Annex 3: Selected comments on particular questions or issues, taken verbatim from the responses

NB Comments have been taken from the responses, and copied to this annex, where they appear to contain points of interest. It is therefore a subjective collection of comments, brought together as they may be of interest.

Questions are only noted where comments have been copied; eg question 1.1 was not one which invited comment, so it does not appear in this annex.

1. Policy and strategic issues

1.2 Do you have

a) a digital preservation strategy, or

b) any other policy or strategy document(s) which refer to digital preservation?

c) any alternative to a digital preservation strategy?

If so, would you be prepared to share these documents or references with us?

BBC: We are committed to migration, because there is no serious implementation of emulation as an approach for digital video and audio formats (very hard to emulate a digital videotape player), and migration has been our approach

1.3 Do you have a corporate risk management process?

If so:

a) does digital preservation figure on it?

b) is it defined as a long, or short, term risk?

CCLRC: In the sense of data storage and loss of data. Data storage is viewed in terms of cost and technology. Loss of data is viewed under the terms of cost and business impact.

NAS: We have a corporate risk register which defines as risks the possible proliferation of electronic record systems and increased demand for records held by NAS, but not specifically digital preservation per se. These risks are perceived to be relatively long term and are seen as equally issues of resource and technology. The proposed countermeasures to these risks include working with others such as the DPC and developing NAS's capacity for dealing with electronic records. In other words, digital preservation is seen as a potential solution to other risks, rather than as something involving risk itself.

1.4 Is your responsibility for digital preservation driven by one (or more) of the following? :

a) a statutory requirement to keep/store/preserve material?

b) A statutory duty to provide public access?

c) Some other statutory requirement or duty?

d) A core business requirement?

AHDS: The AHDS collects, distributes and preserves digital resources created in the arts and humanities. We are an entirely digital research archive, holding digital outputs of academic research and collections of digital surrogates of use in research (and learning and teaching). Our key responsibility is to preserve this material, ensuring that it is available for secondary analysis, reuse, and critique in the future.

CCLRC: (on 1.4 d)) mainly d – data is expensive and often cannot be reproduced

1.5 Are you aware of any statutory or legal issue or problem that might complicate or impede your organisation's digital preservation activities?

CCLRC: Ownership of scientific data is very unclear and a minefield in the UK.

NAS: As with the 1958 Public Records Act, Scottish legislation was passed before electronic records were really thought of, so some updating will be necessary if we are to make sure that our preservation of electronic records is properly covered by statute. As with the UK, plans are being made to consider the need for new public records legislation in Scotland and to develop an overall Scottish public records strategy.

2. Current projects [for all respondents' replies to this question see separate annex 4]

3. Volumes of material

3.1 Have you carried out any survey of digital material which you already hold for long-term storage?

If not, can you offer any approximate estimate as to volume?

AHDS: The AHDS currently holds approximately 1TB of data, comprising about 4,000 distinct collections, containing . This total will rise to 4TB within the next year months

BBC: Have surveyed web archive requirements, and have detailed knowledge of analogue and digital audiovisual material and its preservation requirements. Over a three-year period we will be capturing 3.5 TB of data which represents a selection of the BBC's web output. This will be stored on disk, backed up to digital linear tapes. Our 10-year audiovisual preservation project will produce about 500,000 hours of digital material of various types – approximately 40 Peta bytes of data.

PRONI: has not surveyed 'born digital', but notes that it currently holds 1 Tbyte under its digitisation programme.

UKDA: All our holdings are catalogued and listed. C. 4000 'studies' – each study can (and usually does) consist of multiple datasets and associated metadata materials. Whole collection = c. 3 Tbytes.

3.2 Have you assessed the volumes of digital material you will have to preserve in the future, and the likely rate of growth?

AHDS: No. Difficult to anticipate, given that we neither create, nor fund the creation of material that will come to us for preservation. We anticipate sharply rising volumes of data from a moderately rising number of depositing projects. Initial planning for the new AHDS digital repository will give it a capacity to hold 10TB of data.

TCD: No. Unpredictable at this stage, but likely to be rapid if Irish copyright material is deposited with TCD.

UKDA: We have facilities to increase core collection to c.10 Tbytes over next 4 years. Ad hoc collections will be added from time to time, separately costed (eg. Old Bailey Project; BOPCRIS etc.)

3.3 On what basis you have made these assessments?

UKDA: Requirement to increase capacity three-fold over course of current ESRC award.

3.4 Do you have a selection policy for material for digital preservation?

AHDS: Material may be refused on the basis that it is, a) outside the scope of the AHDS collection policy, or, b) does not meet minimum technical requirements (e.g. very poor structure, use of undocumented custom code etc.). The AHDS does not normally seek to hold source code or executables. The AHDS normally aims to preserve the underlying data of a digital resource, not the systems that provide access to that data.

CCLRC: selective by funding

OU: In the printing process we store copies of everything created but the data attached to these documents

varies. We only add metadata to items perceived as having a reuse value or of learning and teaching standard. Decisions are still being made about other types of resources, such as email, web-based materials, etc

OULS: The application of legal deposit to electronic publications will put selection policies in place for one category of material. The libraries Collection Development policies in certain areas are currently under review, and will incorporate digital preservation arrangements where appropriate.

UKDA: All potential acquisitions are assessed by an Acquisitions Review Committee. Criteria are quite flexible, but main factor is potential use for research/teaching. Under this system, we reject probably 2-3 studies for every one taken in.

4. Types of material **[for all respondents' replies to this question see separate annex 5]**

5. Resource issues

5.1 Have you ever asked your major funding source(s) for extra resources for your digital preservation programme?

BBC: We don't distinguish between analogue and digital preservation – it is just a 'format change'. Even the change from videotape to datatape to servers is a format change. Our Preservation Project sought special funding of £55m pounds (after already spending £5m on smaller projects and pilots).

JISC: formal request to Funding councils to increase JISC funding allocations to digital preservation. £625k p/a for digital preservation programme, £750Kp/a for 3 years for Digital Curation Centre. Details - see JISC Continuing Access and Digital preservation Strategy and Implementation Plan.

PRONI: Currently conducting a Strategic Network Review and an IT Organisation Review. While these will not include detailed resource estimates, they are expected to identify digital preservation as a major issue, outline storage options and identify high-level resource issues.

5.2 Is funding for digital preservation likely to become a major issue for you over the next five years?

AHDS: Only in the sense of ever-increasing costs related to ever-increasing volumes of materials and number of collections.

BBC: Digital *preservation will be cheaper than the conversion from analogue, so digital preservation is a welcome outcome, and will gradually reduce our requirements for preservation funding*

CUL: (a) Legal deposit obligations, (b) The sustainability of the DSpace@Cambridge repository when project funding ceases after Dec 2005.

CCLRC: Yes – funding of science does not currently want to recognize that digital preservation is both necessary and expensive. It's hard enough getting funds to do projects which generate the data.

JISC: Yes. Predicted growth of materials needing long term curation within the sector.

NAS: Increasing amounts of the information NAS takes in will come in digital form, and increasing numbers of the records we hold will be digitised for access purposes. We will need to have robust systems in place for ensuring that this information remains authentic and accessible over time. We will need people who can devote time to developing the necessary policies and procedures and we will need money to develop the IT infrastructure necessary to carry these out.

NeLH: Not funding, but technology!

NHM: It will require management – specifically creation of context-related metadata, resilient storage facilities, user and administrative interfaces and (not least) culture change management. Pressure for enhanced access anticipated as a result of Freedom of Information legislation from Jan 2005.

NLS: Transition from voluntary deposit to legal deposit (and therefore preservation) of electronic material. During the voluntary phase it seems that more emphasis has been placed on what we know, i.e. cataloguing, than what we are less sure about, i.e. digital preservation.

NLW: Yes. NLW is unlikely to receive additional funding from sponsor body (the National Assembly for Wales) in the near future. This would have serious implications on 3 of the Library's 5 core functions,

specifically Collecting, Preserving and Giving Access and Information to the diverse material in its care.

OU: The preserving of online course materials is only just being addressed and the financial implications have yet to be calculated. Finding the resources to even piloting a scheme is proving problematic. The volume of the materials is one issue but the main issue is the systems and human resource necessary to achieve this.

OULS: It will become a resource issue because of the extension to legal deposit. Because of the legislative framework in respect of public records; and because of the imperative to cope with eManuscripts as an extension to our existing repository functions.

TCD: Yes, we will require specialist staff, facilities and equipment.

ULCC: Other people's funding is a major issues, since we depend to some extent on other people being willing to pay us to do digital preservation for them. If they don't have funding, we don't have funding.

Wellcome: The Library is increasingly being offered archives that have a significant digital component. For example, we have just secured Sir John Sultson's archive. Much of this is the form of emails. Thus, we will need a system to "accession" this and make it available – and recognizing that the volumes of digital objects are high, we need some way to automatically extract metadata. For example, if we acquire thousands of emails – there is no way we can manually catalogue these. We need a system which automatically extracts metadata – such as sender, recipient, subject, date etc. Such systems are unlikely to be cheap.

5.3 Have you assessed your funding requirements to meet your digital preservation responsibilities:

a) in the short-medium term (ie 1-5 years)?

b) in the longer term (ie 5-10 years or longer)?

AHDS: The figures are based upon likely staff requirements to set up and manage the AHDS preservation repository, and to put in place the necessary hardware, software and off-site back-up systems. In addition, funding requirements for the acquisition and curation of digital collections are costed and included in the AHDS budget.

BBC: Our requirements were mainly based on the existing analogue 'legacy material'. We will need further work to estimate funding for a second 10-year project, which (except for film) will deal mainly with born-digital and already-migrated-to-digital material.

CUL: Assessments are based on the Oct 2002 impact study on extension of legal deposit. They assume that a single copy of each item is shared among the legal deposit libraries. (a) £0.5 million, (b) £1.1 million. The DSpace@Cambridge project includes provision for a business analyst to assess the long-term cost implications and identify appropriate business models.

JISC: Need to do more to cost this. Position is complicated by the nature of JISC's role within the sector and fact that primary responsibility for preservation often rests elsewhere.

NLS: This is made more difficult because the extent of our role alongside other legal deposit libraries is unclear, as are our associated requirements.

NLW: Longer term funding requirements are more difficult to identify partly due to the unknown impact of the new legal deposit legislation.

5.4 Is your present resource for digital preservation:

a) from within existing resources AND/OR:

b) new money allocated by your funding source specifically for digital preservation

CUL: Thus far funding of non-project-based digital preservation has been provided from within existing resources, but CUL's capacity to cope with growth of this activity is now exhausted.

JISC: As a funding body JISC funds services and projects. Re-allocation of project funding is relatively straight forward. Services with preservation component or any ongoing commitment will clearly reduce this flexibility over time

5.5 Is your digital preservation requirement totally new and additional to what you do now, or will it, in whole or in part, replace existing activities?

BBC: Digital preservation (by migration) is just the next stage in our analogue to digital migration, so in that sense it is a replacement

CUL: CUL's digital preservation requirement will be predominatly additional, because (i) legal deposit of digital material will be in addition to the continuing deposit of printed material, and (ii) the DSpace@Cambridge repository is expected to acquire types of digital content that would not be collected by the library if they were in printed form.

JISC: At present digital is often a supplement rather a replacement for other material so preservation represents an additional cost. However some libraries are beginning to generate space savings where they can de-accession/ suspend subscriptions to paper journals and substitute access to an electronic version

NAS: For the foreseeable future, it will be in addition to more traditional archival activities. Even if we reach a day when only digital material is received by NAS, we will still have hundreds of years worth of paper and parchment to care for. It is possible that our requirement for more shelf space for traditional records might one day reach zero, to be replaced by a requirement for more virtual space, but that day is still a very long way off. We know that there is still about 30 years worth of paper to work its way through many government organisations, and we cannot predict what will come to light from private collections in the future. The near future is likely to be one of additional expense, rather than one of savings, .as digital preservation is developed in tandem with traditional preservation,

NHM: It will complement existing paper-based activities, at least for the foreseeable future.

OU: The Open University archives print materials in both digital and non-digital formats. Non-digital formats have to be produced for students on most courses (although more online only courses are being developed) so the print form of these resources is kept for student access and students have make strong indications that any move to digital only formats would not be well received. The resources we are attempting to develop digital preservation strategies for are web-based and have no non-digital equivalent and will therefore not replace a need we already have.

PRONI: Digital preservation is a new activity for PRONI. However, it is expected that there will always be the need to preserve non-digital records. Public records transferred will shift from non-digital to digital over a period of time. The NI Civil Service is at the procurement stage of its corporate EDRMS project and there is no expected date for the transfer of electronic records yet.

ULCC: digital preservation will replace existing activities in storage of paper records. In the long term this will lead to savings, but they will not be a significant proportion of total digital preservation costs.

5.6 Will you need core or project funding, or both, to cope with digital preservation?

AHDS: Research into new methods of digital preservation, testbeds for advanced approaches will most likely require project funding support.

CCLRC: Future work will require capital expenditure of ~ £300K every 2-3 years depending on growth rates to fund technology upgrade, plus recurrent spend of ~ £350-450K per year to cover staff costs. There may also be other costs such as network upgrade costs arising later Additionally new costs will be added to adhere to upcoming standards e.g. the Open Archival Information System. Funding is unlikely to get easier.

JISC: It is likely that balance will change over time between project and service funding.

NAS: Both. Project funding will be required to develop new procedures and purchase new equipment, and core funding will be required to maintain digital material over time, as this comes as central to NAS's work as the storage and conservation of traditional records. Both – and both types of funding are likely to be hard to get.

NLS: I expect that we will push for core funding. As a legal deposit library it is essential that this is funded in the same way as the traditional preservation activities that are needed to provide access today to material created in the past. Our ICT division will be actively producing a business plan that will highlight storage needs so that we can present our case to the Scottish Executive in 2004, but this will not tackle other associated resources that will be required.

NLW: We will need both core and project funding to cope with DP. Both will prove hard to obtain.

Wellcome: Both. Project funding to get the systems in place; core funding for sustainability.

5.8 Will your digital preservation programme help you to develop any significant additional revenue stream from your digital services

AHDS: We intend to offer digital preservation services to those outside the community we are funded to support.

BBC: Certainly helps provide additional services, principally a digital archive with desktop media delivery internally and externally where commercially appropriate.

CCLRC: Any income generated will help to go towards our running costs and improving the service

NAS: Unlikely in the short term at least – we are undertaking it as an extension of what we do in a paper environment. It is possible that we might find ways of earning revenue in future (eg 'selling' any expertise we gain; providing a preservation service for others) but our first priority has to be to develop systems for ourselves and with limited resources, it is unlikely that we will have spare capacity available for a long time to come.

NHM: Certain digital materials, especially images, will continue to provide revenue.

OU: Yes, the reuse of materials is seen as a way of making a significant saving to the Open University both in terms of production cost and time and has been a policy within the Open University for the past three years. Quantifying this saving has been difficult though and is something we are presently investigating.

PRONI: Not known. The development of a new funding model could be a possibility in a 21st century digital preservation era. However, there could be difficulties in implementing any model that levied a charge depending on the amount of material transferred/preserved - PRONI, in conjunction with an organization, determines what is retained – the retention of records within organizations could be jeopardized on economic grounds.

UKDA: We do undertake a limited amount of separately costed preservation for specific projects. These are mainly in the academic sector. We could undertake 'commercial' preservation contracts and may do so in future, but would not wish this to undermine or dilute our main mission.

Wellcome: Highly unlikely. We do generate some income from our image library (http://medphoto.wellcome.ac.uk) but income generation is not the driver.

5.9 Have you assessed your staffing needs for digital preservation?

AHDS: Hiring staff with prior experience of digital preservation, as opposed to experience of some aspects of IT relevant to preservation, is difficult. The AHDS provides a significant element of training on the job in order to equip staff with the necessary skills.

CUL: Specialised staff are essential. We are planning to provide selected library staff with training in metadata skills, using external trainers to supplement existing in-house expertise.

CCLRC: This is ongoing, specialist staff are difficult to get.

JISC: will aim to build capacity in sector through funding programmes and opportunities these present.

NLS: Not beyond guesswork at this stage. Again the cataloguing resources needed to manage the legal deposit of electronic material will have been investigated, but the support needed for the digital preservation aspect will not have been properly assessed.

PRONI: The availability of suitable training courses will be essential (particularly as the traditional preservation skills have to be blended with IT).

TCD: we need specialised staff from the outset. No funding is available and this is not high on the current priority list. We are conscious of the need to begin planning and to assess our likely needs.

ULCC: We are lucky at present in that we have critical mass of staff with experience. We have proved to ourselves that this allows us to take on people without relevant skills or experience, but merely with the correct attributes, and to train them up rapidly in-house. We have only once had difficulty in attracting staff for a post. Nonetheless, we have a keen interest in the development of training courses at all levels to ensure a larger pool of skilled people, if only because it makes it less likely our staff will be poached!

6. Timing issues

6.1 Do you face any time deadline, or other issue affecting timing, on digital preservation?

BBC: We have decay issues with analogue material, and with digital videotape. We have time-related issues in keeping "access" on current technology. We would lose 3 to 5% of our archive per year if we were not transferring, and it would all become inaccessible in 20 years without transfers.

7. General gaps and priorities **[for all respondents' replies to this question see separate annex 6]**