

DRAFT August 2004

Directory of Digital Preservation Repositories and Services in the UK

<i>PART 1: DPC MEMBERS AND RELATED PUBLIC SECTOR SERVICES</i>	
<i>CATEGORIES:</i>	
A	Archives
B	Data Services
C	Deposit Libraries
D	Libraries
E	Research Centres
F	Research Councils
<i>PART 2: PRIVATE SECTOR SERVICES</i>	
<i>CATEGORIES:</i>	
A	Data services
B	Consultancy support
C	Development services and support
<i>Entries are listed alphabetically by category</i>	

Scope of the Directory

The DPC Directory of Digital Preservation Repositories and Services is designed to help organisations needing information and services relating to their digital preservation storage needs and also wanting to know about resources which exist to provide for digital storage and preservation. The original intention was to list and give detail of digital preservation repositories, which was understood to mean a repository storage service designed and offered in accordance with established digital preservation principles. This approach clearly excluded a lot of potentially useful services and information. It excluded several public sector operations which store, or arrange storage, (for example, of e-journals or other academic information) with long term preservation in mind, even though they do not offer any commercially available storage service to others. It excluded basic data storage, or a storage medium, or software, such as is offered by a number of companies both large and small, but which cannot be shown to have been designed with close attention to long-term digital preservation principles.

It soon also became clear that there are related services available, from both the public and private sectors, which it would be sensible to list. For example, if you are looking to preserve and store digital material in an obsolete format, you might find the service offered by PRONOM to be of use as a part of the process, so that is included. If you are trying to select a commercial digital storage provider, you might want to consult a company which had worked on substantial and relevant digital preservation projects and could help you assess the best choice to make. So we have further broadened the scope to include services such as these.

The unifying element for all entries remains that an understanding of and commitment to long-term digital preservation had to be inherent and of the essence if such services, products and offers were to be included.

Some instances may seem borderline for inclusion; for example, organisations which act more as gateways to other stores of information, than as stores themselves. We have included some services which fall into this category. This is both because it clarifies what it is that they do, and because the line between such services and more storage-based ones can be quite fine. Examples of such a service are MIMAS, or EDINA.

We have not limited entries to UK based institutions. The rationale for this is, chiefly, that if the services are relevant to this register, and available within the UK, then inclusion is justified. We have, though, limited the register to anglophone services, at least for the time being.

If the need for selectivity seems to have resulted in wrong decisions, either about inclusion or exclusion, the DPC is happy to consider representations and changes.

Where we have been able to, we have given a contact name and details. Otherwise we have given the general contact point, normally an email address taken from the website details.

Where organisations listed are DPC members, we have indicated this.

We are not claiming that the directory is comprehensive. It will need to be maintained and updated if it is to go on being useful. In particular, it will be obvious that, in this first edition, we have only scratched the surface of the private sector possibilities. So we welcome suggestions both from DPC members and from others as to products, companies and organisations which should be included. Please send any suggestions,

with an indication why you think they are relevant, and as much detail as you can, to the DPC Secretary, Maggie Jones: maggie@dpconline.org

Finally, the DPC is not offering any recommendation or endorsement whatsoever for any products or services in this directory. Nor is it offering any guarantee of the accuracy or currency of information. For the most part information has been taken at the time of compiling each entry, from published information about the organisations and services included. Most commonly it has been taken from their websites. It has usually been edited and often shortened. It will not reflect any changes since the time it was collected or last updated. It is offered in good faith, but it is up to users to check accuracy and currency, and to make their own decisions before committing to any action arising from any information contained in this directory. The DPC accepts no responsibility or liability for any such actions or for anything arising from them.

Index of contents

Entries are listed under the name they are generally known by, so MIMAS, rather than Manchester Information and Associated Services, LOCKSS rather than Lots of Copies Keep Stuff Safe, except where this might be unclear – so, Arts and Humanities Research Board rather than AHRB, Natural Environment Research Council rather than NERC etc.

Part one: Public sector and not-for-profit services	
A Archives	
BBC	5
Internet Archive	6
The National Archives	7
National Archives of Scotland	8
Public Record Office, Northern Ireland	9
UK Web Archiving Consortium	10
B Data Services	
Arts and Humanities Data Service	11
Atlas Datastore	12
EDINA	13
MIMAS	14
Online Computer Library Center	15
PRONOM	16
UK Data Archive	17
University of London Computer Centre	18
C Deposit Libraries	
British Library	19
Cambridge University Library	20
National Library of Scotland	21
National Library of Wales/ Llyfrgell Genedlaethol Cymru	22
Oxford University Library Services	23
Trinity College Library, Dublin	24
D Libraries	
e-Archives	25
e Prints UK	26
JSTOR	27
LOCKSS	28
E Research Centres	
Digital Curation Centre	29
F Research Councils	
Arts and Humanities Research Board	30
Council for the Central Laboratory of the Research Councils	31
Economic and Social Research Council/Economic and Social Data Service	32
Medical Research Council	33
Natural Environment Research Council	34

Part two: Private sector services	
A	Data services
B	Consultancy support
	<i>Audata</i>
	<i>Tessella</i>
C	Development services and support
	<i>Magus</i>

CATEGORY A: ARCHIVES

CATEGORY: Archives	DPC member
BRITISH BROADCASTING CORPORATION (BBC)	
<p>Location: London</p> <p><i>DPC contact:</i> Cathy Smith (New Media Archivist), tel: 020 8225 9958, e-mail: cathy.smith@bbc.co.uk</p> <p><i>URL:</i> www.bbc.co.uk http://presto.joanneum.ac.at</p>	
<p>Scope of business: About 70% of the BBC archive is at risk and requires digitization as a method of preservation. The BBC is investing £60 million in this project over the decade 2000-2009. Much of the problem concerns material which is at present analogue, and requires conversion. In the next decade the BBC will have to face the continued preservation of large amounts of digital material. The BBC also has a growing volume of ‘born-digital’ preservation issues, eg web-archiving and interactive TV</p>	
<p>Service offered: the BBC does not offer a digital archiving service to others</p>	
Type of material held	
<p>The material is created by the BBC and covers the following:</p> <p>Website: www.bbc.co.uk; audio: mainly analogue, but includes DAT tape, CD and DVD; just about to begin acquisition of ‘born digital’ material from servers;</p> <p>Video: mainly analogue, but will soon need to transfer D3 digital videotape. Have made CD and DVD material during preservation.</p> <p>Core business records</p>	
Volume of material held/planned	
<p>BBC has surveyed web archive requirements, and has detailed knowledge of analogue and digital audiovisual material and its preservation requirements. It will be capturing 3.5 TB of data over a three-year period, a selection of the BBC’s web output. This will be stored on disk, backed up to digital linear tapes.</p> <p>The 10-year audiovisual preservation project will produce about 500,000 hours of digital material of various types – approximately 40 Peta bytes of data.</p> <p>Also BBC is acquiring material at a rate of about 20k hours per year, a growth rate of about 10% per year.</p> <p>[For further information see PRESTO project – http://presto.joanneum.ac.at]</p>	

CATEGORY: Archives	
INTERNET ARCHIVE (IA)	
Location: San Francisco, CA contact: URL: www.archive.org	
Scope of business: Set up in 1996, as a non-profit making public organisation. The mission is to help preserve digital artifacts and to create an internet library of websites for researchers, historians and scholars. The archive collects publicly accessible web pages and offers free access to them. It is a web-crawler based internet archive, operated by Alexa Internet. It respects privacy policies and 'robot exclusion' in its selection. The archive is updated 'every few months', updates taking 6-12 months to reach 'The Wayback Machine', (http://www.archive.org/web/web.php) which is the access method. Storage is on basic PCs, currently about 800, each storing around one terabyte on ATA disks, a cheap and low maintenance system running on Linux. In the process (June 04) of moving to a new system – see www.petabox.org . Funding relies heavily on institutional supporters and on donations.	
Service offered: Free access, though IA says some programming skills may be needed. IA 'actively seeks' donations of digital material for preservation, responding to proposals.	
Type of material held	
Publicly accessible web pages	
Volume of material held/planned	
June 04, over one petabyte and growing at about 20 terabytes per month.	

CATEGORY: Archives	DPC member
THE NATIONAL ARCHIVES (TNA)	
<p>Location: Kew, Surrey and over 200 places of deposit across the UK</p> <p><i>DPC contact:</i> David Ryan, Head of Digital Preservation; e-mail: david.ryan@nationalarchives.gov.uk</p> <p><i>URL:</i> http://www.nationalarchives.gov.uk/preservation/digitalarchive/</p>	
<p>Scope of business:</p> <p>The National Archives holds state and central court documents from the Domesday Book to the present, provides public access at its reading rooms in Kew and at the Family Records Centre in central London for anyone and provides access to a growing collection of documents online. It is also the UK's central advisory body on archives and manuscripts relating to British history. It works with central government to help select documents now, which will be opened to the public in 30 years time and also provides advice for central and local government. It contracts with the ULCC (qv) for the management, storage and preservation of government datasets. It gives high priority to digital preservation.</p>	
<p>Service offered: TNA has an acquisitions policy and operates a selection process for the records which it accepts. It does not offer any commercial management or storage service.</p>	
Type of material held	
Digital material held consists of UK Government records and archived government websites	
Volume of material held/planned	
<p>TNA Digital Archive and National Digital Archive of Datasets hold 184.5 Gbytes at present. Includes datasets, websites, CD-ROM publications, office documents, digital video, 44,374 individual files in about 150 different file formats. Anticipated new accessions for the coming year about 800 Gbytes. Substantial rates of increase expected thereafter. Current actual storage capacity of the TNA Digital Archive is 4 Tbytes, split between master and open systems and offsite backup, 1.5 Tbytes being currently given over to master record storage.</p>	

CATEGORY: Archives	DPC member
NATIONAL ARCHIVES OF SCOTLAND (NAS)	
Location: Edinburgh <i>DPC contact:</i> Laura Mitchell, tel: 0131 535 1412; e-mail: laura.mitchell@nas.gov.uk <i>URL:</i> http://www.nas.gov.uk	
<p>Scope of business: As the repository for Scotland's national archives, NAS has responsibilities for collecting, preserving and giving access to, archive material of national importance. Original archive material is increasingly likely to come, and be stored and managed, in digital form. Also one of the main ways of giving access to delicate hard copy originals is through digitisation. There is therefore a need to preserve long term both born-digital and digital surrogate material.</p>	
<p>Service offered: NAS is responsible for the storage and preservation of its own collections; it does not offer any digital storage or other services to other organisations.</p>	
Type of material held	
<p><i>Deposited under the 1937 and 1948 Public Records Acts:</i> Public records, mainly from Scottish government and agencies. Public registers (eg the Register of Sasines , Scotland's land register). Not yet in digital form, but will be. 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)</p> <p><i>Deposited under other agreements:</i> Records created by private individuals or organisations (these will, increasingly, be in electronic formats)</p> <p><i>Records created by NAS:</i> NAS's own administrative records. Digitised copies, known as surrogates, of material held in traditional formats in the archives. SCAN, which has digitised over 520,000 Scottish wills and testaments dating from 1500 to 1901, is the biggest of these.</p>	
Volume of material held/planned	
<p>Approx 8.15 gigabytes of material from the Scottish Executive and the Scottish Parliament. Also about 216 gigabytes of surrogate digital images, currently kept on CD. The SCAN project has so far produced about 1.4 Tb of surrogate digital material held on line and 26 Tb on tape. Eventually all court, government and Scottish Parliament records will come in digital form. Current rate of accession about 1800m of paper a year, which will continue in digital formats.</p>	

CATEGORY: Archives	DPC member
THE PUBLIC RECORD OFFICE OF NORTHERN IRELAND (PRONI)	
Location: <u>Belfast</u> , Armagh, Ballymena, Blacklion, Londonderry DPC contact: Hugh Campbell, e-mail: Hugh.Campbell@dcalni.gov.uk URL: www.proni.gov.uk	
<p>Scope of business: PRONI is the official place of deposit for public records in Northern Ireland. It accepts both official and private records. The records fall into three general categories:</p> <p>Records of Government Departments which in many cases go back to the early nineteenth century; Records of courts of law, local authorities and other non-departmental public bodies; Records deposited by private individuals, churches, businesses and institutions.</p> <p>PRONI to select and preserve those records which provide a legal or historical record of the past and to make these available to the public for consultation and research, regardless of format.</p> <p>It recognises the importance of electronic records management and of digital preservation.</p>	
<p>Service offered: PRONI operates a selection process for the records, both official and private, which it accepts. It does not offer any commercial management or storage service.</p>	
Type of material held	
Digital materials include 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. PRONI also holds images created by digitisation projects.	
Volume of material held/planned	
<p>At present 700Gbytes, comprising 69,000 surrogate digital images. This is stored on a dedicated server. Expected growth in the short term is likely to result from future digitisation projects.</p> <p>PRONI is implementing an Electronic Document and Records Management System (EDRMS) for its administrative records. When this is implemented in Autumn 2004, records will be created and stored electronically and will have to be preserved. The volume of electronic records to be preserved will increase as other public bodies in Northern Ireland implement EDRMS.</p>	

CATEGORY: Archives	
UK WEB ARCHIVING CONSORTIUM (UKWAC)	
<p>Location: partnership of British Library (qv), Joint Information Systems Committee of the Higher and Further Education Councils, The National Archives (qv), the National Library of Wales (qv), the National Library of Scotland (qv) and the Wellcome Trust.</p> <p><i>contact:</i></p> <p>URL: www.webarchive.org.uk</p>	
<p>Scope of business: Launched June 2004, to go live early 2005, UKWAC is an experimental system, initially for two years, for archiving selected key UK websites, with an initial target of 6,000 sites. With permission from rights holders, each partner will capture content relevant to its own remit. Infrastructure costs are shared. The archive uses a development of HTTrack, an open source web crawler using software developed for a similar consortium archiving Australian websites called PANDAS (http://pandora.nla.gov.au/index/html). Adaptation of the software, plus hardware and technical support, are provided by Magus Research (qv)</p>	
<p>Service offered: archiving of, and online access to, selected UK websites, according to selection decisions made by the partners. UKWAC does not offer any commercial management or storage service.</p>	
Type of material held	
UK websites	
Volume of material held/planned	
Initially 6000 websites	

CATEGORY B: DATA SERVICES

CATEGORY: Data Services		DPC member
ARTS AND HUMANITIES DATA SERVICE (AHDS)		
<p>Location: <u>London</u>, Oxford, York, Colchester (Essex), Glasgow, Farnham (Surrey)</p> <p>DPC contact: Hamish James, +44(0)20 7928 7371; hamish.james@ahds.ac.uk</p> <p>URL: http://ahds.ac.uk</p>		
<p>Scope of business: The AHDS collects, preserves and promotes electronic resources in the arts and humanities. It is funded by the Joint Information Systems Committee (JISC) and the Arts and Humanities Research Board (AHRB). It covers five subject areas: archaeology; history; visual arts; literature, languages and linguistics; performing arts. Digital preservation has been a core activity of the AHDS since its establishment in 1996. AHDS centres collect, preserve, catalogue, and distribute digital resources which are relevant to their subject areas, facilitate good practice in their creation and use, and offer some user services.</p>		
<p>Service offered: preserves material deposited in accordance with policies; no commercial service offered</p>		
Type of material held		
<p>The AHDS is a distributed service and preserves material deposited 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.</p> <p>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).</p> <p>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 represents digital resources based on, but not direct surrogates of, non-digital sources, such as collections of information taken from historical documents. Some also represents born digital research papers, reports, field work notes etc.</p>		
Volume of material held/planned		
<p>Currently approximately 1TB of data, comprising about 4,000 distinct collections, will rise to 4TB by the end of 2004.</p> <p>Anticipate sharply rising volumes of data from a moderately rising number of depositing projects. Initially planning a capacity of 10Tbytes of data for the new digital repository</p>		

CATEGORY: Data Services	
ATLAS DATASTORE (ADS)	
See entry under Category F, Research Councils, Council for the Central Laboratory of the Research Councils (CCLRC)	

CATEGORY: Data Services	
EDINA	
Location: University of Edinburgh contact: edina@ed.ac.uk URL: http://edina.ac.uk/	
Scope of business: a provider of specialist data services, based at Edinburgh University Data Library, EDINA is a JISC-funded national datacentre. It provides to the UK tertiary education and research community networked access to a library of data, information and research resources. All EDINA services are available free of charge to members of these institutions, though with subscription and registration for most services. Subjects include health, agriculture, arts, humanities, social sciences, engineering, physical sciences and general reference topics.	
Service offered: access to a wide range of materials – see http://edina.ac.uk/sitemap.shtml	
Type of material held	
EDINA is an access service; it does not hold or store its own materials	
Volume of material held/planned	
Not applicable	

CATEGORY: Data services	
MANCHESTER INFORMATION AND ASSOCIATED SERVICES (MIMAS)	
Location: run by Manchester Computing, at the University of Manchester contact: URL: www.mimas.ac.uk	
Scope of business: A national data centre, supported by JISC, to provide dataset services for UK HE, FE and research through networked access. Available to all FE and HE institutions mostly free of charge though some services (eg JSTOR, qv) require subscription. MIMAS offers access to archive and publications catalogues from over 70 institutions in the UK; catalogues of 26 major university libraries; plus many databases including, for example, and digital publications such as e-journals. For fuller details http://www.mimas.ac.uk/reports/annual/year0203/mimas-services.html . Also provides dataset services in collaboration with others eg British Library (qv), UK Data Archive (qv) and others.	
Service offered: access to a wide range of materials – see http://www.mimas.ac.uk/reports/annual/year0203/mimas-services.html .	
Type of material held	
MIMAS is almost entirely an access service; it does not hold its own materials	
Volume of material held/planned	
Not applicable	

CATEGORY: Data services	DPC member
ONLINE COMPUTER LIBRARY CENTER (OCLC)	
Location: Dublin, Ohio, USA DPC contact: Liz Bishoff, Vice President, Digital Collections and Metadata Services +1-800-848-5878 URL: http://www.oclc.org/digitalarchive/default.htm	
Scope of business: OCLC is a nonprofit, membership, computer library service and research organization whose aim is to further access to the world's information and reducing information costs. More than 45,000 libraries worldwide use and support OCLC services to locate, acquire, catalogue, lend and preserve library materials. OCLC operates a digital archive service to all its members, with an emphasis on digital preservation of the material. The service is in two forms: web archiving for item-by-item harvesting and submission of web pages and web-based documents, or batch archiving to submit collections on various storage media for ingest and automated metadata creation at OCLC. These are then made accessible to users through a range of means.	
Service offered: OCLC members can place material in the digital archive service as described above. There is no commercial service to others.	
Type of material held	
Any of the library materials held by OCLC members	
Volume of material held/planned	
No information	

CATEGORY: Data services	
PRONOM	
Location: The National Archives, Kew, London <i>contact:</i> pronom@nationalarchives.gov.uk <i>URL:</i> http://www.nationalarchives.gov.uk/pronom/	
Scope of business: PRONOM is an online file format registry, created and maintained by The National Archives in Kew, London (qv). It is a free resource for technical information about the file formats used to store electronic records, and the software products that are required to create, render, or migrate these formats. It is constantly updated, relying chiefly on input of information from others about the software products of which it contains details. Details are available from the website on how to submit details of relevant products for inclusion.	
Service offered: free access online to information on mainly obsolete software products, to help enable their migration or other steps to ensure preservation of material	
Type of material held	
Information on software products; access to PRONOM content is via a search facility	
Volume of material held/planned	
Not relevant	

CATEGORY: Data services	DPC member
UNITED KINGDOM DATA ARCHIVE (UKDA)	
Location: University of Essex DPC contact: K. Schürer; schurer@essex.ac.uk URL for additional information: www.data-archive.ac.uk www.esds.ac.uk	
<p>Scope of business: UKDA is a centre of expertise in data acquisition, preservation, dissemination and promotion. It curates a large collection of digital data in the social sciences and humanities. It provides resource discovery and support for secondary use of quantitative and qualitative data in research, teaching and learning. It provides preservation services for other data organisations and facilitates international data exchange. UKDA collects material according to its collections development policy, though other data may also be accessioned if of exceptional merit and capable of being handled within current resources.</p> <p>It also houses the Economic and Social Data Service, a joint venture with the Economic and Social Research Council (see ESRC entry under category F, Research Councils)</p>	
<p>Service offered: subject to its collection policy and other objectives the data archive accepts digital material for storage and access on a commercial basis.</p>	
Type of material held	
<p>A variety of data types for academic research and teaching. These are created by academics, government departments and agencies and commercial companies. They include databases and associated metadata. They can take the form of statistical databases, relational databases, text files, image files, audio files. All those are 'born' digital. In addition, the data archive preserves digital copies of mainly historical documents, mainly in image formats.</p>	
Volume of material held/planned	
<p>Current collection = c. 3 Tbytes; facilities to increase core collection to c.10 Tbytes over next 4 years.</p>	

CATEGORY: Data services	DPC member
UNIVERSITY OF LONDON COMPUTER CENTRE (ULCC)	
Location: London DPC contact: Kevin Ashley tel: 020 7692 1338, e-mail: K.Ashley@ulcc.ac.uk URL: http://www.ulcc.ac.uk http://ndad.ulcc.ac.uk/	
Scope of business: The National Data Repository at ULCC provides a network-accessible digital archive and filestore, based on a robotic tape system with access to up to 300 Terabytes of data. held on high-speed digital tape, and brought online automatically whenever it is required. Data is automatically migrated to new media as required. ULCC provides services for the British Library's Initiatives for Access programme and hosts the National Digital Archive of Datasets for The National Archives. Also services and consultancy to a wide range of organisations in the public and private sectors, and a major role in developing digital preservation.	
Service offered: The service is available to customers on a commercial basis. Costs are dependent on a number of factors other than data volume, with frequency of access and level of security required also being involved.	
Type of material held	
Research publications/journals, books, primary (i.e. unpublished) research material, public records and other public sector record material held as an official repository, websites, other material. The product of own organisation's research programmes, material acquired for research or other purposes, created by others, created by others and passed to ULCC for reasons other than statutory deposit, the result of digitization programmes. Both 'born digital' and digital copy.	
Volume of material held/planned	
Currently small number of Tbytes; millions of separate objects; over 200 formats. Cannot make assessments at present of likely future growth	

CATEGORY C: UK DEPOSIT LIBRARIES

CATEGORY: Deposit Library	DPC member
THE BRITISH LIBRARY (BL)	
<p>Location: <u>London</u> and Boston Spa (Yorks)</p> <p>DPC contact: Helen Shenton, tel: 020 7412 7594, e-mail: Helen.Shenton@bl.uk</p> <p>URL: http://www.bl.uk/</p>	
<p>Scope of business: <i>New Strategic Directions 2001</i>, the Library's vision for the following five years, lists its key responsibilities. They include 'ensuring the comprehensive coverage, recording and preservation of the UK national published archive'. This incorporates a growing proportion of digital materials; therefore the Library has developed a dedicated policy and set of strategies for digital preservation. The BL is a legal deposit library and new legal deposit legislation requires it to preserve digital publications by law. Preservation of digital material is a high priority for the British Library, which is also a partner in UKWAC (qv)</p>	
<p>Service offered: The BL collects and holds large volumes of digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.</p>	
Type of material held	
<p>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 are also collected. This is subject to new Legal Deposit legislation concerning British digital materials, and within the framework of the Library's collection development policy, as approved by the Board and its selection policy.</p>	
Volume of material held/planned	
<p>The estimates provided are based on a partial survey so they should be taken as indicative only. Items received under voluntary deposit: c.100,000 items, c.1Tb. Digitised material: at least 10Tb</p> <p>Expected by 2005: Legal Deposit: hand-held (i.e. CD-ROM): > 1200 monographs, >1000 serial titles with over 3700 serial issues / parts. Purely digital: >1300 e-monographs; >7000 serial titles with nearly 200000 serial issues/parts.</p> <p>Digitised: in excess of 30Tbytes</p> <p>Audio: At April 2003, estimated that Sound Archive holdings are 622 terabytes of data (622,259,225Mb - calculation factor was that a 70 min CD uses 650Mb of digital space) using our current preservation standard as the benchmark. Estimated 25% of this is born-digital. Current rate of growth estimated at 23 terabytes per annum, 99% of this being born digital.</p>	

CATEGORY: Deposit library	
CAMBRIDGE UNIVERSITY LIBRARY (CUL)	
Location: Cambridge <i>contact:</i> Peter Morgan, Project Director, DSpace@Cambridge , tel: 01223 333130, e-mail: pbm2@cam.ac.uk <i>URL:</i> www.lib.cam.ac.uk/ , www.lib.cam.ac.uk/dspace/	
Scope of business: As a legal deposit library with research collections of international importance, CUL is committed to the long-term storage and preservation of information in all media; and it is now developing a commitment to curation of the University's intellectual output in digital formats.	
Service offered: CUL collects and holds digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.	
Type of material held	
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 are included.	
Volume of material held/planned	
Currently storing approx. 400 GB of digital material in long-term storage, mostly in-house digitization projects. Growth depends on the extent of CUL's commitments with regard to legal deposit intake and the uptake of DSpace services by the University.	

CATEGORY: Deposit library	DPC member
NATIONAL LIBRARY OF SCOTLAND (NLS)	
Location: Edinburgh DPC contact: Rab Jackson tel: 0131 226 4531, e-mail: r.jackson@nls.uk URL: http://www.nls.uk	
Scope of business: NLS is a legal deposit library and therefore involved in legal deposit and the long term storage and preservation of large quantities of digitised material. The introduction of new legal deposit legislation also requires the Library to preserve digital publications by law. Preservation of digital material is a high priority.	
Service offered: The NLS collects and holds digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.	
Type of material held	
<p>i) Scottish websites in the future.</p> <p>ii) Voluntary deposit and legal deposit of electronic material that could be born digital, or supplied as digital copies and in a variety of formats that reflect the personal or institutional expressions of the creators.</p> <p>iii) The result of an internal digitisation programme, with some further related files being produced externally (i.e. OCRd text using in-house TIFFs).</p> <p>The first two categories are expected to have the highest demands and form the bulk of digital preservation needs.</p>	
Volume of material held/planned	
At present only surveys of in-house digitised content, not legal deposit of electronic material. c. 4 Terabytes on hard disk backed up onto tape. A more detailed survey is being undertaken in 2004.	

CATEGORY: Deposit library	DPC member
NATIONAL LIBRARY OF WALES/ LLYFRGELL GENEDLAETHOL CYMRU	
Location: Aberystwyth DPC contact: : Mared Owen tel: 01970 632876, e-mail: mared.owen@llgc.org.uk URL: http://www.llgc.org.uk	
Scope of business: The National Library of Wales is the memory of the Welsh nation; it is a legal deposit library. Traditionally it has collected, preserved and provided access to a wide variety of formats such as books, periodicals, newspapers, manuscripts and archives, maps, paintings, drawings and prints, photographs, sound and moving images. During the last few years electronic media have accounted for an increasing percentage of the material that the Library receives, and we now face the enormous challenge of preserving and protecting the digital memory of Wales, in accordance with the new legislation governing legal deposit of digital materials.	
Service offered: The NLW collects and holds digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.	
Type of material held	
This includes both digitised material (from the Library's Digitisation Programme) and born digital e.g. <ul style="list-style-type: none"> • 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 	
Volume of material held/planned	
Impossible to estimate the volume at present, without knowing quantity of material under the new legal deposit legislation.	

CATEGORY: Deposit library	
OXFORD UNIVERSITY LIBRARY SERVICES (OULS)	
Location: Oxford <i>contact:</i> Richard Ovenden, Keeper of Special Collections, tel: 01865 277158, e-mail: Richard.ovenden@ouls.ox.ac.uk <i>URL:</i> www.lib.ox.ac.uk/	
Scope of business: Oxford University Library Services, which includes the Bodleian Library at its centre, is a legal deposit library. Its purpose is to collect, preserve, and make available information for the scholarly community in the University of Oxford, and to the wider world of scholarship. It has engaged in the world of digital information from the earliest days both as a creator and a consumer, and recognizes that ensuring long-term accessibility of both categories of digital information is an activity critical to its mission both now and in the future. It must also now operate in accordance with new legal deposit legislation for digital materials.	
Service offered: OULS collects and holds digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.	
Type of material held	
<ol style="list-style-type: none"> 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. 	
Volume of material held/planned	
No information available at present	

CATEGORY: Deposit library	DPC member
TRINITY COLLEGE LIBRARY, DUBLIN (TCD)	
<p>Location: Dublin</p> <p><i>DPC contact:</i> Susie Bioletti, Keeper Of Preservation And Conservation, Trinity College Library; tel: 00 353 1 6082203, e-mail: Susie.bioletti@tcd.ie</p> <p><i>URL:</i> http://www.tcd.ie/library/</p>	
<p>Scope of business: The Library of Trinity College is the largest research library in Ireland. In addition to the purchases and donations of almost four centuries, the Library has the right to legal deposit of British and Irish publications. As a Legal Deposit Library TCD library has a lead role as a repository for Irish electronic collections and has responsibility for preserving this digital material</p>	
<p>Service offered: TCD collects and holds digital material in accordance with legal requirements and its collection policies. No commercial service for other material is offered.</p>	
Type of material held	
<p>Legal deposit, academic research, library records, surrogate copies of collection material. Digital holdings consist of a combination of digital copy material and a large proportion of 'born digital' material.</p>	
Volume of material held/planned	
<p>A survey exercise is planned. Growth is unpredictable at this stage, but likely to be rapid if Irish copyright material is deposited</p>	

CATEGORY D: LIBRARIES

CATEGORY: Libraries	
Electronic Archiving Initiative (e-Archive)	
Location: Princeton, NJ Contact: Eileen Fenton ecfenton@ithaka.org URL: www.ithaka.org	
<p>Scope of business: e-Archive aims to extend the JSTOR (qv) service for digitised journals to encompass electronic journals. It is a not-for-profit organisation sponsored by Ithaka (www.ithaka.org) now working (June 04) to</p> <ol style="list-style-type: none"> 1. define the archival service to be delivered to participating publishers and libraries; 2. formulate an economic model to support E-Archive's long term sustainability; 3. develop relationships with publishers who will entrust their content to E-Archive's care; 4. design and construct the appropriate technological infrastructure; 5. design and build a production infrastructure which supports the required content processing procedures; 6. conduct research into the financial impact that an electronic archive may have on libraries and publishers 	
Service offered: not yet available, but will offer access to e-journals	
Type of material held	
e-journals	
Volume of material held/planned	
No information yet	

CATEGORY: Libraries
e-PRINTS UK
<p>Location: UKOLN and King's College London</p> <p>contact: m.guy@ukoln.ac.uk for repositories information</p> <p>URL: http://www.rdn.ac.uk/projects/eprints-uk/;</p>
<p>Scope of business: the e-Prints UK project is developing a series of services through which the HE and FE community can access e-print papers, held, managed and made available by a group of compliant Open Archive repositories, 26 in January 04, which are listed at www.rdn.ac.uk/projects/eprints-uk/repositories/, particularly those provided by UK universities and colleges. OCLC will provide an automatic subject-classification Web service. The project is run by the Resource Discovery Network, a JISC service with ESRC and AHRB support – www.rdn.ac.uk. RDN is a cooperative network of over 70 research and educational organisations.</p>
<p>Service offered: Further repositories are invited to join the e-Prints UK list. There is no commercial data storage or management service offered.</p>
Type of material held
e-prints on all subject areas covered by the member repositories (it is they who hold the information)
Volume of material held/planned
13,000 e-prints at September 03

CATEGORY: Libraries	
JSTOR	
Location: JSTOR Main Office, New York, JSTOR Production, Ann Arbor, Michigan Contact: http://www.jstor.org/servlets/FirstContact URL: www.jstor.org	
Scope of business: JSTOR is a not-for-profit organisation which aims to provide access to digitised journal articles through a large digitisation, storage and access programme, for the benefit of the scholarly community. It has over 2000 participating institutions in the USA and across the world, 230 participating publishers and almost 400 journals. Its users access the digitised archive through payment of a licence fee. It holds, as of May 2004, over 2.5m articles online, amounting to over 15m pages. It undertakes the conversion of the backfile archive of scholarly journal literature, digitising loaned, donated or purchased material. It is committed to providing the entire back run of the journals it selects. JSTOR is solely concerned with digitisation of journals; JSTOR is also developing its Electronic Archiving Initiative (known as e-Archive) to explore the archiving of born-digital journal material – see separate entry.	
Service offered: Access to the journals held by JSTOR, which is by licence. It does not offer any commercial data storage or repository service to others.	
Type of material held	
Digitised journals, a wide range, both multidisciplinary and discipline-specific	
Volume of material held/planned	
See above	

CATEGORY: Libraries	
LOCKSS	
Location: Stanford University, Cal. USA contact: Vicky Reich vreich@stanford.edu URL: http://lockss.stanford.edu/index.html	
<p>Scope of business: A web publishing and access system, LOCKSS (Lots Of Copies Keep Stuff Safe) allows libraries to safeguard the digital journals they subscribe to. It creates low-cost, persistent digital "caches" of authoritative versions of http-delivered content, by using a web-crawler. These are held on the library's own computers using LOCKSS software – 'distributed repository model technology'. All file formats delivered through HTTP are included (html, jpg, gif, wav, pdf, etc.). The LOCKSS software is free to participating institutions; it enables them to collect, store, preserve and archive authorized content locally. It runs on inexpensive hardware and requires little technical administration, and is distributed as open source through http://www.sourceforge.net. At June 04 about 100 universities and libraries were participating, and about 80 publishers. The content is continually and automatically validated against the same content in other caches, to ensure that it doesn't get corrupted or lost. If it does, it can be replaced from the publisher or the other caches. LOCKSS's main purpose is to store older material, where access demands are not heavy; until that stage the material can be accessed in other ways, normally direct from the publisher.</p> <p>LOCKSS is a five year project and is seeking new distributed management arrangements to take over then from the present ones.</p>	
Service offered: Open source software to enable local libraries to operate data stores, with local access, for digital journals	
Type of material held	
Digital journals	
Volume of material held/planned	
No detailed information	

CATEGORY E: RESEARCH CENTRE

CATEGORY: Research centre	
DIGITAL CURATION CENTRE (DCC)	
<p>Location: e-Science Institute, University of Edinburgh, with some functions to be at other partner sites (Glasgow, Bath and CCLRC (qv) centres)</p> <p>contact: digitalcuration@ed.ac.uk</p> <p>URL: http://www.dcc.ac.uk/</p>	
<p>Scope of business: The Centre has two aims: to be an organisation that is research proficient and to be one that is service orientated; each internationally respected, and to a standard that would warrant leadership and advocacy across disciplines. It intends to do the following:</p> <ol style="list-style-type: none"> 1. Establish a vibrant research programme - addressing wider issues of data curation 2. Nurture strong community relationships - forming and extending the Associates Network, engaging with scientific digital curators 3. Develop services - for testing and evaluating tools, methods, standards and policies in realistic settings, and to offer a repository of tools and technical information, a focal point for digital curators 4. Achieve the 'virtuous circle' - feeding expertise, experience and need into its research programme on data curation, and transforming research-led innovation into services that enhance <p>The DCC is setting up and planning to launch in the last quarter of 2004. The JISC and the eScience Core Programme have appointed, to set up the DCC, a consortium consisting of University of Edinburgh (lead partner) and the University of Glasgow, which together host the NeSC; UKOLN, at the University of Bath; the Council for the Central Laboratory of the Research Councils (which operate the Rutherford and Daresbury Laboratories).</p>	
Service offered: The DCC will not offer digital repository services	
Type of material held	
No materials held	
Volume of material held/planned	
No materials held	

CATEGORY F: RESEARCH COUNCIL

<i>CATEGORY: Research Council</i>	
ARTS AND HUMANITIES RESEARCH BOARD (AHRB)	
See entry under Category B, Data Services, Arts and Humanities Data Service (AHDS)	

CATEGORY: Research Council	DPC member
COUNCIL FOR THE CENTRAL LABORATORY OF THE RESEARCH COUNCILS (CCLRC)	
Location: <u>Didcot</u> , Abingdon, Daresbury, Stockbridge	
DPC contact: David Corney, tel: +44 1235 445993, e-mail: D.R.Corney@rl.ac.uk	
URL: http://www.e-science.clrc.ac.uk/web/projects/Data_Curation_Centre	
<p>Scope of business: : The CCLRC, formed in 1995, is a public body of the Office of Science and Technology, part of the DTI; it is one of the UK's Research Councils that, between them, provide the support required for university science and engineering research programmes. Scientific data and its preservation is CCLRC's core business it therefore needs to provide digital preservation advice, tools and facilities to CCLRC users and to research council users.</p> <p>The CCLRC owns and operates the Rutherford Appleton Laboratory in Oxfordshire, the Daresbury Laboratory in Cheshire and the Chilbolton Observatory in Hampshire. These world-class institutions support the research community by providing access to advanced facilities and an extensive scientific and technical expertise, including: to generate public awareness; to communicate research outcomes; to encourage public engagement and dialogue; to disseminate knowledge; and to provide advice. Other research councils, notably EPSRC, BBSRC, NERC and PPARC, are involved in digital curation projects with CCLRC.</p> <p>The Atlas DataStore (ADS), set up in 1983, operated by the e-Science Centre, has a maximum capacity of 1.2 Petabytes, used for projects such as the Particle Physics, but with many other users within CCLRC and externally. DataStore can be used remotely from anywhere in the world, via client software or GRID interfaces. More information is available at: http://www.e-science.clrc.ac.uk/web/services/datastore. It stores user data independently of the physical medium, offering cost effectiveness and the security that data can always be accessed even if the original media become unavailable.</p>	
Service offered: Atlas Data Services (ADS) provides storage for user institutions; the latter also provides storage on request on a commercial fee-paying basis.	
Type of material held	
Scientific data in a variety of formats: Monte carlo, HDF, flat files ASCII etc, mostly born digital	
Volume of material held/planned	
Currently 100 terabytes, with a capacity of over 1 petabyte Expected to increase to ~ 2+ Peta bytes within the next 2-4 years	

CATEGORY: Research Council

ECONOMIC AND SOCIAL RESEARCH COUNCIL (ESRC): ECONOMIC AND SOCIAL DATA SERVICE (ESDS)

See the entry under category B, Data Services, for the UK Data Archive.

url: <http://www.esds.ac.uk>

ESRC jointly-funds and sponsors, with the Joint Information Systems Committee (JISC) the ESDS, which is a distributed, national data service. It brings together a number of centres of expertise in data creation, preservation and use:

UK Data Archive (UKDA), University of Essex

Institute for Social and Economic Research (ISER), University of Essex

Manchester Information and Associated Services (MIMAS), University of Manchester

Cathie Marsh Centre for Census and Survey Research (CCSR), University of Manchester

These centres will work collaboratively to provide preservation, dissemination, user support and training for an extensive range of key economic and social data. ESDS replaces the social science data services offered previously by the UKDA and MIMAS with a single joined-up service, housed at the UKDA.

CATEGORY: Research Council

MEDICAL RESEARCH COUNCIL (MRC)

The MRC has launched a on data sharing and preservation initiative, and produced a draft policy statement (www.mrc.ac.uk/strategy-data_sharing_policy-link). This policy emphasises the importance of data sharing and preservation. It states that 'MRC plans to facilitate development of generic technical standards and tools to enable dataset discovery and, for high-value datasets established with MRC funding, data sharing and preservation' Also that 'Investigators requesting funds for new data collection or renewal of funding will be required to include a data sharing and preservation plan in their proposals'. The policy assumes that data sharing will be normal practice: Applicants whose data are not amenable to sharing should include in the research proposal reasons for not making the data available.

In support of its policy, the Council also issued in February 2004, closing in March 2004, an Invitation to Tender for a data support service, whose intention is 'to create a source of appraisal, guidance and development of strategies for data preservation and for data sharing'. The ITT also said: 'The service will work both on "data rescue" (by applying preservation strategies retrospectively to existing datasets) and on prospective strategies for preservation in new studies'. The Council plans to have a sustainable curation strategy in place by 2005.

CATEGORY: Research Council	
NATURAL ENVIRONMENT RESEARCH COUNCIL (NERC)	
Location: data held at seven data centres across the UK – see website for details contact: URL: www.nerc.ac.uk/data	
<p>Scope of business: Environmental science datasets are collected or generated by NERC scientists and NERC-funded Higher Education Institutions. Datasets are also placed into the custody of NERC as a result of statutory obligations, voluntary deposits, negotiated exchanges or purchase. There are seven designated Data Centres with delegated responsibility for NERC data and implementation of its data policies. All major projects are required to prepare a written data management plan in line with NERC policy. If a dataset is to form part of NERC’s enduring data resource, minimum standards are required for its management.</p> <p>NERC has a formal data policy statement (http://www.nerc.ac.uk/data/documents/datahandbook.pdf) requiring consideration of the ‘post-project’ stewardship of data before approval will be given for a ‘project. The policy also requires that recipients of NERC grants offer to deposit with NERC a copy of any resulting datasets.</p>	
Service offered: data is held in accordance with the data policy. No purely commercial service is offered.	
Type of material held	
Environmental science datasets	
Volume of material held/planned	
No information	

CATEGORY B: PRIVATE SECTOR SERVICES

CATEGORY: Consultancy support	
AUDATA	
Location: Ashford and Canterbury contact: info@audata.co.uk URL: www.audata.co.uk/	
<p>Scope of business: Audata offers strategic and practical services and advice about Information Management, Records Management and Document Management and has experience in digital preservation. It is included primarily because of its work on two major public sector projects relating to digital preservation. It was part of the team working on the Dutch government's digital preservation Testbed project, where its role was to provide strategic consulting including practical experiments and implementation. For further information see http://www.digitaleduurzaamheid.nl/home.cfm. It also project managed the project to assemble a digital archive for The National Archives at Kew. For further detail see http://www.pro.gov.uk/about/preservation/digital/archive/default.htm</p>	
Service offered: customised development relating to digital preservation and storage needs	
Type of material held	
Not relevant	
Volume of material held/planned	
Not relevant	

CATEGORY: Consultancy support	
TESSELLA	
Location: Abingdon, Berks contact: info@tessella.com URL: http://www.tessella.com/tessella/index.htm	
Scope of business: Tessella is a software services company specialising in the support of scientific, technical and engineering establishments. It is included because of its work on three important public sector projects in the digital preservation field: PRONOM for The National Archives (qv); being a team member in setting up a digital archive system for The National Archives - see http://www.pro.gov.uk/about/preservation/digital/archive/default.htm ; and being a team member of the Testbed development project for the Dutch national archives – see http://www.digitaleduurzaamheid.nl/home.cfm . For further details on this aspect of Tessella’s work see http://www.tessella.com/Services/Sector/public_digitalarchiving.htm .	
Service offered: customised development relating to digital preservation and storage needs	
Type of material held	
Not relevant	
Volume of material held/planned	
Not relevant	

CATEGORY: Development services and support	
MAGUS RESEARCH LTD	
Location: Highgate, London Contact: contact@magus.co.uk URL: www.magusresearch.com	
Scope of business: Magus offers services as an Internet content and information management specialist, with strong capabilities in search, retrieval and managed applications. It offers “a broad range of technology solutions including Internet and Intranet projects from concept through deployment and beyond, with expertise from the front-end interface and information architecture through to the back-end content management and database systems”. Magus is included because it was chosen to support the UK Web Archiving project (UKWAC) (qv) – see www.magusresearch.com/webarchive.html	
Service offered: internet related development services	
Type of material held	
Not relevant	
Volume of material held/planned	
Not relevant	