies, again in keeping with public expectations and the original means in which the sites were made available. The current proposal would clearly be an unprecedented check on the sharing of information: in fact it is more restrictive than the law as set out in the Copyright Designs and Patents Act 1988. That Act, the Library believes, should remain the authority on the extent of copying allowed; by the same token, publishers must not be permitted to impose technical barriers that block legal copying. Similarly, the Library would like to make temporary copies of excerpts or pages of archived electronic publications, subject to existing copyright law, for their use in the Library's exhibitions, and for loans to other institutions for their temporary exhibitions.

The Library wishes to draw attention to the issue of "metadata" associated with electronic publications. This is information associated with a publication that may describe what a publication contains, how and by whom it was created, how it is structured, how it should be used and how it relates to other works. In terms of the Legal Deposit Libraries Act this constitutes "information necessary in order to access the work" (Section 6 (b)) and is also liable for deposit. The Library needs to have the explicit legal authority not just to receive metadata but to be able to extract, use, adapt and enhance it, to expose it for resource discovery internally and externally, and to share it with other organisations. It is only by doing this that the Library can give full and proper access to the deposited electronic publications in its care, and further the sharing of knowledge.

As set out in the above document, the Library urges the Secretary of State to seek an amendment for freely available UK online publications only that would afford legal deposit protections to the national collection of archived websites accessible to all via the web itself.

Question 21: Do you agree with these cost assumptions? Explain why. What needs to be included or excluded?

Yes: the British Library has substantially contributed to these costings based on experience and information gathering and believes they provide a reasonable basis to predict and plan. The Library believes the evidence they present points directly at regulation for harvesting as by far the best option for securing the UK live web for posterity.

For further information please contact:

Victoria Carson
Public Affairs Manager
The British Library
0207 412 7790
Victoria.carson@bl.uk