

Legal Deposit of Non Print Works: A Response from the Digital Preservation Coalition

Introduction

1. The **Digital Preservation Coalition (DPC)** is a not-for-profit membership organisation whose primary objective is to raise awareness of the importance of the preservation of digital material and the attendant strategic, cultural and technological issues. Its vision is to make **our digital memory accessible tomorrow**.
2. In response to your previous consultation on electronic legal deposit we welcomed the proposal for regulation based harvesting of free to web publications which we believed offered a significant step forward to ensuring long term access to vital digital resources. We noted the challenge of capacity building required to implement these proposals and we noted that, although there were different views on access to collections, **future access will be impossible without the safeguards that rigorous attention to preservation provides**.
3. In our response to this consultation we note again the **essential relationship between preservation and access**. We note and **welcome the proposal that extend legal deposit** to include charged content as well as content to which access is restricted. This will create the conditions where a more rounded and more valuable national archive can be created.
4. Experience in digital preservation shows that normalization and adherence to standards in the creation of digital resources are advantageous to long term access. Therefore we have some **questions pertaining to the practicality of provisions regarding deposit of materials**, in particular those regulations that leave the medium and quality of electronic deposits at the discretion of publishers, and those regulations that pertain to adapting content for preservation. If poorly implemented, these provisions could have the inadvertent result of making preservation intractable or excessively complicated.
5. We recognise that recommendations from the DPC are best focussed on those topics where we can offer specialist commentary. Therefore it is not our intention to provide a detailed scrutiny of each element of these regulations. However it is our view that **preservation is only sensible within the context of access, and that preservation should be configured around the impact that comes with access**. Therefore we have commented on a small number of access issues that we believe have a bearing on the case for preservation.
6. The membership of the DPC includes museums, libraries, data centres and archives in a variety of sectors; public sector agencies that fund content development; rights holders; professional bodies; and researchers. This response has benefitted from discussion amongst this diverse community.

Introductory Questions

7. You asked (1), 'Will these regulations provide for a meaningful national archive of non-print publications to be deposited with the Legal Deposit Libraries?'
8. We welcome these new regulations and believe they take a considerable and timely step towards the maintenance and ongoing development of a national digital archive. Electronic production processes have transformed the print industry and electronic distribution has transformed the way that the public engages with published content. These great changes have so far had only limited impact on the operation of legal

deposit which has remained almost entirely based in printed output. The rapid rise of electronic production and dissemination has been matched by a growth in the capacity of Legal Deposit Libraries to manage increasingly large and increasingly complicated digital publications. So the opportunity to develop a meaningful national archive of non-print publications will not be lost.

9. You asked (2), 'The Regulations are intended to cover all types of non-print works: Do you agree with this approach?'
10. We welcome the approach of ensuring that a broad range of work falls within the purview of the regulations. Although we recognise that the term 'publication' as used in the regulations derives from the 2003 Legal Deposit Libraries Act, the term seems increasingly unhelpful to describe the diverse holdings within libraries as the nature and boundaries of digital resources become more complicated. It is important that the regulations do not inadvertently prevent the Deposit Libraries from collecting and preserving material in emerging formats or media.
11. You asked (3) 'The regulations provides illustrative examples of the types of non-print publication which are covered. Are there other examples of non-print publications which should be expressly included?'
12. Although this list is helpful in the context of illustrating the regulations we think it might be sensible to provide other examples of emerging forms of publication – such as e-book, e-journal and website. The rapid evolution of non-print publication means that a prescriptive list would likely be damaging to the purposes of the regulations. The context makes it clear that the list is illustrative rather than prescriptive.
13. Concerning regulation 2(3), you asked (4) whether offline publications, on line publications (both free and charged, restricted and unrestricted) and electronic publications comprising material packaged in response to a query should be included in the regulations.
14. Yes. In our view the regulations may create complications for the preservation of 'orphan' digital sound and video recordings – specifically those sound and video recordings that meet the criteria pertaining to web harvesting, which have no obvious archival home and no clear copyright owner. It is appropriate that the vast majority of sound and video – such as published by the music industry or by broadcasters – remain outside the scope of the legislation. But there remains a small amount of orphaned sound and video content which has no home and which could be preserved as part of a web harvest if the legal deposit regime were sufficiently sensitive to this need. By definition, it is hard to license orphan works, making permissions-based deposit or voluntary deposit problematic. We recognise that sound and video recordings are currently outside the definitions of the Act, and therefore could not be included here by regulation alone. Moreover it is unlikely that the Legal Deposit Libraries would welcome such a substantial addition to their mandate, nor have the producers of such content been consulted. Nonetheless, we believe that the review of the regulations (proposed in paragraph 65 below) should examine the state of preservation of 'orphan' sound and video. Such a review should measure the impact which this perceived complication has had on the goal of creating a 'meaningful national archive of non-print publications'; the relative merits of voluntary or regulation-based deposit of orphaned sound and video; and the relative ease with which existing tools for web harvesting could be extended to gather sound and video.

Draft Regulations: Deposit

15. You asked (5) 'when substantially the same work is published in both print and non-print the regulations must provide that medium for delivery is print unless an alternative medium has been agreed?'

16. This regulation defaults to the status quo and as such it implies the least amount of disruption to an existing system of deposit. While reasonable in the short term it is our view that, as mechanisms for electronic deposit become more sophisticated, efficient, dependable and trusted, so the justification of defaulting to the status quo will become less convincing. If this prediction is right, a regulation that defaults to print only copies will become a hindrance for publishers and libraries alike but that some may use this regulation to resist sensible change. Consequently we believe that permission to harvest in electronic form should not be withheld unreasonably and we recommend that this regulation be reviewed within a fixed time reference to ensure that it does not inadvertently become an obstacle to change (see paragraph 65).
17. You asked (6) concerning regulation 5(2), whether it was appropriate that publishers and Legal Deposit Libraries should agree on how to establish whether works are to be regarded as substantially the same, or whether this should be provided for in the regulations.
18. It seems unlikely that the regulations could encompass the rapid change of technology and even in the technologies currently available it can be hard to decide what constitutes a 'significant variation'. Consequently it is best that this regulation be left as a discussion between publisher and deposit library.
19. You asked (7) whether there were any consequences made regulation 5(3) impracticable, whereby if the Legal Deposit Libraries and Publishers cannot agree on a medium of delivery the publisher will be left to decide how to deposit a work?
20. As currently written, we believe this regulation could give rise to unnecessary complications and at worst may be impracticable. Such complications can be avoided but only if a trusted mechanism can be developed within the governance of legal deposit that deals effectively with disagreements.
21. Experience in digital preservation shows that media – especially file format - can be a significant challenge to long term preservation. Strictly speaking very few file formats are impossible to preserve, but the proliferation of formats and the supply of malformed or non-standard implementations of file format standards creates a significant management overhead for the preservation facility since it will be necessary to test and transform each file on ingest. Moreover, inefficient formats which contain redundant code can become very large as against the size of the information content being preserved creating unnecessary storage and delivery overheads. While these issues will not of themselves prevent effective preservation, they could have a serious effect on the cost and effectiveness of the preservation actions. Simultaneously evidence suggests that the costs of preservation are loaded to ingest and therefore anything which adds to the costs of ingest could have a multiplier effect on costs overall. Although most file formats in the print industry are well understood, it should be recognised that there are very many localised or specialised implementations of standards. Therefore the default position of simply allowing publishers to decide on the medium of delivery could create a situation which is unworkable, preventing the effective implementation of the other parts of the regulatory framework.
22. This question of file formats for ingest has been widely discussed within the preservation community. For example the National Library of Australia has invested in creating a set of tools which provide for migration at ingest, thus reducing the proliferation of file formats – but these tools are only effective for a small set of formats and only where the code is well formed. Some institutions have tackled the proliferation of file formats by creating a canonical list of file formats that they will accept, either rejecting those which do not meet the criteria, or refusing to guarantee long term access. Other institutions have sought to resolve the question of file formats by examining the mission of the institution and matching the selection of file

formats to the needs of their designated community, or through pragmatic decisions based on affordability and available skills (see Todd 2009 and references for a full discussion of this issue).

23. Experience in digital preservation would suggest that obsolescence of file formats is less of a challenge than once thought (eg Rusbridge 2006). However experience also shows that the effort and expense of retrieving data from non-standard or malformed formats can be great while standardisation and automation reduce costs. Consequently, it is reasonable to expect Legal Deposit Libraries to be able to cope with any format: but only if they are very generously resourced.
24. The options outline in paragraph 22 are unlikely to be available to the deposit libraries and even if they were, we do not believe that regulations should be framed in a way that could drive up costs needlessly or put the preservation facilities in a position where they will not be able to deliver the service required of them.
25. By far the most satisfactory way to resolve this issue is to ensure that there is consensus and dialogue between publishers and Legal Deposit Libraries and as second best for there to be trusted and robust mechanisms to resolve disputes built into the governance structure of Legal Deposit. As far as file formats are concerned, it may help if publishers and deposit libraries could formalise their consensus through flexible but clear code of practice which balances the resources of preservation facilities while minimizing the costs to publishers. Such a code of practice could be based on practical experience of preservation, define the minimum standards required for preservation and illustrate the benefits to publisher and library alike. The code of practice could be subject to periodic revision as new standards are adopted and could be subject to peer review by the preservation community.
26. You asked (8) concerning regulation 5(3), whether it was appropriate that publishers and Legal Deposit Libraries should agree on how to establish whether works are to be regarded as substantially the same, or whether this should be provided for in the regulations.
27. See paragraphs 17 and 18 above.
28. You asked (9) concerning regulation 13, whether in the event that Legal Deposit Libraries and Publishers cannot agree on the quality of a work to be preserved it was appropriate that the publishers should decide on the quality of works to be deposited
29. See paragraphs 19-25 above. We believe that the issues associated with quality are analogous to the issues associated with medium and therefore that there are risks with the current proposals which should not be allowed to complicate or prevent their operation. We firmly believe that the best way to address this question is by achieving consensus between libraries and publishers, and failing that to build robust and trusted mechanisms to resolve disputes into the governance of the legal deposit. As with file formats there may be scope for the development of a formal but flexible code of practice to balance the reasonable expectations of the publishers for no additional costs as against the requirements of the preservation facility to provide viable long-term service.
30. You asked (10) concerning regulation 14 whether we agree with the approach proposed for online publications which are free of charge and which are not subject to access restrictions.
31. We have commented in more detail in our response to your earlier consultation on online free publications. Details are available at: <http://www.dpconline.org/advocacy/direct-advocacy/dpc-response-to-electronic-legal-deposit>

32. You asked (11) concerning regulation 17 whether we agree with the approach proposed for online publications for which there is a charge or which are subject to access restrictions.
33. Note our comments on quality and media which we believe need to be further addressed (paragraph 19-24).
34. You asked (12) concerning regulation 19 whether 'A request for deposit must be made in writing (whether sent by a web harvester or other means). Are there consequences that make this approach impracticable?
35. We do not know of any consequences that make this approach impracticable. We can see that the first request may require a small amount of work from the publisher to permit access, and we can also see that there may be a need to agree a schedule of harvest to maximize effectiveness of the process – especially for the first harvest which may be exceptionally large. But subsequent harvests should be simpler.
36. You asked (13) concerning regulation 21 whether it was appropriate that publishers and libraries agree on the most suitable quality for preservation of online publications and, in the event that agreement cannot be reached whether it is appropriate for the publisher to decide?
37. See our comments in paragraphs 19-25 above. We believe that the issues associated with quality are analogous to the issues associated with medium and that there are risks with the current proposals. We are clear that consensus between publishers and libraries is the best solution, and failing that, robust and trusted mechanisms for dispute resolution will be advantageous. A formal but flexible code of practice could be drawn up to balance the reasonable expectations of the publishers for no additional costs as against the requirements of the preservation facility to provide viable long-term service.

Draft Regulations: Permitted Activities

38. You asked (14) concerning regulation 23 whether we agree with the approach to access, limiting concurrent access to one terminal in each of the deposit libraries, a regulation that seeks to mirror the system for printed publications.
39. There are different views in this proposal within the Coalition but we are agreed that overly strict restrictions on access undermine the case of preservation. Members have been encouraged to respond directly with their own views. However we are concerned that this proposal does not adequately mirror the current system for printed publications which historically have been available for inter-library loan and document supply. Therefore a closer match with the existing regulations would result in a method for preservation copies being made available under regulations outside the premises of the deposit libraries.
40. You asked (15) concerning regulation 24 whether we agree with the approach that allows publishers to request an embargo for a period not exceeding three years for online publications, and a mechanism (regulation 27) by which that embargo might be extended
41. There are different views in this proposal within the Coalition but we are agreed that overly strict restrictions on access undermine the case of preservation. Members have been encouraged to respond directly with their own views. However the case for preservation is about the impact that can be derived through access. Therefore it is our view that the facility for an indefinite embargo that seems to be implied by this regulation could undermine the case for preservation.

Research and Private Study

42. You asked (16) concerning regulations 29 and 30 whether we agreed with the provision that a library may only grant access to a print copy of a work unless permission has been given by the publisher to provide a copy on another medium.

43. We have no view on these regulations.

Copying and adapting for preservation purposes

44. You asked (18) concerning regulation 32 and 33 whether we agreed with the approach for copying or adapting works for preservation purposes?

45. We agree with the provisions made for copying and adapting works but seek clarification on a number of points.

46. Firstly the outcome Gower's Review of the preservation exceptions to copyright is likely to have a direct bearing on these regulations. This review recognised that good practice within computing was to provide multiple copies as a way of reducing risks of data loss. In our response to the Gower's Review we agreed with this provision but also noted that there was virtue in permitting institutions for preservation copies to be replicated outside of an institution, and that too narrow a definition of the permanent collection could inadvertently constrain reasonable endeavours to assess and stabilise content.

47. Therefore provided the recommendations of the Gower's review are taken forward to permit multiple preservation copies, and provided the Gower's review also permits preservation copies to be lodged with third parties, and for reasonable appraisal and stabilisation to take place, then the regulations proposed here in respect of copying are likely to be sufficient. If the Gower's review is not taken forward or if these special conditions are omitted from its ultimate implementation then the regulations will not be sufficient. Provision will need to be made for multiple, off site copies and for those exceptions to include copying before items are accessioned into a permanent collection.

48. We note also the assumption that databases and computer programmes may be subject to adaptation for preservation purposes. This is welcome but seems like a special dispensation for databases or computer programs. Experience in digital preservation places emphasis on the need for all data types to be available to some degree of adaptation. Needlessly constraining the adaptations that can take place by restricting them only to computer programs and databases greatly complicates the process of preservation planning and is likely to prevent effective preservation from being delivered.

49. Consequently we seek clarification of these regulations. We believe as currently framed they will inadvertently but comprehensively prevent reasonable preservation actions for certain classes of digital object.

Disposing of Copies of Deposited Works

50. You asked (19) concerning regulation 34 which permits Deposit Libraries to dispose of duplicate works by destroying them.

51. We have not comment on this regulation

Draft Regulations: Exemption from Liability

52. You asked (20) concerning regulation 35 whether the geographical scope of the exemptions for liability were appropriate?

53. We have no comment on this regulation.

54. You asked (20) concerning regulation 35 whether the geographical scope of the exemptions from liability were appropriate?

55. We have no comment on this regulation.

56. You asked (21) concerning regulation 36 whether it was appropriate that the regulation should not cover works which are accessible to readers in the UK unless they are published in the UK by publishers based in the UK?

57. We have no comment on this regulation.

58. You asked (22) concerning guidance provided under section 8 if the scope of collecting was appropriate.

59. We have no comment on this regulation.

Impact Assessment

60. You asked (23-29) about the financial and regulatory impact of the proposed regulations on publishers?

61. We have no comment on these question

Ireland

62. You asked (31) whether the approach to Trinity College Dublin was appropriate in the current regulations?

63. We have discussed this with representatives of Trinity College Dublin who are members of the Coalition and encouraged them to respond to you directly on this matter.

Other

64. You asked (32) if there were any other comments issues or concerns that we wanted to raise?

65. We wish to comment on the Sunset Clause associated with this regulation. While we understand the requirement for regulations to reviewed, we believe that there is need to take account of the enduring mission of memory institutions like the Deposit Libraries. For example we believe it would be unhelpful if one consequence of the sunset clause was to require the Deposit Libraries to retain but no longer adapt or provide access to collections developed using the regulation. Therefore we believe it would be more appropriate in this case to remove the sunset clause and replace this with a requirement for the Secretary of State to review the operation and effect of the regulations (as described in paragraph 1.12 of the consultation). In doing so, the Secretary of State would take account of the specific long term mission of the deposit libraries in reviewing and recommending any changes to the system of legal deposit.

66. In considering such a review we have already identified certain items which we think need to be reviewed once the regulations are in operation: deposit of 'orphaned' sound and video recordings (paragraph 14); and transition to electronic only deposit (paragraph 16).

Citations

Rusbridge, C 2006 'Excuse Me... Some Digital Preservation Fallacies?' in *Ariadne* 46, online at: <http://www.ariadne.ac.uk/issue46/rusbridge/intro.html> last visited 20/12/2010

Todd, M 2009 'File formats for Preservation', DPC Technology Watch Report 09-02 online at: <http://www.dpconline.org/advice/technology-watch-reports> last visited 20/12/2010

Dr William Kilbride FSA,

Executive Director of the Digital Preservation Coalition

December 2010

About this document

Version 1	Document initiated	20/11/2010	WK
Version 2	Distributed to Board for comment and discussion at meeting on 1/12	01/12/2010	Board
Version 3	Distributed to members for comments	06/12/2010	WK
Version 4	Comments received and edited into preview	13/12/2010	Membership
Version 5	Final version submitted	21/12/2010	WK