Annex 5: responses to question 4.1

4.1 What type of digital material will you mostly be preserving?

Arts and Humanities Data Service

The AHDS preserves material deposited with us voluntarily by individuals and research groups within Higher Education, or as a condition of awards granted by the Arts and Humanities Research Board. Some material created outside Higher Education is also actively pursued for deposit by AHDS staff.

The AHDS holds electronic texts, databases, still images, moving image, audio, GIS data, Geophysics data (archaeology) metadata sets (catalogues deposited with us, as opposed to our own catalogue).

Some of this material represents digital surrogates for still and moving images, and audio recordings, transcriptions of original literary works, transcriptions of original statistical works. Some of this material represents digital resources based on, but not direct surrogates of, non-digital sources, such as collections of information taken from historical documents. Some of this material represents born digital research papers, reports, field work notes etc.

BBC

The material is created by the BBC and covers the following:

Website: www.bbc.co.uk

Audio: mainly analogue, but includes DAT tape, CD and DVD; we are just about to begin acquisition of 'born digital' material from servers.

Video: mainly analogue, but will soon need to transfer D3 digital videotape. We have made CD and DVD material during preservation.

Core business records

The British Library

The nature of the Library's national and international standing means that all the following will be part of its digital collection: research publications, journals, books, primary research material (e.g. e-manuscripts) ,records (though here there is particular scope for collaboration with other bodies, e.g. The National Archives, so that the Library is unlikely to collect records of this kind comprehensively) , the Library's own management records and web pages, external websites, e-mail newsletters, and other materials.

The scale of online and offline digital publication in the UK, and the priority given to UK publications within the British Library, means that the bulk of digital material will be British, but foreign research publications, journals, and books will also be collected.

This will be subject to proposed Legal Deposit legislation concerning British digital materials, and within the framework of the Library's collection development policy, as approved by the Board and the selection policy as formulated by the Web Archivist (See 3.4 above).

Specific projects will specialize as follows:

Domain UK: UK websites, selected by BL and captured with permission of 'publisher'.

SHERPA: unpublished research reports and papers generated by independent scholars who use BL reading rooms.

Purchased publications: possibly scanned versions of journal articles in a PDF wrapper. How we provide hot links to these articles remains a question.

Manuscripts: it is predominantly unpublished born digital material acquired by donation or purchase under the terms stated in 3.4.

Sound Archive: Born-digital sound recordings and digital transfers of born-analogue sound recordings.

Newspapers: born digital (current) newspapers, those newspapers scanned as a result of BL led Projects, older newspapers scanned by other publishers, and websites which contain significant amounts of news content (naturally after permission has been received)

Cambridge University Library

For digital material originating within Cambridge University, all types of material (publications, learning objects, datasets, digitized library collections, admin records, etc.) are potentially included.

Material of external origin will largely be restricted to items received by legal deposit and, to a lesser degree, purchased digital material.

Both born-digital and digitized copies will be included.

Council for the Central Laboratory of the Research Councils

Scientific data in a variety of formats: Monte carlo, HDF, flat files ASCII etc etc. most of which will be born digital

National Archives of Scotland

Deposited with us under the 1937 and 1948 Acts:

Public records (mainly from Scottish government and agencies. The Scottish Executive is developing an electronic records management system and aims to be as fully electronic as possible by 2005)

Public registers (eg the Register of Sasines , Scotland's land register. Not yet being created digitally, but it is only a matter of time)

Possibly some websites or parts of websites (eg of particularly significant organisations like the Scottish Parliament)

Court records (some areas are already exploring imaging paper records and disposing of the paper, though this has not yet begun on a large scale)

Deposited with us under other agreements:

Records created by private individuals or organisations (we expect these to be increasingly electronic)

Records created by NAS:

1. Our own administrative records – it is only a matter of time before we too will need to develop an ERM system, at which point, the electronic version of a record will be the master copy and will have to be preserved.

2. Digital surrogates – NAS is running / is a partner in many projects which are digitising material held in our archives. SCAN, which has digitised over 520,000 Scottish wills and testaments dating from 1500 to 1901, is the biggest of these.

National electronic Library for Health

Research material, created by third parties, all in web format. All acquired by us to inform our stakeholders. All "born digital".

Natural History Museum

There will be a mix of all materials, including public records, research material, original art and manuscripts along with rare publications. Virtually all of it will be created by this organization and will be a mix of "born digital" and "preservation digitization" material. In future, we may become a repository for digital data from similar organizations elsewhere in the world as the result of reciprocal agreements.

National Library of Scotland

i) Scottish websites in the future. ii) Voluntary deposit and legal deposit of electronic material that could be born digital, or supplied to us as digital copies and in a variety of formats that reflect the personal or institutional expressions of the creators. iii) The result of an internal digitisation programme, with some further related files being produced externally (i.e. OCRd text using in-house TIFFs). It is the first two categories that are expected to have the highest demands and form the bulk of digital preservation. I am uncertain what plans are to preserve our own management records beyond short-term business continuity.

National Library of Wales

Digitised (from the Library's Digitisation Programme) and born digital e.g.

- Digital publications received through voluntary deposit agreements e.g. CD ROMs, disks
- E-journals, e-books
- Databases
- Disks that accompany printed material
- Online publications received via e-mail, etc
- Disks that form part of archival collections
- Electronic records deposited by institutions as part of their archives
- Websites
- Time based materials e.g. sound and video

Online Computer Library Center

In terms of the Digital Archive service, our intent is to preserve whatever digital materials our members need preserved. However, at this early stage, there is no clear view of what these materials will comprise. In terms of

OCLC's internal digital preservation needs, some of the content currently being preserved includes OCLC's Annual Review of Research, OCLC Annual Reports, OCLC Newsletters, and past news releases. This content includes both born-digital and digitized material.

Open University

Born digital: Teaching materials, including published materials that are accessed by students in paper form but created and stored digitally (we also archive the paper copies but only the first and last presentation of a course) Web sites, email conferencing; images. These materials are generates in-house

Digital copy: the library's Jennie Lee Archives. Papers from Jenni Lee, Walter Perry and Betty Boothroyd and our research collection. Thesis collection

The National Archives: UK Government records, archived government websites

Oxford University Library Services:

1. Research publications (traditionally defined very broadly by the Bodleian Library).

2. Records: Oxford University digital records (in collaboration with OU Archives); public records of organizations which deposit material; business process records of the Library.

3. Websites, especially those produced within the Oxford domain (in collaboration with OUCS)

4. eManuscripts (eg author's papers, email etc in digital form)

5. Digitised materials, eg images derived from analogue originals within the Oxford collections.

The Public Record Office of Northern Ireland

It will comprise public records, mainly expected to be office documents but could include websites. PRONI is also the repository for private records which could be in any form, journals, books, publications, papers etc.

It will also comprise material created by the PRONI digitization programme.

Trinity College Library, Dublin: legal deposit, academic research, library records, surrogate copies of collection material. A combination of digital copy material and a large proportion of 'born digital' material.

UK Data Archive

The UKDA holds a variety of data types which are made available for academic research and teaching. These are created by academics, government departments and agencies and commercial companies. They include databases and associated metadata. These can take the form of statistical databases, relational databases, text files, image files, audio files. All those are 'born' digital. In addition, it preserves digital copies of mainly historical documents, mainly in image formats.

ULCC

Research publications/journals, books, primary (i.e. unpublished) research material, public records and other public sector record material held as an official repository, own organisation's management records etc), websites, other material. Created by own organisation, the product of own research programmes, acquired for research or other purposes, created by others, created by others and passed to you for reasons other than statutory deposit, the result of a digitization programme. Both 'born digital' and digital copy.

Wellcome

a) Images

b) Archives (documents, spreadsheets, emails)

c) Web sites

d) From the Trust's perspective – documents, policy statements, grant applications