

**Archives On-Line**  
**THE ESTABLISHMENT OF A UNITED KINGDOM**  
**ARCHIVAL NETWORK**

National Council on Archives

1998

## **Contents:**

Executive Summary	p 5
1. Origin and Status of the Report	p 7
2. The need for a National Archival Network	p 19
3. The archives community and Information and Communications Technology in 1997	p 24
4. Data standards for a national network	p 33
5. The technology of a National Archival Network	p 41
6. The management of a National Archival Network	p 45
7. Development and implementation plan	p 48
8. Summary of recommendations	p 50
Appendix 1: Membership of the Networking Policy Committee	p 51
Appendix 2: Glossary	p 52

## EXECUTIVE SUMMARY

This report has been written in the context of the current debate about the need for increased public access to information technology and the creation of networked information resources, to stimulate the emergence of the UK as an Information Society (see Section 2). Although in the light of the widely-acclaimed report *New Library - the People's Network* this debate has focused largely on public libraries, information and communications technology (ICT) also has profound implications for the nation's archival repositories. In particular, the growing tendency of educational users at all levels from primary schools to higher education to regard the Internet as a preferred tool for resource discovery makes it increasingly vital that detailed information about archival holdings should be accessible in this way; a need that will be felt the more strongly as libraries place the Internet at the disposal of every member of the public. In the face of such profound changes to ways of accessing information, it is vital both to Society and to archive professionals that appropriate measures are taken to provide detailed and accurate information in digital form about the nation's great archival heritage. The new technology also offers the opportunity to make archival information more accessible than ever before, to the benefit of a variety of social, education and commercial purposes, and the Council would urge the Government to seize the opportunity to include the creation of a National Archival Network in its emerging ICT strategy.

The report considers whether it is necessary to develop provision specifically for the archival community, or whether a resource discovery mechanism in this area might be integrated with a wider project aimed at cultural resources more generally. It concludes that the present state of both technology and professional standards favours an independent structure, although it is possible - perhaps even likely - that in the future this can be integrated into a broader framework.

The report recommends the creation of a network that aims ultimately to provide access from a single gateway to *all* archival catalogues in the United Kingdom (see Section 5). It proposes a model for the network by which data from individual repositories would be mounted on a series of regional (for England) and national (for Scotland, Wales and Northern Ireland) servers, which would be linked to an interface with the Internet and other public networks. Although the aim of the Network is comprehensive coverage, the Report acknowledges that this cannot be achieved in the short term, and that the Network must begin with those repositories which are most ready to participate.

The creation and implementation of such a National Archival Network would require significant investment in the purchase, maintenance and future upgrading of hardware, software, and communications lines; it would require the widespread introduction of digital cataloguing in the high proportion of archival repositories which have not yet secured this technology for themselves; it would require increased acceptance of and familiarity with national and international cataloguing and indexing standards on the part of archivists; and it would require, over a long period, a large-scale programme for the retrospective upgrading and conversion of existing manual finding aids to digital form (see Sections 3-4).

The report recommends management arrangements for the National Archival Network (see Section 6). It proposes that the Network should be commissioned by a Consortium of the national repositories and professional bodies, and that tenders should be invited for the creation and management of the Network to a specification from this consortium. The report recognises the significant role a National Name

Authority File would play in any archival network and recommends that the Historical Manuscripts Commission should take responsibility for its development.

The report considers the level of funding which would be required to deliver a National Archival Network, and the means by which this might become available (see Sections 5-6). The largest costs will be incurred in generating a sufficient quantity of catalogue data in digital form to make the network a valuable tool. Altogether, we believe that a sum in the order of £38.5m over an extended period would be required for the improvement and retrospective conversion of existing manual lists. The infrastructure costs are heavily dependent upon the technology selected for the implementation of the network, but the minimum estimates of the sums required are £3m for equipping repositories with digital cataloguing systems, £1.5m for training, and £3.1m for the establishment of the hardware and software of a National Network, including regional servers and the central gateway. Annual running costs will be between £5.5m and £7.7m using dedicated leased lines, and around £2.6m using ISDN2 lines for the majority of the network connections. In the light of the Government's response to the *New Library* report, it may be that linking archive repositories and the servers and gateway we propose to the National Grid for Learning would help to reduce the infrastructural costs. Meeting the costs of a national archival network is indeed the main obstacle to its creation, for the archival community cannot simply divert the necessary funding from existing activities; the existence of the network will not release resources elsewhere. The report concludes that since the benefits of the network will be widely felt throughout Society, the funding for it must be similarly broadly based, and might thus come appropriately from the National distributors, from the education sector and from sponsorship generated with the assistance of Central Government.

This report is the first step towards a UK archival network, and is therefore able only to approach the concept at a very general level; section 7 sets out the next steps the Council believes should be taken. The concept and planning of a national network is already much further advanced in Scotland, where detailed proposals have already been accepted by the archival community. Whilst the Scottish model cannot simply be scaled up to the rest of the UK for various technical and professional reasons, the existence of detailed, costed proposals which address matters such as the impact of the increased use of archive services generated by networked access is of the greatest value. We envisage building the experience of the Scottish Archive Network (SCAN) into our proposals by that network's participation in our proposed First Stage Implementation Project (see Section 7).

# 1 ORIGIN AND STATUS OF THE REPORT

## 1.1

The National Council on Archives (NCA) [NOTE 1: Key acronyms and technical terms are explained in the [Glossary](#) at the end of this report] has produced this report following a seminar entitled *Into the ether: vision or void?*, organized by the Archives Sub-Committee of the Joint Information Systems Committee (JISC) and hosted jointly with the NCA at King's College London, in November 1996. That seminar caught the imagination of the wider archival community by bringing to centre stage the potential of digital networks to provide entirely new ways of accessing archives. It considered that the first objective in networking had to be the creation of a gateway through which on-line researchers could perform simultaneous searches on the finding aids of UK archival institutions, and gain access to the detailed catalogue descriptions of individual items. It charged the NCA with exploring the feasibility of developing a National Archival Network to make this possible.

## 1.2

A Networking Policy Committee, appointed by the NCA in December 1996, undertook the study leading to this report. Its membership has been drawn from as broad a field as possible: from the university sector, local authorities, the main national agencies - the Royal Commission on Historical Manuscripts (HMC) and the Public Record Office (PRO) - and from the national archival networking initiatives planned in Scotland and Wales. In view of the historic and geographical links between the United Kingdom and the Republic of Ireland, a representative from the Republic was also invited to join the group. A list of the members of the committee is given in [Appendix 1](#). The specific brief of the committee was to draw up a technical and intellectual model for a national network structure involving authority controls, levels of data capture and the acceptance, subscription to, and incorporation of, appropriate standards for archival description; to make proposals for how such a structure might be managed in the future; and to consider funding issues.

## 1.3

This report sets out the NCA's view of how a National Archival Network can best be achieved, and what steps need to be taken to make it a reality. It has become clear during the work of the Networking Policy Committee that the proposals are both relevant and timely in the context of developments elsewhere in the library and information sector, such as the publication of the UKOLN scoping study *Towards a National Agency for Resource Discovery* and the Library & Information Commission report, *New Library - the People's Network*; and the emergence of more detail about the Government's proposals for a National Grid for Learning. For this reason the NCA has determined to make this document a report both to the archival profession and to the wider world of users, providers and potential funders. The governing bodies of archival repositories and other organizations interested in the future of UK archives are invited to comment on, and to endorse, the recommendations in the report, and the action plan in section 7.

## 1.4

For the wider world, the report is intended as a statement by the archival community of its priorities in relation to networking, how these are justified, and how they fit into the current strategic debates about resource discovery and the creation of an Information Society. It is hoped that the publication of this report, and the widespread endorsement of its views which we anticipate will follow publication, will encourage the Government, local authorities, distributors and other potential funders to work in

partnership with the archival community to make the proposals set out in the report a reality in the near future.

## 2 THE NEED FOR A NATIONAL ARCHIVAL NETWORK

### 2.1

In many walks of life, archives are needed on a daily basis as the source of information about actions taken or decisions recorded in the past which continue to affect our rights and entitlements today [NOTE 2: Paragraphs 2.1-2.3 are based upon C. Kitching, *Archives: the very essence of our heritage*, Phillimore for the National Council on Archives, 1996. The NCA is grateful to Dr Kitching for permission to quote from his book at such length in this report.]. Registers of births, marriages, deaths and adoptions, for example, enable us to establish our very identity, our ancestry and birthright. Title deeds contain evidence of individuals' rights to property, as well as information about the age of a building and succession of its ownership. Charters define corporate rights and privileges. Trust deeds spell out the terms of reference of a charity. Patents for designs and inventions protect commercial interests. Historic architects' drawings of a country house or garden may be aesthetically pleasing in their own right, but assume a vital new role when it comes to a restoration project. Businesses, too, may successfully recycle designs from their archives of historic labels, posters or packaging for today's advertising campaigns. Maps and plans can have many practical uses, from verifying rights of way to identifying old mine workings which might cause building subsidence, contaminated land which might be a hazard to future generations, the lines of drains, or the structure of buildings. For many archive services, these and similarly urgent or important types of enquiry are a large and growing area of work.

### 2.2

Whilst the records of individual government departments, local communities, businesses or families may tell us a lot about their respective originators, they become much more valuable when brought together with complementary records by an archives service to achieve, as it were, a 'critical mass' that can sustain wider research. At a national level, archives are the key to an understanding of past governmental policies and decisions which have affected every sphere of life, in both domestic and foreign policy. We are reminded of this each January as the press eagerly descends upon the Public Records newly released under the thirty-year rule. But archives service a wider cultural need, too. They provide a framework for our understanding of the past: how our forbears thought and behaved; what life was like for them; how they worked and played; the social, religious and political context of their lives. At the level of the local community, they can help to explain how a building, a street, a village or a town, or for that matter a business or other organisation, took shape and why it is as it is now. They have enormous, and currently underexploited, potential for educational use at all levels from primary school to higher education. Ultimately, they provide a major cornerstone of our historical knowledge, and in the form of information analysed and interpreted by historians, archives reach a vast reading public who might never, personally, visit an archive service.

### 2.3

The increasing availability of archives for study has undeniably broadened the kind of research that can be undertaken. This is not only because more evidence makes for better-informed research, but also because familiarity with a wider range of records and the data they contain has helped historians and others to know where to turn for evidence that was previously thought to be in acutely short supply, for example for the study of women and children in the Middle Ages or of ethnic communities in more recent times. Knowledge of what records exist and where they are located is

fundamental to their effective exploitation: we have not really appropriated our archival heritage at all until we have access to it. Many record repositories have therefore been working hard at public outreach, through lectures, group visits, publications and exhibitions, and through roadshows and other public events taking archives to the heart of the community. As a direct result, more and more people find a professional or leisure interest in studying the past, and the utility of archives for research and teaching in disciplines from architecture to medicine is ever more widely recognised. In schools, the educational power of archives has long been recognised, but the difficulty of locating relevant resources and then obtaining copies has often held back wide exploitation of this resource. Awareness about the availability and location of particular records - rather than archives in general - requires, however, a co-ordinated national strategy. For more than half a century, repositories have voluntarily co-operated in ensuring that the National Registers of Archives for the UK and Scotland act as central locations to which the researcher can first address him or herself in order to locate relevant documents. Fifty years on, the challenge is to provide a similar point of access for a generation of researchers who wish to undertake their research from a terminal in a classroom in Aberdeen, a public library in Aylesbury or a university in Adelaide.

#### 2.4

Responsibility for the preservation of and public access to historical documents in every format from Anglo-Saxon parchments to digital information and videotapes is widely shared between national and local government, the universities, specialist institutes and private owners. There are about 2,000 recognized archival repositories, although many of these are dedicated to the care of a single (often very significant) archive. Some 300 institutions have the main responsibility of collecting and preserving our documentary heritage, and it is these that should form the backbone of any National Archival Network. At the top level are the bodies with a defined national responsibility: the Public Record Office, the National Archives of Scotland, and the Public Record Office of Northern Ireland. National libraries - the British Library and the National Libraries of Scotland and Wales - house collections of broad national significance as well as a host of smaller archives. Local authority archive services function as repositories for the collection of historical records within and about their local communities, and often also have material of national or wider significance [NOTE 3: Some, but not all, of these services are provided as part of local authority library services.]. The university sector houses an enormous diversity of archives, as well as maintaining the records of the universities themselves, and there are a number of specialist institutes funded from government or charitable sources, such as the British Film Institute or the Wellcome Institute. It is thus a key characteristic of the archival community in the UK that it is cross-sectoral, and this has important implications for the design, funding and implementation of a national archival network which will be considered later in this report.

#### 2.5

Not only are archival collections widely dispersed, but the co-ordination of their work is largely collaborative in nature and not directed by any national monitoring agency. Such regulatory functions as do exist are generally exercised by different bodies in the constituent countries of the United Kingdom, and are very limited in scope. It is accordingly another key characteristic of the archival community that it has to proceed by co-operation; in the present legislative framework, there is no mechanism for enforcing change or development. The community has, however, proved itself to be increasingly capable of collective action to achieve common objectives in recent years, a fact of which the formation of the National Council on Archives in 1987 is perhaps a symbol. In 1996 the Council issued *An archives policy for the United Kingdom* [NOTE 4: National Council on Archives, *An archives policy for the United*

*Kingdom*, NCA, 1996], which was widely endorsed by British archival institutions and the various professional associations, and which itself provides part of the policy context for this report. There will need to be similar widespread professional endorsement of the findings of this report if its recommendations are to be carried into effect.

## 2.6

The preceding paragraphs have demonstrated that access to our documentary heritage is essential to Society for the contribution it makes to education from primary level, through lifelong learning to academic research. The increasingly widespread availability of access to communications networks such as the Internet has stimulated a good deal of research and debate about the potential such networks may have to improve access to the vast storehouse of information in libraries and archives. The power of modern communications technologies opens up the possibility of achieving a degree of co-ordination in the presentation of these resources to the end-user which has never been possible in the past. Such improved co-ordination would make the process of resource discovery both simpler and more efficient, and the first and best argument for the creation of a national archival network is thus that it will make it easier for users to locate the materials they need.

## 2.7

The corollary of this opportunity is also of significance to archivists. The vast heritage of published and unpublished information in the nation's libraries and archives that is not in digital form remains as relevant as ever to society's information needs, but risks becoming progressively less visible to users as the Internet becomes increasingly the public's, and especially the research community's, tool of choice for resource discovery. The longer the archives profession is unable to respond to the opportunities of Information and Communications Technologies (ICT), the more risk there is that this lack of visibility will progressively reduce the use of British archives, and it will certainly have an adverse effect on the quality of research. It is therefore of great importance for improving access to and understanding of our documentary heritage, as well as for the health of the UK archival community, that the opportunities presented by ICT for enhancing public access to archival information are seized not just by the larger and most generously financed repositories but by the archival community as a whole.

## 2.8

The goal, essentially, remains the same as when the National Registers of Archives for the UK (NRA) and for Scotland (NRAS) were first envisaged: to connect the researcher and the relevant document in the most efficient way possible. In modern terms, this means that a researcher anywhere in the world who has access to the Internet should be able to contact a common gateway, submit a single enquiry and receive a single integrated response, listing the relevant source materials housed in all UK archive repositories. The Historical Manuscripts Commission (HMC) has allocated some resources so that the NRA can continue to play its traditional role in a quickly changing digital environment. As a commitment to this, the HMC has added a new corporate priority to its current Business Plan: "To maintain the central lead role of the NRA in the context of the nationwide networking of digital data about archives and manuscripts". The resources of the HMC alone, however, cannot bring into being the infrastructure and management arrangements required to deliver an effective and comprehensive national network and make this aspiration a reality.

## 2.9

Any remote-access network set up to enable researchers to pursue their interests and needs will be immeasurably enhanced by being constructed so as to transcend sectoral boundaries within the archival community. The search for the papers of an

individual politician or writer, for the evidence of a social, political or religious movement, or for references to a specific place or business, will lead the researcher to collections stored in repositories in several different sectors. The essential archival quality of uniqueness ensures that the collections of different repositories cannot duplicate one another, but together constitute the fabric of the national memory. The sectors of the archival community are thus unlike those of the library world, which build collections tailored to the needs of their clientele, but which are overlapping subsets of the total body of published information. Any national archival network must be constructed to encompass all sorts of archives.

#### 2.10

In terms of geographical coverage, the NCA has produced recommendations for a national archival network to cover the four countries of the United Kingdom. In view of the historical and geographical links between the UK and The Republic of Ireland an Irish archivist participated in the discussions leading to the production of this report. The Council's view is that it would be beneficial to develop a joint British and Irish Archival Network if the formidable problems posed by trans-national funding, divergent telecommunications systems and legislation could be overcome. National initiatives for archival networking are already being developed in Scotland and Wales, and it is not the intention of the NCA to propose a model which will in any way undermine, circumvent or duplicate these initiatives. Rather, it is the Council's hope that they can play a full and early part in the development of a broader network. It may be possible for the objective of a joint British and Irish Archival Network to be achieved by creating similar linkages between a national network for The Republic of Ireland and an overall British & Irish project.

#### 2.11

The Library and Information Commission (LIC) has made opening up digital access to the national memory a key issue in its 2020 Vision, by aiming to "support the enabling of a digital library collection in which the UK's heritage of intellectual property will be available" [NOTE 5: Library & Information Commission, 2020 Vision, LIC, 1997], and it is also the principle behind the [British Library's Digital Library initiative](#) [NOTE 6: See the Digital Libraries programme website: <<http://www.ukoln.ac.uk/services/bl/digital-library-research-projects>>]. The LIC commissioned a report from Information North in 1997 to identify the scope and extent of digitization projects in the local authority library and archive sector, and to recommend what new projects should be initiated [NOTE 7: Information North, *Virtually new: creating the digital collection*, LIC, 1998].

#### 2.12

It is the view of the National Council on Archives that the sheer scale of the UK's archival holdings makes any scheme for their *wholesale* digitisation impractical in the context of existing technology. Whilst there are already many worthwhile projects underway to put particular resources into digital form, and whilst this number will undoubtedly increase in future, the only project that holds out any prospect of putting a comprehensive representation of the nation's archival heritage into the digital domain is one which focuses on catalogues and indexes, not on images of documents. Accordingly, this report is concerned to identify a route by which such a digital resource could be created.

#### 2.13

In concentrating on a network of finding aids in this way, the Council believes that it is laying the foundations for a broader range of services in future, since digital catalogue entries will provide the essential indexing tools required for other forms of digital delivery. The NCA has no intention of playing down the value of other digitisation initiatives, and indeed co-hosted a seminar at the Public Record Office in

February 1998 to discuss the potential for collaborative projects in this field. Digital imaging offers archive services the opportunity to copy fragile originals once and to make available good quality copies indefinitely thereafter without further damage to the original. The NCA believes, however, that the most useful and exciting opportunities for digital imaging lie in creating thematic selections of (sometimes interdisciplinary) resources rather than in digitising whole collections or classes of records. Some such projects will be commercially viable, as has already been demonstrated by the [Research Libraries Group's](#) 'Studies in Scarlet' programme [NOTE 8: See the Research Libraries Group website: <<http://www.rlg.org>>], while others will be tailored specifically to the needs of the National Curriculum or to the history of particular communities, and will attract public funding because of their value to educational or social objectives.

## **The policy context of the Report**

### **2.14**

Perhaps the most significant background document to this Report is a recent scoping study entitled [Towards a National Agency for Resource Discovery](#), which was funded jointly by the British Library Research & Innovation Centre (BLRIC) and JISC [NOTE 9: P. Brophy *et al.*, *Towards a National Agency for Resource Discovery*, BL Research & Innovation Report 58, July 1997; available on the website:

<<http://www.ukoln.ac.uk/services/papers/bl/blri058/>>]. The study proposes that a National Agency for Resource Discovery should be set up to act as a facilitator to ensure that scholarly resources are visible and accessible across sectors and other traditional boundaries in an effective and sustainable way. The report envisages that the first priority of the Agency would be resource discovery for library and on-line information, but 85% of respondents to a survey of potential users welcomed the suggestion that archival data might also be included. The report envisages that the Agency would focus on encouraging the creation and accessibility of collection-level data, and promoting the Z39.50 interoperability profile, while encouraging individual sectors and disciplines to progress work on standards. The report proposes that the agency should be partly, but not exclusively, funded by the Higher Education sector and that it should be established in 1998 to provide timely support to eLib phase 3, the public library networking initiatives and the proposed National Archival Network. The proposed Agency is viewed by the Council as broadly helpful in that it could provide strategic co-ordination between the National Archival Network and parallel initiatives in related disciplines.

### **2.15**

The [Anderson Report](#) for the Joint [Higher Education] Funding Councils' Library Review in 1995 [NOTE 10: Joint Funding Councils' Library Review: Report of the Group on a National/Regional Strategy for Library Provision for Researchers (The Anderson Report); available on the website: <<http://www.ukoln.ac.uk/services/elib/papers/other/anderson/>>] reported on national and regional strategy for library provision for researchers. It acknowledged that a very significant proportion of the primary sources for research used by the Higher Education sector was located outside the sector in national and local authority libraries and archives, and said that a national strategy should provide 'the means to locate and gain access to material with reasonable ease, reasonable speed and at reasonable cost'. The report emphasised the provision of 'adequately co-ordinated information on the location and current availability of research material'.

### **2.16**

The LIC report, [New Library - the People's Network](#) [NOTE 12: Library & Information

Commission, *New Library: the People's Network*, LIC, 1997; available on the website: <<http://www.ukoln.ac.uk/services/lic/newlibrary/>>; see also the discussion forum on the same website.], which was commissioned by the Government, has received a very positive response at all levels from the Prime Minister downwards and is likely to form the basis of developments in public library policy in the medium term. One of its key recommendations is the establishment of a Public Library Networking Agency, among the functions of which will be to 'procure and develop content' for a national digital network, including a programme to digitise rare and special collections in public libraries. Despite the paucity of references to archives in the *New Library* report, many of its recommendations about the accessibility of 'rare/special collections' would apply equally well to archival material, and it is understood from the LIC that the language of the report should not be read in an exclusive way. The report singles out local history collections as amenable to digitisation to 'make these resources more widely accessible and...facilitate the security and conservation of the original, often inherently valuable, documents', and says 'In this area the content to be delivered will include...records of births, marriages and deaths and local newspapers; digitised collections of maps and photographs...; catalogues of local history libraries across the world [etc.]'. It is clearly only a small and uncontentious step from this vision to providing access to the National Archival Network as part of the content development for the public library network.

#### 2.17

In its consultation document on the National Grid for Learning, [\*Connecting the Learning Society\*](#) [NOTE 12: DfEE, *Connecting the Learning Society: National Grid for Learning*, DfEE, 1997; available on the website: <<http://www.open.gov.uk/dfee/grid/>>], the DfEE set out a series of phased objectives for the creation of the National Grid for Learning. In the second, main development phase, the Grid is envisaged as providing resources for higher education and independent learners, not just for schools. The report stated that 'developing a nucleus of good quality content will be [the] prime concern' [NOTE 13: *Ibid*, p. 16.], and given the educational value of archives, it is clear that the catalogue information on a National Archival Network would be of value to the National Grid for Learning. In its response to the [\*New Library\*](#) report, the Government has made it clear that it sees the proposed educational and public libraries networks as being integrated into a single entity [NOTE 14: "*New Library: the People's Network*": *the Government's response*, Cm 3887, HMSO, 1998 ; available on the website: <<http://www.culture.gov.uk/new-library.htm>>]. It recognises that much of the content proposed for a public libraries network will be of broad educational value, but bringing together the Public Libraries Network and the National Grid for Learning will also enhance the relevance of the latter to lifelong learners, and extend its significance beyond the educational sector. A national archival network would seem to fit logically into the same framework, since schools, libraries and independent learners will make up such a significant part of its audience.

#### 2.18

The Heritage Fund commissioned a report in 1997 on [\*Funding Information and Communications Technology in the Heritage Sector\*](#) [NOTE 15: S. Ross, *Funding Information and Communications Technology in the Heritage Sector: policy recommendations to the Heritage Fund*, Humanities Advanced Technology & Information Institute, University of Glasgow, 1998; available on the website: <<http://www.arts.gla.ac.uk/HATII/>>], which sought to define some priorities for the funding of projects involving the use of ICT. The report argues that the use of ICT to support the creation of digital collections describing heritage resources, the retrospective conversion of existing heritage catalogues and records, and the

digitisation of resources to improve their public access should all attract support. Among the specific recommendations in the report are that:

- the Heritage Fund (HLF) should not normally fund the hardware and software requirements for network infrastructure, although some modest components of this kind should not be precluded;
- that the conversion of existing catalogues, inventories and finding aids should be supported where value is clearly added as a result (e.g. through inter-organisational projects);
- that the establishment of websites on the Internet should only be funded as a by-product of a consistent and coherent information service;
- that the Fund should require that access to all ICT-based resources should be free for educational purposes;
- that preference should be given to collaborative projects;
- that the Fund should investigate the feasibility of supporting national data services, in conjunction with the statutory bodies in the field;
- and that the Fund should require applicants to use appropriate national and international standards.

The NCA understands that these recommendations have been accepted, and notes that the creation of a National Archival Network as proposed in this report, and especially the retrospective conversion of catalogues for inclusion in it, falls neatly within the range of projects which the Heritage Fund wishes to support.

#### 2.19

The Higher Education sector has been developing its thinking about the selection of library and archival material, and other data, for digitisation and inclusion in national networked information resources [NOTE 16: Joint Information Systems Committee, *An integrated information environment for Higher Education: developing the distributed, national electronic resource (DNER)*, 1997; available on the website: <[http://www.jisc.ac.uk/cei/dner\\_colpol.html](http://www.jisc.ac.uk/cei/dner_colpol.html)>]. The Committee on Electronic Information has concluded that among the criteria for selection would be:

- a mix of traditional and non-traditional resources, including categories of material such as manuscripts, maps, moving and still images, and sound recordings.
- resources which meet teaching and learning needs, and which support research work.
- resources which build up a critical mass of information in individual disciplines
- the support of disciplines where there are financial or other barriers to the development and use of electronic resources, through the identification of strategic investments capable of providing an impetus to the adoption of networked resources services in these disciplines.
- resources supported by the availability of appropriate metadata (such as would be provided for archival material by on-line catalogues) would be highly desirable for inclusion.

The report also notes that the Joint Information Systems Committee will seek collaborative opportunities to develop the national electronic resource, and that JISC is eager to work with non-HE agencies and institutions that create and use data of interest to the HE sector.

#### 2.20

Nearly all the recommendations made in the Foresight report on Leisure and Learning [NOTE 17: *Leisure and Learning*, Technology Foresight 14, Office of Science & Technology, 1995, pp. 75-79] depend upon the generation of substantial quantities of digital content if developments are to take place in the demand for information and learning, skills development, the social impact of new technologies, and technology and wealth creation. In a wider social context digital content forms the raw material on which a range of interactive online education programmes depend. In its summary report, Foresight recommended that the UK 'encourage the development of new multi-disciplined content-based electronic businesses including...remote learning...[and] leisure products' and acknowledged that a key long-term issue to the economic success of the UK depended on 'Capitalising on the UK's artistic and creative strengths in the content-based information industries'. Archivists curate our documentary heritage and have a central role to play in the creation and provision of digital content that must be better represented in the overall strategies for access to resources.

#### 2.21

The National Council on Archives published *An Archives Policy for the United Kingdom* in 1996 [NOTE 18: National Archives Policy Liaison Group, *An archives policy for the United Kingdom*, NCA, 1996]. This noted that "archivists have been much slower than other information professionals to make use of modern information technology, partly because the unique nature of records and archives denies archivists the benefits which librarians have been able to derive from shared cataloguing, partly because they are too few in number to tempt software houses into the development of turnkey archives systems, and partly because the scale and nature of the national archival heritage precludes the full-text processing of anything more than a small proportion of it in electronic form". The report noted the "inter-communicability of data about archival holdings" as the first of three challenges to be met in the ICT area, and recommended national and international co-operation to find ways in which these could be addressed. This report is the result of just such co-operative activity.

#### 2.22

At the end of 1996, the Library & Information Commission commissioned a review of digitisation activity in the local authority library and archives sector [NOTE 19: Information North, *Virtually new: creating the digital collection*, LIC, 1998]. The report found that most digitisation projects in the local authority sector related to archives and library special collections, such as local studies materials. A focus group discussion carried out as part of the study resulted in agreement that improving access to these collections should be the key motive for digitisation, but the report noted that this did not necessarily imply direct access to digital images of documents or printed works: "creating, automating and networking catalogues and finding aids has to be recognised as the priority for many local studies collections, special collections and archives". A seminar for librarians, archivists and museum professionals organised jointly by UKOLN and BLRIC at the British Library on 25th February 1998 also concluded that whilst there were significant advantages to the selective digitisation of material such as newspapers, maps and photographs in archives and local studies collections, the most basic and essential requirement was for digital catalogues which were capable of being networked nationally and

internationally. This applied both to archives and to local studies libraries, which were often significantly behind other areas of local authority library services in terms of retrospective conversion.

## **The research and development context of the Report**

### 2.23

The NCA recognises that it is by no means starting with a blank sheet in setting out to design a national archival network. The Council is conscious of considerable activity already being undertaken in this field. In particular, the work of the Archives Sub-Committee of JISC in exploring the technology by which searches might be undertaken across a range of proprietary finding aid databases, developing user-friendly screens for searching and effective display formats for responses, and exploring issues around subject index authority lists, is proving invaluable as a testing ground for the archival profession (see paragraph 2.27 below). The HMC, working in partnership with the NCA, has been the moving force in developing personal, corporate and place-name authority rules and in exploring the possibilities for maintaining central authority lists. The Commission, which has had responsibility for the National Register of Archives since its inception, naturally also sees itself as playing a leading role in the migration of the principle of the Register to a new generation of technology, and is currently seeking to define an 'archival intranet'. It has mounted on its web-server the contact details of the 2000 UK repositories, and as each institution mounts a home page on the Web, the Historical Manuscripts Commission substitutes a direct link for its own entry. At present, about 11% have such links from the Commission's web site. The Commission also plans to acquire web-crawling software, which will be used to provide a central index to the contents of all pages mounted by UK archival institutions, and a dedicated archival search engine will enable users to discover what new information has been added. These developments would continue to be valuable even if a national archives network is created.

### 2.24

There are a number of current national and international research projects which are highly relevant to the establishment of a national network. These include the American Heritage Project, the Arts & Humanities Data Service on-line retrieval system, and the National Networking Demonstrator Project, described in the paragraphs which follow.

### 2.25

The [American Heritage Project](http://sunsite.berkeley.edu/FindingAids/EAD/ameriher.html) [NOTE 20: See the American Heritage Project website, <<http://sunsite.berkeley.edu/FindingAids/EAD/ameriher.html>>] is a demonstrator project to create a national union catalogue of finding aids relating to the American heritage, with special reference to the nineteenth and early twentieth centuries. The intention is to bring together hundreds of finding aids comprising thousands of pages of text. By concentrating on a specific subject area and historical period, the database will enable the project to study the encoding and content issues associated with combining information on related subject matter from different institutions in the same database. It is hoped that it will provide sufficient information on a topic to attract significant numbers of users, permitting accurate user behaviour studies. The project is based at the University of California, Berkeley, and also includes material from Stanford University, Duke University, and the University of Virginia. Locally created records are passed to the central database. Collection-level material is given in the US MARC format, while finding aids are represented using Encoded Archival Description (EAD) [NOTE 21: Encoded Archival Description is a

standardised format for the representation of archival finding aids as SGML (Standard Generalised Mark-Up Language) encoded text. The format is based on and capable of being mapped to the ISAD(G) data content. For a further explanation see below, paragraph 4.3.]. The union database allows a user to search a bibliographic catalogue displaying collection-level records and, from within the bibliographic record, to click on a user interface button that will launch a browser to navigate the collection's finding aid. The project will report in 1998 and is designed to look at a range of issues:

- intellectual issues, including the development of finding aid content standards to enable finding aids from different institutions to coexist in the same database, and even to be integrated;
- political issues, including looking at the problems of decentralised creation and maintenance of finding aids, and at issues surrounding the ownership and responsibility for creating local catalogue records which will be consistent with the central database;
- technical issues, including access, and the description and control of finding aids representing collections with related subject matter from different institutions. The remote creation and maintenance of finding aids will be considered, as will the potential use of natural language retrieval technology on the union database;
- economic issues, such as the cost of finding aid conversion, data input, database maintenance, training and documentation.

## 2.26

The construction of interoperable catalogues remains a matter of considerable technical complexity and a focus for research. As part of the search for better techniques, the Arts and Humanities Data Service has awarded a contract to a consortium of Fretwell Downing Informatics, the University of Liverpool and the University of California, Berkeley, to deliver an on-line information retrieval system for cross-domain searching of bibliographic, full text and multimedia resources. This system, based on international standards, SGML and Z39.50 and a variant of Fretwell Downing's VDX gateway software (also used in the National Networking Demonstrator Project) is used for cross-domain searching and uses the Cheshire information retrieval system which is the basis of the Berkeley Digital Libraries project [NOTE 22: The Cheshire II Information Retrieval Project is a "next generation online catalogue and full-text information retrieval system", and was designed to overcome the problems of search failure and information overload which bedevil online catalogue searching. The system incorporates a client/server architecture, and uses Z39.50 and SGML. For further information, see the [website](http://sherlock.berkeley.edu/asis_paper/paper.html): <[http://sherlock.berkeley.edu/asis\\_paper/paper.html](http://sherlock.berkeley.edu/asis_paper/paper.html)>]. The project was designed to allow simultaneous access to archives, museums and libraries, containing a mixture of text, images and sound recordings. The system is designed to cope with a network based distribution system with local servers responsible for maintaining individual collections of digital documents which will conform to a specific set of standards for document description, representation and communication protocols based on SGML and Z39.50. The system is due for implementation in August 1998, and whilst it may not prove to deliver the best possible and most comprehensive solution, it will certainly be a valuable contribution to technical progress in this area.

## 2.27

The National Networking Demonstrator Project for Archives has been mounted as a working test of the potential for multilevel cross-searching of archival catalogues,

using the Z39.50 protocol [NOTE 23: See the [project specification](http://www.kcl.ac.uk/projects/srch/reports/z3950.htm), mounted on the website: <<http://www.kcl.ac.uk/projects/srch/reports/z3950.htm>>]. Led by the JISC Archives Sub-Committee, it takes as its starting point the international standard for archival description, ISAD(G), and will build on work already undertaken in mapping ISAD(G) to Z39.50, develop a uniform front end to permit seamless public access to multilevel catalogue data from a wide variety of repositories, and report by May 1998 on the implications of bringing together data in different formats, from different sources, in these ways. The Project has the participation of some two dozen repositories, mainly from the Higher Education sector, but also including the Public Record Office, local government, specialist and business institutions. It aims to work with the widest possible variety of source data: fielded data from proprietary finding aid systems, data in commonly used formats including MARC and EAD, and also word processed texts, tagged for project compliance. Whilst recognising the potential of EAD, and the strength of moves towards it among the better-funded Higher Education institutions, the project team considers it would be inappropriate, for a practical demonstration of what is currently possible, to prescribe its use: data in EAD format is of course supported by the Demonstrator. The project has identified the provision of collection-level records as mandatory for inclusion, but aims to achieve functionality over all available levels of description. Participants have a range of options in submitting data: those with their own servers can elect to link them into the Project, making their data available directly, through system adaptation, or indirectly, via export to a 'reference target'. The reference target will also be made available to those without access to a server. It is hoped that suppliers of proprietary systems will be willing to co-operate through their own Z39.50 development work.

#### 2.28

Joint ventures have begun to emerge. In Scotland, the Scottish Archives Network, a proposal led by the National Archives of Scotland, will provide a national network of all Scottish archives. A digital search room will provide the top-level descriptions of all collections in Scottish archives with hypertext links to more detailed item level descriptions when these are available. All will be searchable electronically using free text retrieval and will be available on the Internet. In Wales, a scoping study for the establishment of a Welsh Archival Network was approved by the Archives Council Wales at the end of 1997, and detailed proposals will be developed during the coming year.

#### 2.29

As familiarity with, and access to, Web technology increases in the coming months and years, it seems possible that increasing numbers of individual repositories will begin to provide some form of remote access to meaningful finding aids via their Web-sites, certainly giving data at collection level, and frequently more detailed information. Given the gathering pace of development, it is timely to examine the potential for broad-based co-operation, not least in order to try and avoid the problems of piecemeal investment in systems and practices which, in the event, may prove to be mutually incompatible; a problem which has beset the archival world for decades.

#### 2.30

One of the questions the Council has considered is whether it is sensible to develop a specifically archival solution to the challenge of networking, given that the user at the monitor may well be indifferent to whether his evidence is found in documents, photographs, printed books or museum objects. On the whole the Council believes that, given the hierarchical structure and evidential properties of archival data, it will be better for the foreseeable future to concentrate on developing a model which will reflect the particular needs of the archive community. There is, however, no reason

why such a model should not form part of a larger project devoted to exploiting the needs of the UK library and information world if this is able to encompass the ambitions of the archival community without compromising its particular needs and standards. Work done in recent years (the Dublin Core project [NOTE 24: For information about the Dublin Core metadata project see the following websites: <[http://purl.org/metadata/dublin\\_core](http://purl.org/metadata/dublin_core)> and <<http://vancouver-webpages.com/VWbot/VW-dublin-core.html>>.] has examined the feasibility of establishing a standardised metadata element set to facilitate digital resource discovery across disciplinary boundaries, which might open up the possibility of such projects. It will certainly behove the archive profession to remain conscious of developments on the broader front and to ensure that the NCA initiative is widely recognised and understood outside the boundaries of the profession.

### 2.31

The national archival network outlined below is one in which every repository in the country will eventually be able to store catalogues down to item level. Clearly the speed with which individual archives will wish or be able to move towards this position will vary according to local circumstances. The next section of this report reveals that there are enormous variations in the ict expertise of archivists, and in the expertise to which they have access through their organisations.

### 3 THE ARCHIVES COMMUNITY AND INFORMATION & COMMUNICATIONS TECHNOLOGY (ICT) IN 1997

#### 3.1

In order to provide a factual foundation for its study the NCA carried out a survey during the summer of 1997 into the availability of information technology and network connectivity and the adherence to key listing and indexing standards in the archival community. A questionnaire was circulated to around 250 national, local authority, university and specialist repositories, and a total of 183 responses was received [NOTE 25: The results of the research were summarised in a paper delivered by Nicholas Kingsley to the Annual Conference of the Society of Archivists in September 1997 entitled "Are we ready for IT?". The sample size of 250 was based on the repositories appearing in the Royal Commission on Historical Manuscripts' guide, *Record repositories in the UK*, 9th edn. Since the research was undertaken, a 10th edition of this guide has been published, which includes some fifty additional repositories, and the higher figure of 300 has been used as the basis for costing the National Archival Network. It is believed, however, that the findings given in this section of the report remain approximately representative of the repositories proposed for inclusion in the Network.]. Of the returns received, 8.5% came from the national repositories, 52.5% from local authorities, 18.6% from university repositories and 20.4% from specialist repositories. 79% of these responses came from England, 12.8% from Scotland and 6.2% from Wales. There were also 2 responses from the Republic of Ireland, 2 from the Channel Islands and 1 from Northern Ireland. Non-respondents were drawn from all sectors of the profession, but most of them were among the smaller institutions in their sector; Scottish local authority repositories were particularly weakly represented. The not unexpected bias towards larger repositories needs to be taken into account in interpreting the questionnaire results.

#### 3.2

In addition to this survey, the NCA simultaneously employed a consultant to study the readiness of the archival profession to use and contribute to National Name Authority Files, and he also looked at the extent of automated finding aids in archival institutions. His report confirms many of the findings of the NCA's own survey, and the findings from the two pieces of research have been integrated in the paragraphs which follow [NOTE 26: Peter Gillman, [\*National Name Authority Files: a report to the National Council on Archives\*](#), BL Research & Innovation Report 91, 1998; the report is available on the website: <<http://www.hmc.gov.uk/nca/nnaftit.htm>>].

#### 3.3

The first part of the NCA questionnaire looked at present ict equipment and firm future plans for its provision or upgrading. 5% of respondents reported that they had no computers: these were mainly the smaller English local authority record offices. The remainder mostly had from one to twenty machines, but the figures were complicated by the fact that some respondents (mainly local authority libraries and university libraries, and specialist repositories forming parts of larger institutions) included in their figures machines not *dedicated* to archival use. It was impossible to determine whether some or all of these machines were genuinely part of the ict provision for archives, without knowing how they were used in individual cases. The report showed that while most repositories have some ict capability, the bulk of the equipment reported is relatively obsolescent. The pattern is of a mix of stand-alone and networked machines, often running incompatible software that prevents any effective data transfer between machines, even within a single institution. Approximately 45% of PCs are fitted with 386 or lower specification processors,

which lack the processor power, memory size and hard disk capacity to run Windows applications.

#### 3.4

The NCA questionnaire also asked whether the public had access to any of the ict equipment listed, and whether it was networked. 65% of respondents had some or all of their machines networked, but in a few cases it was clear from additional comments that the network facilities were used to provide access to non-archival databases, such as financial management information or library catalogues. Quite a lot of repositories had a networked accessions register system. 38% of repositories had machines to which public access was given, but there was no significant correlation between this response and the repositories which used databases for cataloguing; public access was as likely to be for access to the Internet or to a specialist tool, such as a photographic index, as to a general finding aid for the repository.

#### 3.5

The report showed that the predominant software application on all ict equipment reported is the word processor. From the responses, amplified by visits and interviews conducted by the consultant, it is evident that one of the principal uses of ICT in repositories is to automate the production of otherwise conventional finding aids. Typically this is by using a word processor to hold the text of lists and descriptions. Instead of printing out this text, it can then be searched using the character-string search function of the word processor. In some instances the finding aids are printed as well, for use in public search rooms and dissemination in hard copy form. This type of ict application is adequate when its products are confined to the originating body, but as a means for data exchange it is of limited value; the 'flat' word processed file generally has no structure in the sense of fields to hold the various types of data associated with a record. It may be possible to tag the word-processed file in such a way that the data can be transferred to a database, but this is not always straightforward, and it may be as quick to rekey the data from scratch.

#### 3.6

The survey asked about the availability of database software and whether this was used for cataloguing archives, in an effort to discover how widely the profession has yet adopted digital cataloguing and indexing systems. Only 16% of the respondents were using archival cataloguing software. They were using a range of proprietary applications, and only about half were using systems of the current generation. Most repositories were clearly only just beginning to take decisions about the purchase of software, and to consider how to tackle the issues of backlog conversion. A number of respondents referred to decisions on the purchase of new systems being pending, and it is clear that the 'new generation' of proprietary systems has tempted a number of repositories into computerisation for the first time. Local authority record offices were less likely to be using databases for cataloguing than national and specialist repositories, but the divergence is not great. Some major institutions have not gone down the database route, holding their finding aids in textbases (or in reality a mixture of both). There are major differences between the two approaches, which are explained in paragraphs 4.1-5

#### 3.7

The delay in the adoption of digital cataloguing and indexing has been as much due to the absence until recently of satisfactory proprietary products geared to the needs of archives as to the lack of resources, but the resource constraints are undoubtedly real and keenly felt. Many repositories, but especially the smaller and less well-funded ones, will not be able to afford proprietary products in the foreseeable future if they have to rely exclusively on the resources made available by their parent body.

### 3.8

The Council was surprised to find that as many as 27% of respondents were undertaking image capture, although a good many noted that this work was still experimental. The university sector was clearly most advanced in this field, with nearly double its proportionate share of the number of positive responses. Only about one in five local authority repositories and one in seven specialist repositories had undertaken work in this area, compared with 50% of universities and 35% of nationals. Repositories were not asked whether they were undertaking image capture as a routine part of their cataloguing process, but no repository mentioned doing so, and it seems unlikely. The present lack of recognised technical or documentary standards for the inclusion of images in archival catalogues in any case makes such initiatives of doubtful long-term or networking value.

### 3.9

The second part of the questionnaire focused on use of the Internet, to determine the existing accessibility of information about repositories and their holdings, and how aware repositories were about the potential for remote access. 62% of respondents had access to the Internet, but this figure includes both some who had mounted pages on the server of a host organisation but did not have direct access to the Net themselves, and some who had access for searching, but did not maintain a web site. A further 26% expected to have Internet access within two years. Only 12% had no current plans for Internet access, and 98% of respondents saw the Internet as a desirable part of their future strategy. That suggests that the profession as a whole is very alive to the possibilities presented by the Internet, and an openness to the idea of making use of it. Internet access was not evenly spread through the profession, however. Only 47% of local authorities and 55% of specialist repositories had access, compared with 93% of nationals and 97% of university repositories. 15% of local authorities had no plans for Internet access. Clearly, this reflects the capital famine in local authorities in recent years, and is a measure of the extent to which local authority facilities have fallen behind those available to other types of repository.

### 3.10

The survey asked some technical questions about Internet linkages, which were again important for assessing the readiness of the profession to participate in a national network. 36% used a modem link for access, mostly at 28.8 kbps, 25% used a faster ISDN line, while 19% connected via JANET and 20% via a leased line. 45% had access via a commercial Internet Service Provider and 55% via their own web server. 83% of those with Internet access maintained a home page about the archive service, although generally this was mounted on the server of a parent organisation, rather than on a server being maintained directly by the repository. 82% gave details of their access arrangements, 46% gave at least some collection-level information about their holdings, and 22% provided access to word-processed catalogues or catalogue databases. The answers to other questions suggest that this on-line access must often be on an experimental basis: for example, a Midlands local authority repository has one small section of a Guide available through its web site. 88% of repositories with Internet access used it to send and receive e-mail, and 85% used the Internet to find details of other collections. Only 33% of those with Internet access allowed the public to share this facility, and it was clear from additional comments that in many cases where such access was given, it was via either a dedicated terminal or a network that was delivering primarily non-archival information, such as a library catalogue. A number of repositories commented that they anticipated providing public access to the Internet in future.

### 3.11

Many repositories' Internet access arrangements had involved work by both their

parent organisation and members of staff. 9% had employed paid consultants to set up their web site. A few respondents also referred to voluntary consultants, such as project work by ICT students, or co-operation with enthusiasts in Friends' organisations or local family history societies. Relatively few repositories seem to have in-house skills in the design and mounting of web pages, or an independent web-site. A number of respondents commented that because their web pages were part of a larger local authority or university site, there were restrictions on the style, content and size of their contribution, which they were obliged to work within. Overall, 59% of respondents cited reasons for their not having Internet access, or for not having developed it more fully than they had. Lack of resources featured in 91% of these responses, and were especially prominent in those from local authorities. 39% cited lack of accessible skills, 22% lack of understanding and 13% discouragement by their parent organisation.

### 3.12

The third part of the questionnaire concerned the standards for cataloguing and indexing which have been promulgated in recent years. The Council sees the widespread adoption of these standards as best practice for the profession, and as essential to the effective operation of a national cross-searching network, since without some consistency about the nature of data provided, searches performed across a number of databases will yield inconsistent results. The questionnaire asked about these standards using the acronyms by which they have become known - ISAD(G) [NOTE 27: International Council on Archives, *ISAD(G): General International Standard Archival Description*, ICA, 1994; currently under review.], ISAAR(CPF) [NOTE 28: International Council on Archives, *ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families*, ICA, 1995], and the NCA Rules [NOTE 29: National Council on Archives, *Rules for the construction of personal, place and corporate names*, NCA, 1997]- and an alarming number of people did not know what the acronyms stood for, let alone whether their finding aids conformed to these standards or not. 33% of respondents did not know whether they met ISAAR or the NCA Rules, and 23% were similarly uncertain about ISAD(G). On the other hand, about a quarter of repositories believed that their present listing practices conformed to ISAD(G), 10% felt they met the NCA Rules for personal and place name indexing, and 3% were using ISAAR(CPF). These findings highlight the need for an extensive training programme among professional archivists on standards and the potential of ict systems, and the Society of Archivists has responded to this by working with the Universities of London and Liverpool to offer appropriate courses, albeit on a limited scale, during 1998.

### 3.13

The final part of the questionnaire concerned repositories' willingness to participate in a national archival network if one was created. It was very encouraging to find 95% of respondents saying they would wish to participate in some form, and only two respondents definitely that they would not. In answering questions about the sort of data they would be willing to contribute to a national network, 93% of respondents said they would be willing to contribute a home page, and 89% that they would give collection-level data, but only 45% offered database information and 36% access to images. In the light of the additional comments people made about this question, it would appear that the responses were strongly coloured by the existing circumstances of repositories - this, after all, is what the rest of the questionnaire was about - and that when 55% of would-be participants said they could not offer database information, this was a comment on their present lack of data in digital form and their inability to see that this situation would change in the near future than an unwillingness. Having said that, a number of respondents did voice concerns about

the publishing of valuable copyright information on the Net, and this was more frequently mentioned in connection with the supply of images.

### 3.14

There is an enormous diversity in the hardware and software available and in use in different repositories, and in the availability and adequacy of network connections. To bring every repository up to the standard where they would be able to participate in the national network will require large-scale investment in equipment and training costs. Simply equipping repositories to begin creating their catalogues and indexes in digital form from now onwards would not alone be sufficient to create a useful national network, however, as in most repositories the majority of material (often including the most important collections) is described in existing manual lists prepared over the last century or more without any standardisation of layout or content. Work to transfer the vast backlogs of existing catalogues into machine-readable format will represent much the largest cost of creating a national archival network, albeit one that can be spread over a considerable number of years.

## 4 DATA STANDARDS FOR A NATIONAL NETWORK

### Relational databases and EAD

#### 4.1

The storage of data in digital form is most commonly done either as structured data or as free or structured text. Where it is stored as structured data the most flexible approach is to use a relational database. The data is held in fields which are grouped into linked tables. Because the data is stored in a structured manner it is possible to search the data in a variety of ways and to generate output for presentation on screen or paper. An alternative to using a structured database is to store the data as free or structured text. Text database applications are particularly suited to the storage, indexing, manipulation and searching of lengthy free-text descriptions. Some proprietary systems use the latter approach.

#### 4.2

These database systems have several advantages:

- they are based on common software products, which many programmers are familiar with;
- simple input templates can be created;
- they are relatively cheap;
- it is quick and easy to amend data already entered;
- text-based systems have good text handling and indexing facilities.

Conversely, there are some disadvantages:

- some relational databases cannot handle large text fields;
- all rely on commercial software products which may cease to be supported in the future (although adequate import-export facilities are now commonly included in such systems that allow backups to be made in standard, non-proprietary and easily transferable forms);
- some apparently excellent systems have not yet been tested on huge data volumes;
- some database systems have difficulty in adequately displaying the hierarchical structure of archival data.

#### 4.3

Storing data as free text often sacrifices much information about its composition. One solution to this problem is to add internal markers to the text to flag the meaningful units of information and their meaning. The fundamental standard for this kind of encoding is based on Standard Generalized Mark-Up Language (SGML). Once encoded, software products can control the searching, retrieval and structured display of the finding aids. SGML is a set of rules for defining and expressing the logical structure of documents; in other words a means of designing output formats, and encourages consistency of markup by introducing the concept of 'document type definition' which prescribes an ordered set of SGML tags available for encoding the

parts of documents in a similar class. Encoded Archival Description (EAD) is a document type definition for archival finding aids. The use of EAD is increasingly coming to be seen as a standard in the USA and is being adopted in the UK by a number of institutions, including the Bodleian Library, the Universities of Durham, Glasgow and Warwick, and the Public Record Office, while other archives, including the National Library of Wales, are showing interest in adopting it as a standard. Other countries in Europe, however, are less keen, and it has been specifically rejected by archivists in France and Germany, possibly because it is seen as too Anglo-centric.

#### 4.4

EAD also has advantages and disadvantages. On the positive side:

- it is not software dependent; it simply consists of encoded text files (although proprietary software is used to assist with encoding and with the display of the text);
- it is excellent at displaying the hierarchical structure of archive data;
- it is easy to use.

On the other hand:

- it has not yet been tested with really large volumes of data;
- the search facilities it offers require further development
- it may be time-consuming to implement (although encoding can be done automatically by generating output in EAD format from a database)

Furthermore, current Web browser technology cannot take advantage of text marked-up using EAD, and additional software must be downloaded to enable it to be read across the Internet. If EAD marked-up finding aids are to be effectively distributed over the Internet there is a need for suitable application modules to be developed and made freely accessible which can handle EAD. This problem may be resolved in the medium term as the mark-up language XML is adopted on the World Wide Web. There remains the issue of the 'broadcastable unit', i.e. how much of a document to transmit in response to a query, although in practice document sizes in EAD are usually fairly small, so the whole document is normally transmitted.

#### 4.5

The Public Record Office will use a hybrid solution for its ambitious project to get all its finding aids into digital form by 2001. Catalogue data will be stored on a large relational database, giving the PRO the power and ease of updating afforded by such systems, while output for users will be generated in SGML/EAD, where the display facilities of EAD will enable the hierarchical structure of archives to be expressed. The suppliers of one of the most widely adopted proprietary systems, CALM 2000, have recently adapted their product to import and export data to and from the EAD format, using Z39.50. Thus an EAD document may be broken down into component records which are capable of being manipulated efficiently in an integrated archive management system without losing the structure of the collection. CALM 2000 is completely compatible with ISAD(G) and record structures follow the ISAD(G) standard. ISAD(G) fields can be directly mapped to EAD so it is possible for CALM users to conform to both standards without altering their methods of working. The CALM 2000 record structures are also designed to be very flexible and changes in standards can be easily accommodated. An Internet display format is also being developed for another of the modern proprietary systems, SEAX, and work is in progress to enable the CAIRS system to generate output in EAD format.

4.6 The diversity of systems in use and the potential of EAD as the output standard will need to be taken into account when considering the architecture of a national network.

### **The need for data standards**

#### 4.7

The archival profession in the United Kingdom developed until the late 20th century without engendering any substantial body of professional theory. Its practice was largely based on the experience of heads of repository, who in some cases codified this practice to establish a house style. However, few if any repositories maintained listing practices designed in the circumstances of the 1930s or 1940s unchanged into the present day, and most repositories are able to show several generations of professional practice. This steadily increasing inconsistency in finding aids occasionally produced irritation in users, but seems to have had little professional impact until, in the late 1970s and 1980s, interest in the application of computer technology to archival listing and indexing stimulated a more theoretical interest in the nature and structure of archival finding aids. For a decade or more, the principles and concepts first established by Michael Cook in his *Manual of Archival Description* have been taught in the archive training schools, and a generation of archivists is thus emerging who are both familiar and comfortable with a more theoretically-based approach to cataloguing.

#### 4.8

In the 1980s, the imminent prospect of using relational databases for cataloguing made it necessary to consider how the finding aids of an individual repository might be consistently structured, so as to allow their input into a fielded record. At the same time, it was realised that much of the apparent diversity of traditional cataloguing practice between repositories was due to different traditions of layout and presentation, and that the information content of catalogue records was similar, even from repositories dealing with very different kinds of archives. This opened up the possibility of establishing a common structure for archival data between repositories, and of the future interchange of data between repositories. Work in the 1990s has therefore concentrated on developing national and international standards for archival data, and the standards proposed - ISAD(G), ISAAR(CPF) and the NCA *Rules for the construction of personal, place and corporate names* - have been accepted by many of the national institutions, but have yet to be widely adopted by the rest of the profession. Standards such as these are crucial for the interchange of data between repositories, and also facilitate the continued understanding of finding aids over time. The standards themselves will continue to see periodic amendment and improvement, but this will not invalidate their usefulness to the profession provided that each successive generation of standard is mapped to its predecessor, and proper records are kept by the profession of how and why the standards have been altered.

#### 4.9

Any expectation that technological developments will obviate the necessity for consistency of practice is illusory; unless data is consistently presented, the searching power of digital systems is hugely impaired, and their ability to represent the hierarchical structure of archival data compromised. The text-string or keyword searches which alone are possible on unstructured data, are valuable, but even with fuzzy matching and other probabilistic enhancements, are an inadequate substitute for the power and speed of dedicated field searches, especially as data volumes grow. For these reasons, it is recommended that the general acceptance and

adoption of these standards is an essential precondition to the successful development of a national network.

**Recommendation 1:**

**That the national and international data standards recognised by the Society of Archivists and the other professional bodies are generally adopted by archive repositories in the UK.**

This should not be taken as an argument for not implementing keyword searches, which are vital as a means of locating specific bits of information. A combination of keyword and defined-field searching will usually generate a better result than either approach could deliver on its own.

4.10

As noted above, the survey undertaken for the Networking Policy Committee in the summer of 1997 showed many archivists had not adopted the recent national and international standards - ISAD(G), ISAAR(CPF), and the NCA *Rules*. To some extent, the non-adoption of standards may simply reflect lack of awareness on the part of archivists. The Council recommends that training opportunities be made available to ensure that the UK archive profession is equipped to apply current international and national standards for archival data.

**Recommendation 2:**

**That training opportunities are made available by the professional associations to ensure that UK archivists are ready and equipped to apply the national and international data standards.**

4.11

Work on establishing an authority list of subject indexing terms is at a less advanced stage than the fully developed standards mentioned in the previous paragraph, but the work being done by Janet Foster on behalf of the JISC Archives Sub-Committee on this topic suggests that broad agreement can be reached on subject indexing too, at least for subject indexing at collection level. A thesaurus of around 500 terms is in preparation. Any commonality of approach which can be adopted across the profession on this matter will also help to ensure the consistency of cross-searching between the finding aids of different repositories.

4.12

The situation in Britain is in strong contrast with the situation in the USA, where in the mid 1980s an archival version of the library MARC standard was developed as an encoding standard for collection level descriptions (MARC-AMC). Standards for the content of the descriptions themselves were set out by S.L. Hensen in *Archives, personal papers and manuscripts* and these have been widely adopted. Encoded Archival Description was developed to provide a standard for rendering finding aids at all levels in digital form, and has made it possible for US archivists to develop large-scale joint on-line projects, like those described in paragraphs 1.8-9 above. It remains true, however, that the US experience has been most successful at the

collection level, and that some difficulties have been experienced (especially prior to the introduction of EAD) with attempts at item-level data interchange based on the MARC-AMC and APPM standards.

#### 4.13

Returning to the research carried out in 1997, a number of respondents pointed out that whilst their present practice was in conformity with one or more of these standards, they had large backlogs of older finding aids which were not, and which would require a substantial amount of revision or total recataloguing to be made so. It seems clear from the responses that a large part of the profession is still unclear about the nature and role of such standards, and is unwilling to make a commitment to them in the face of concerns about enforced uniformity of cataloguing practice, or the need for major revision of existing finding aids. They have, however, now been widely endorsed by the relevant professional bodies, and in the opinion of the Council, the adoption of these best practice standards will require a proactive approach to training but only manageable changes to current practice for most repositories. Their retrospective adoption can be undertaken as one element of the much larger task of converting existing catalogues to digital form. Obviously, the whole retrospective conversion process is not going to be cheap, but it needs to be undertaken over a period of time, and the creation of a national archival network may improve the strength of the arguments for funding it, both locally and nationally.

### **The content of a National Archival Network**

#### 4.14

There are two further issues relating to the data to be displayed in a national archival network: the level and volume of the data. Catalogues are hierarchically structured documents proceeding from the general to the specific. At the highest level, they provide information about the holding repository. Below this, there are collection-level descriptions, giving information about the provenance and content of a particular group of archives. At the lowest level, they describe individual items, or even parts of items. The higher the level in the collection, the less specific is the information. Clearly the long term aim must be to establish access from the Internet to the catalogues of the nation's archives at all levels, but it may not be feasible for all repositories to supply data in such detail from the beginning. The Scottish Archive Network will hold copies of the top-level finding aid of every repository in Scotland. This will be marked up to show whether more detailed catalogues are available on-line, and if so, there will be hypertext links to them. It is recommended that the UK national network aims for coverage down to piece level descriptions in the long term, but that repository and collection level descriptions should be acceptable initially if more detailed information is not available.

#### **Recommendation 3:**

**That the UK National Archival Network aims for coverage down to item/piece level descriptions in the long term, but that collection and repository level descriptions should be acceptable initially.**

#### 4.15

The need to progressively make available a large volume of data down to item level

will require a large retrospective conversion exercise. The Public Record Office has recently begun to convert its paper lists to digital form using an off-shore keying contractor. The PRO's experience is that there is a considerable cost in preparing lists for conversion. All lists have to be carefully marked up before they are keyed by the converters to ensure that data items (date, reference code, class title, headers, sub-headers and piece descriptions) are treated consistently. The more complex lists are marked up in house, while simpler ones are marked up by the conversion contractor overseas. The data is returned on magnetic tapes in text format files with delimited fields. It is first checked to ensure that the correct type of data is in each field and it is then uploaded into an SQL SERVER database. Each entry is then validated by a member of the PRO's staff. The major issues have centered around the complexity of the lists - which have been produced over 150 years and lack consistency. There have also been difficulties in developing a relational database system to check, validate, and store the list data. The database which the PRO is using has had to be specially designed to store varying types of data in a range of fields. The cost of data conversion is less than £1 a page, but the costs of marking up, checking and database development have just about doubled this. However, the PRO is willing to share its expertise and to allow other institutions to make use of its coding for its database and this should reduce costs in the future.

#### 4.16

The HMC estimate that the lists of holdings of British archives would extend to about 3,000,000 or 3,500,000 pages if they were fully catalogued. Based on the PRO's experience, retrospective conversion of this volume of lists would cost some £7,000,000, including marking up but exclusive of hardware costs. However, there is some evidence that this figure could be a substantial underestimate. Problems with the quality and consistency of inherited data are known to be very significant in many repositories; these include known, extensive and readily verifiable inaccuracies in older lists arising from

- poor palaeography;
- incomplete and inconsistent data, arising from a lack of adherence to standards in cataloguing;
- the inconsistent use of abbreviations, especially in personal names.

The need to ensure that data entered on to a national archival network conforms to the key national and international standards would appear to require substantial checking and improvement of this inherited data.

#### 4.17

Work done by Birmingham City Archives to quantify their retrospective conversion costs, suggested that many older catalogues (perhaps 20%) were beyond economic improvement, and that the staffing costs of improving and rekeying the remainder would be about £220,000 for 12,000 pages of text. The NRA estimate of a national total of 3-3.5m pages of text includes an allowance of 25% for uncatalogued material, for which by definition no lists are available. Taking out of the calculation this percentage, and a further 20% for that proportion of lists which Birmingham felt it would be uneconomic to convert, Birmingham's experience suggests that the total cost of a national retrospective conversion exercise would actually be nearer to £33.0-£38.5m. Whilst it is perhaps impractical to consider the *whole* of these works as an essential element in the creation of a National Archival Network, there would be little point in establishing such a network without also generating a substantial volume of data to sit on it. It is also true that much of the historically most significant material held by archive services around the country is described in these older lists,

which need to be converted to digital form if these key resources are to be fully visible on the network. The Council therefore recommends that a substantial programme of retrospective data conversion should be part of the project for establishing a national archival network.

**Recommendation 4:**

**That a substantial programme of retrospective data conversion should be part of the project for establishing a national archival network**

## **National Name Authority Files**

### 4.18

The *NCA Rules* are designed to be applicable as a standard for archival description in both manual and automated finding aids. However, many of those participating in drawing up the *Rules* assumed that they were merely a precursor to the creation of a National Name Authority File. As each repository proceeds with the automation of its catalogues, the names of persons, families, places and corporate bodies appear at varying levels of description. To achieve consistency both in description and in retrieval by users, it is necessary to copy and retain this data separately from the context of the catalogues. Each repository will build up significant quantities of name authority records, which should conform to ISAAR(CPF) as well as the *NCA Rules*. This data may be exchanged and shared by contributing repositories and merged into a single authoritative National Name Authority File. As a passive look-up file, the National Name Authority File would provide a valuable common cataloguing resource for archivists. As an active file, with hypertext 'hot links' to catalogues, it would form a valuable gateway to the National Archival Network.

### 4.19

With the assistance of a grant from the BLRIC, a consultant has been employed to investigate the preparedness of archivists to co-operate with the creation of national name authority files and the requirements and costings for establishing a central server to maintain and disseminate the authority files. The report indicates that nationally, record offices report a total of 23.3 million index entries, with six institutions reporting more than one million entries each. These six institutions account for just over 40% of the total entries. Rates of addition were frequently estimated in the same way as were the total numbers of entries. It would appear from the total figure reported (1.4 million p.a.) that there are either phenomenal growth rates or that the volumes have been greatly overestimated. Figures for index entry deletions (2,362) and amendments (21,917) are both quite low, probably indicating that for most institutions the emphasis is on recording and indexing new material, rather than revisiting entries already created.

### 4.20

The report estimates that the total number of personal and family names that might be contributed if all archives were able to supply their entries would be about 50 million, the number of corporate names around 25 million, and the number of place names around 5 million. The Council believes that these figures probably err on the high side, but as the real totals for the UK as a whole are unknowable, these estimates at least indicate the formidable challenge to data collection and maintenance, to say nothing of storage and searching, that the creation of National Name Authority Files represent. In reality, not all repositories will wish to contribute

all of their entries, and not all personal names are likely to be included. The Public Record Office of Northern Ireland (PRONI) catalogues contain about 200,000 personal names but yield only 10,000 entries for their Prominent Persons Index; this may be a useful guide to the proportion of index entries which might be worth including in a name authority file.

#### 4.21

The evidence collected by the report does not provide a strong case for a National Name Authority File as a passive look-up tool. There is evidence of such a requirement, but not enough to indicate a justification on its own for the effort and investment required. The report concludes that an active file would have a clearer justification. A user would see a list of the collections holding material related to the name in question, and would be able to move directly to the site(s) concerned by a hypertext link. This does not mean that pointers should be made only to sites that are accessible over the network; three possibilities are available: the link could lead straight into a digital finding aid at the site; the link could lead to a passive, directory-type site to give the user access information; or there could be no link, only the indication that a relevant collection was available. The 'hot links' could be brought in as sites became available. The active file approach presupposes the use of Web technology to deliver the National Name Authority File.

#### 4.22

Centralised data capture and conversion offers the opportunity to enforce common standards of format and content. However, it dissociates the process from the collections which have given rise to the entries. This raises the distinct possibility of meaning being lost in the conversion through lack of contact with the source material. Central data capture and conversion loads all of the costs on to one authority. Local data capture and conversion places the task adjacent to the source material, increasing the likelihood that the logical conversion will be done in a proper relationship to it. The major drawback is that the task will represent, for most repositories, a whole new and additional area of work for which they may not have the time or the staff and machine resources. Repositories already engaged in the conversion of finding aids to digital form may be able to produce fielded, consistent data with little additional effort. It is important to understand that many repositories do not already have the full range of name indexes as proposed for the National Name Authority File, not because they have no resources to build them, but because they are not needed. Name indexes, such as place names, have been created where there are particular needs for just that type of index. Creation of entries for a National Name Authority File may therefore be an additional workload for many archivists. The NNAF must be able to offer value in return for the effort. On balance, the most workable arrangement would seem to be to specify centrally the physical (format layout) and logical (content standard) requirements, and the nature of the machine-readable record to be generated, and invite local conformity with these among participants, but to encourage local data conversion and capture. The SCAN model, of providing equipment, software and training in order to encourage participation, would seem to be a good one to follow. Once data is received, the central organization must monitor its quality and apply any necessary corrective action. The central organization must offer back to its data suppliers local data sets, extracted from the NNAF as a whole, as part of a package that provides benefits to both parties.

#### 4.23

If an NNAF is built as an active file then it will, in time, take on many of the functions of the NRA. It will act as the central UK collecting organization for pointers to collections. It appears logical, therefore, that the NRA should take the lead in

developing, implementing and delivering the NNAF, and the Council therefore recommends that the Historical Manuscripts Commission should, with the participation of representatives from archival institutions with significant quantities of data in digital form, develop and implement an active National Name Authority File.

**Recommendation 5:**

**That the HMC should, with the participation of representatives from archival institutions with significant quantities of data in digital form, develop and implement an active National Name Authority File.**

## 5 THE TECHNOLOGY OF A NATIONAL ARCHIVAL NETWORK

### Alternative networking strategies: the piecemeal approach

#### 5.1

The Networking Policy Committee identified several different technological approaches which could deliver some or all of the benefits of a National Archival Network. One option has to be simply to recommend that all archive repositories wishing to make resources accessible to networked users join the Internet by establishing their own web-site or mounting data on the site of a parent organisation (the piecemeal approach). There could be some hypertext links between related sites, and a degree of connectivity could be provided by the HMC's repositories file, available through its Web server, with its already existing lists of contact addresses, supplemented where possible by links to each archive's home page. This approach would have the advantage of being relatively simple, and of employing easily understood technology such as the familiar web browsers. Access to services could be free or charged for as each repository preferred. Each site could be developed incrementally and each archive could develop in the way it wished. This is the sort of network which will come into being *de facto* if no collective planning or investment takes place.

#### 5.2

Unfortunately, the Internet is neither predictable nor reliable. The response times vary enormously depending on the level of traffic the system is carrying when a search is performed, and while e-mail messages almost always successfully traverse the network from sender to recipient, browsers and information seekers often find it difficult to locate and retrieve information in a timely fashion. Moreover, the corollary of the flexibility which a piecemeal approach would offer to network participants is that the network would offer only a extremely patchy representation of UK archives, with wealthier organisations possessed of more substantial ict services coming on stream first and less well resourced archives limping onto the network as they increasingly discover the disadvantages of not being connected. The look and feel of each site and the way in which information was structured within it would be different, and could probably not be effectively co-ordinated. There would be little to guide users towards relevant sources of information. Users would be provided with a variable picture of archival holdings across the range of repositories in the UK and would have to consult multiple sites in order to locate information about particular categories of materials: a comprehensive search might involve accessing hundreds of different web-sites. Repositories would use a wide variety of interfaces to their finding aids, and users would require both guidance and support in their use. The cost of local support of this kind would, of course, vary from place to place depending upon the nature of the public search screens, their ease of use, the quality of online help facilities, the structure of the finding aids, and the kind of network connection in use (e.g. dial-up or dedicated line).

#### 5.3

In such an environment, it would be difficult, if not impossible, to create inter-operable catalogues that would support a single point of access to all archival finding aids and provide such facilities as the ability for users to control the geographical focus of their search (e.g. to locate material on a particular subject in archive repositories near their home). While it would still be likely that as archives create new descriptions and retroconvert older finding aids they would apply international standards such as ISAD(G), the impetus for such consistency and for the rapid development of

improved finding aids would be significantly less than in a more structured environment.

#### 5.4

The other major problem with the piecemeal approach relates to future redevelopment. If each repository goes its own way then it is likely that the growing network will depend on different levels of infrastructure on a local level, put in place at varying times, reflecting advice from different classes of experts and not representing a single national objective. Once started down this path it would be difficult to renovate and upgrade the network to form a consistent infrastructure. Above all, this approach seems increasingly old-fashioned, given the power and sophistication of the web-searching tools which are becoming available. The NCA does not recommend the adoption of an unco-ordinated, piecemeal approach, but if no more sophisticated solution of the type described below can be funded, it believes that efforts should be made to co-ordinate the design of archival web-sites so as to make them as easy to use as possible.

#### **Recommendation 6:**

**That a piecemeal approach to the creation of a Network is not recommended, but that if the profession is obliged by funding constraints to proceed on this basis, efforts are made to co-ordinate the design of archival web-sites to make them as easy to use as possible**

Alternative networking strategies: Gateway approaches

#### 5.5

The preferred alternative to a purely Internet-based network growing up haphazardly would be an archival network, accessed via a central gateway from other networks such as the Internet, JANET, or the National Grid for Learning, or perhaps integrated with the National Grid for Learning, which appears to be developing as a more inclusive public sector information network than the first proposals envisaged. Such an approach would bring the benefit to potential users that the services and resources available across the network could be developed to a consistent standard both in terms of quality and format, and be delivered and accessed over a predictable and reliable network. The network might consist of a central gateway, transparent to the wider public networks, with links to regional and national servers holding data contributed by the participating institutions; some repositories with large volumes of digital data might link their own servers directly to the central gateway. The data volumes involved as the Network grows towards comprehensive coverage appear to preclude all the data being mounted on a single, national, server. Users would approach the central site, on which the National Name Authority Files could be mounted as active files, with hot links to relevant entries in the catalogues of individual repositories. The central gateway would also offer a more flexible search option, allowing users to construct their own searches, e.g. combining a place-name and a subject and a date range. There are two options for the manner in which such searches of the local sites would be performed from the central gateway, and these determine the structure of the system. These options are HTML searches and Z39.50 searches. The operation of the site would be invisible to the user, who would simply see it as another web-site.

## 5.6

In order to allow the performance of HTML searches across the network, each individual local site would have to use a common record template. This template would include tags which assign values to the documents for intelligent indexing and can enforce 'rules' of description or cataloguing schemes. The obvious template for use in archives would be ISAD(G). These templates would be implemented as HTML pages to be mounted on each participating web server. The central gateway site would have an advanced web-crawler which would search the network at intervals and drill down into web-sites to retrieve documents for indexing, and in this way build an index to the data mounted on the servers of the participating institutions. Users would follow links from index entries to the original records on the servers of individual institutions, which (despite using a common template) would be displayed in a range of different formats.

## 5.7

Two problems with this approach are that there would be considerable effort in creating the tags required to make the system function as required, and that it is not clear that adequate information retrieval tools exist to dig down into websites and index all the relevant names. It might be possible to generate the relevant tags automatically from local databases using proprietary software, or to derive them from SGML/EAD encoded documents. The introduction of XML as the basis for the Internet would simplify this issue. The network could then adopt EAD as a presentational standard and simply lead users via the gateway to the relevant EAD text on individual servers, which would then look much alike.

## 5.8

The feel of a Z39.50 search would be quite different. In this scenario, a user would access the central gateway site, and complete a standard search profile form. This query would be passed to a Z39.50 web gateway, linked to the regional and national servers. The Z39.50 server would be able to conduct searches across local proprietary systems and return the results to the user. The advantages of this approach are that users would see information from different sources with the same user interface and the same structure and terminology. Users would be able to query several (or all) archives at the same time, but there would be a feeling of a single co-ordinated virtual archive.

## 5.9

A Z39.50 search will, it is true, require work in the mapping of data and the establishment of servers. One Z39.50 site which already allows simultaneous searching across a number of databases is [Europagate](http://europagate.dtv.dk/) [NOTE 30: Located at <http://europagate.dtv.dk/>], which conducts standard searches on MARC library records. At the time of writing, the archival Z39.50 components are under development by JISC's National Networking Demonstrator Project for Archives. Once the outcome of projects like this and the application of the Cheshire system to the Arts & Humanities Data Service is known, a more mature assessment of the technical possibilities will be within the archive community's reach.

## 5.10

The preferred approach could be inclusive of small, medium and large institutions, all of which could be provided with the same level of support and with an infrastructure matched against predicted traffic, the functionality of local services, and planned development pathways. Issues of compatibility would be addressed in advance because they would be central to the provision of access to resources. A network for the entire community should be more cost-effective to establish and run than one built up piecemeal by different institutions, because the critical mass of members will make it feasible to lease lines in bulk, and to achieve cost savings in network

management. Some of these benefits could perhaps be increased further by participation in the National Grid for Learning. An independent network would, however, be technically easier to manage once it was in place: for example, it would be easy to use diagnostic tools and other metrics to measure performance, to isolate bottlenecks, and to reconfigure the network dynamically to meet changing patterns of use and need. Whilst the network would effectively be transparent to the users of other public networks, including the Internet, the participating repositories would have the benefits of control and support for archival services in a structured space. The specialist resources wanted by archivists such as name authority files and subject index thesauri could be held centrally and developed as an expanding national resource in a way which has been difficult in the past.

#### 5.11

Creating a structured network on the model of this preferred approach depends upon focused community action and considerable planning at both the pre- and post-network stages. Membership conditions would need to be established governing the obligations of members and the legitimate uses to which the network can be put. Standards and protocols are at the core of constructing a service of this kind. This includes infrastructure standards such as the kinds of routers or switches which will be used, network standards, such as whether the network will be TCP/IP based, and standards for the data contributed by individual repositories. Once established, the specification for the network would need to be monitored so that as changes and technological possibilities emerge the specification can be upgraded. In turn, once the protocols and standards have been agreed, members would be required to abide by and promote them.

#### 5.12

The NCA regards a system composed of a central gateway and regional/national servers using Z39.50 searching in tandem with active National Name Authority Files as being the most elegant means of establishing a National Archival Network. It represents a technically advanced solution, but the core structure of the network would be flexible enough to allow the system to expand and develop in response to technical advances. The network could also respond over time to increases in content availability, with the capacity of the servers and the leased lines being increased as more repositories begin to generate digital data, and to deal with the retrospective conversion of their backlogs. In a world where the archives profession is rapidly developing its utilisation of information and communications technology and the technical horizon is as short as 3-5 years, the flexibility of this approach seems particularly valuable.

#### 5.13

Individual archive repositories would supply data to the appropriate regional server (in England; the national server in Scotland, Wales, Northern Ireland and perhaps the Republic of Ireland), where it would be mounted on their behalf with an appropriate search engine. Data could be entered at collection level or item level, depending on the preferred strategy and data availability at each individual institution (but with the ultimate goal of mounting item level data for all collections), and be added to progressively as content availability increased. It would be essential for institutions contributing data to the network to do so in accordance with agreed standards, and ensuring that British archivists can do so is perhaps the greatest challenge facing a National Archival Network. A mechanism would be needed to ensure that agreed data standards were in practice being met (see paragraph 6.4 below)

#### 5.14

The main capital cost of adopting the preferred approach would lie in the establishment of the regional and national servers and the network connections

between them and with the participating repositories; the main professional change would lie in the progressive adoption of agreed listing and indexing standards like ISAD(G), and a common format for the presentation and transport of data, such as EAD. The NCA regards it as feasible that a National Archival Network could be established and made viable on the basis of the preferred approach, and therefore recommends to the archive profession that the establishment of a series of interlinked regional and national servers, connected to a central gateway, be adopted as the best way of establishing a national archival network at the present time, and as a way of creating a flexible structure that will be hospitable to future technical and professional developments.

**Recommendation 7:**

**That a series of interlinked regional and national servers connected to a central gateway be adopted as the best structure for a national archival network at this time.**

5.15

It has already been indicated that a great deal of research and experiment is currently in progress in relation to the networking of archival data. In advance of the results of some of this work, and particularly of the National Networking Demonstrator Project, it is perhaps premature for this report to stipulate in too much detail the mechanism for the implementation of the proposed national network. Nevertheless, the NCA recommends that work starts immediately on the creation of the nucleus of the network as a proving ground for the project.

**Recommendation 8:**

**That the suggested model for a National Archival Network should be tested by a limited scale implementation project.**

This study should build on the results of the National Networking Demonstrator Project by incorporating the use of Z39.50 cross-searching and the provision of common query and results templates, and also model the relationship between the central gateway and multiple regional servers, and the formation of National Name Authority Files by the export of index data from the regional servers to the central gateway. Recommendations are made in paragraph 7.3 about the management and funding of this first stage implementation. While progress will be limited by the current lack of resources, this activity is seen as being the springboard for further development, as and when funding can be identified.

### **The costing of a National Archival Network**

5.16

The costings detailed below assume that a National Archival Network would come to serve at least the 300 main collecting repositories referred to in paragraph 2.4 above. The costs associated with implementing, managing and running an archival network

along the lines set out above depend upon a range of factors: the level of service provided, the location of the connected institutions, the availability and nature of the local infrastructure, and the type of traffic. Of the 300 archives included in this discussion, 57 are based in higher education institutions, and these will have high-speed connections through their institution to the Internet already. If these existing connections were utilised, the network costs could be reduced, but there are significant potential benefits to creating an independent network which should not lightly be discarded. The figures given below therefore include the costs of connecting these repositories to an independent network.

#### 5.17

A number of measures could be used to determine the size and costs of providing the Network, and in any full procurement the formula for calculating the costs would need to take these into account. Archive services could be divided into three categories (small, medium and large) on any of the following grounds: financial size (i.e. income and expenditure); number of staff; number of users; or the volume of records held. For the purposes of this, fairly simplistic, analysis of the network costs, institutions have been divided into three groups based on a subjective assessment of the numbers of staff, users and records, income/expenditure estimates and predicted infrastructure. We have categorised 180 institutions as small, 74 as medium and 46 as large. In addition to this, the location of the institution is critical, as the costs of connecting to high speed networks will depend upon how far the institutions are from the nearest Switched Multi-megabit Data Service (SMDS) service point.

#### 5.18

The ideal network would be based around dedicated leased lines interconnecting all the UK institutions, but this model would be the most expensive to set up and probably to run. The costs of establishing a network in this way would only be justified if a sufficient level of traffic across the network was experienced, and it is difficult at present to predict the level of this traffic. Whilst this is the preferred solution, it is not the only possibility. An alternative would be a mixed network, which depended upon a set of core institutions connected by dedicated leased lines, and a larger body of smaller institutions with ISDN2 connections. This would significantly reduce the set up costs, but would leave the annual running costs uncontrolled, and potentially higher, depending upon the level of network traffic; it would also decrease the flexibility and robustness of the network. If the Government is successful in arranging minimal connection charges for institutions connected to the National Grid for Learning, and the National Archival Network was to form part of the NGL, it is possible that the costs identified below could be sharply reduced.

#### 5.19

In calculating the costs of these alternative approaches, the following costs have been included in the calculations:

- infrastructure at the local institutions
- leased line installation and annual rental charges/ISDN connections, based upon an average of 5 hours per day connection time
- annual support and maintenance charges
- replacement costs over five year.

To provide participating institutions with a minimum level of comparable computing facilities for both staff and visitors, and to offer an incentive for participation, the costings for local infrastructure include 3 PCs (one for staff and two for visitors) at

small institutions, 2 PCs (one for staff and one for visitors) at medium-size institutions, and 1 PC (for visitors) at large institutions. This allocation has been worked out on the basis that smaller institutions are likely to find it more difficult to participate, and will require a bigger incentive to do so. Each institution would receive support in the configuring of their PCs, and would be supplied with suitable routers to permit the connection of their local area network (LAN) to the National Archives Network; smaller institutions would receive support for the interconnection of their local PCs through the development of a local LAN. In addition, it is estimated that to make this equipment useful, about £500 of software for network access, and WWW software will be required. The total cost of putting the necessary infrastructure in place would be around £6.1 million.

#### 5.20

The costs of creating the National Archival Network based upon dedicated leased lines offering a 2 megabit per second level of service would be £2.7m for the initial connection costs, with annual running costs of around £7.2m. These estimates are based on the rental of leased lines. The annual costs include £6.4m p.a. for the network costs and £0.8m p.a. for annual maintenance and support costs; a very rough figure to provide suitably trained staff to provide guidance on network usage and to support and manage further technical development. The annual network costs include an annual charge for each line, and for the distance between sites (at £3,290 for the first 15 km and £310 per km above that). The Networking Policy Committee has estimated the cost of the network using a worst case scenario where 150 of the institutions are an average of up to 100 km from the nearest service point; producing an annual line rental charge of £4.6m. If the average distance proved to be much smaller, say 50 km, then the costing would be significantly reduced, the annual line rental charge coming down to £3.3m, and the total annual running costs being reduced by a similar amount.

#### 5.21

The alternative to using dedicated leased lines would be to connect all the participating institutions to the ten nodes identified in paragraph 5.22 by ISDN lines. These would incur charges per minute of use during the day, and if one assumes an average connection time of 5 hours per day per institution, then the total costs of running the network would be around £2.1m p.a. including networking costs of £1.3m and annual maintenance and support costs of £0.8m.

#### 5.22

Although this model could support local servers at each institution holding the finding aids for that institution, this would not be practical because of the increase in the support cost overheads and the fact that to deliver responses to searches across the network within an acceptable time it will be essential to reduce the number of servers that have to be searched. The Networking Policy Committee has recommended the creation of about ten geographically dispersed nodes, taking advantage of existing local authority or higher education institution infrastructure provision where possible. For the purpose of the costings in this report, it has been assumed that these nodes might be located in Aberystwyth, Belfast, Birmingham, Cambridge, Edinburgh, Exeter, London, Manchester, Newcastle and Winchester. In practice, the optimum locations would need to be determined in relation to a variety of political, data management and costing factors. These sites would form the backbone of the archives network, provide the fast server technology and gateways to external networks such as the Internet. Each of these sites would either be equipped with either servers provided by the project or depend upon storage and processor time rented on a local server. In the worst case, that of providing a server and router at each of these ten locations, it is estimated that the cost of setting up the system

would be in the order of £30,000 per site; a figure which includes hardware, network server software and the necessary database tools. The maintenance would be around £4,500 per site per annum, and the support costs for the entire network would be around £100,000 per annum. This would result in annual charges for the ten nodes of around £145,000 per annum. The cost of digital storage has fallen every year for at least the last 30 years and will continue to do so for the foreseeable future. In 1998 a conservative estimate of the cost of reliable online disk-based digital storage is about £40 per gigabyte per annum. This means that the cost of actually maintaining finding aids online, and later even adding images of archival material will become increasingly affordable over time.

#### 5.23

If the option of using dedicated leased lines were to be implemented, the total cost of establishing a properly resourced and supported network would be around £9m and the annual charges would be between £5m and £7.2m. In addition, the Networking Policy Committee recommends that 20% of the initial hardware costs, excluding set-up costs, should be set aside each year to support the replacement of the infrastructure at the end of the first five years. This is a figure of about £0.5m per year. This would increase the running costs to between £5.5m and £7.7m per annum. As noted above, these figures are obviously based on a number of assumptions, which probably err on the side of pessimism; as a result the actual costs could be significantly lower.

#### 5.24

If the alternative option of using ISDN lines to connect all the participating repositories to one of the ten main Network nodes were adopted, the total cost of establishing a properly resourced and supported network would be around £6.1m and the annual charges around £2.1m, to which once again replacement provision should be added, in this case around £0.5m p.a., taking the running costs to around £2.6m per annum. Of course, the pattern of usage of ISDN would be different and the available bandwidth much lower than would be available with the leased-line option.

5.25 Both these alternative costings are based on the assumption that all of the 300 institutions whose participation in the National Archival Network is d participate from the beginning. In practice, it is anticipated that a core of institutions will be ready to join the Network at its inception, and that others will join over time, as they acquire systems for the digital cataloguing and indexing of records, and begin to generate data in a format which can be contributed to the Network. The traffic carried by the Network will depend very much on how fast the network grows, both in terms of participation and in terms of the data volumes held for each repository. Two critical points are likely to be reached when the Network is widely perceived as a useful tool for resource discovery, and when it is accepted as the best and most comprehensive tool for resource discovery. In order to enable a more accurate modelling of the network costs, it is recommended that research should be undertaken into the availability of digital catalogues for inclusion in the Network and the readiness of individual repositories to participate in the project.

#### **Recommendation 9:**

**That research should be undertaken into the availability of digital catalogues and the readiness of repositories to participate in the project.**

## 6 THE MANAGEMENT OF A NATIONAL ARCHIVAL NETWORK

### Management structures and controls

#### 6.1

The implementation of a National Archival Network would need to be the responsibility of a single body that was responsive to the collective will of the participating repositories, and which was constituted in such a way as to maximise its freedom to operate and to receive grants from other agencies. In practice, this is likely to mean that it should be constituted as a company limited by guarantee which is also registered as a charity, the Directors and Trustees of which are nominees of the professional associations and the main national institutions. The management association of the network would need to develop the network, oversee its implementation, and evaluate its success or failure.

#### 6.2

It is probably desirable that the actual work of creating and running the Network should either be carried out by a single network service provider or by a franchising arrangement, because these options limit the financial risk to the archival community in general and to the management association in particular, by creating a situation where the costs are fixed and subject to stringent contractual terms. The various contracts which would need to be issued to develop the network would then be managed by a series of service level agreements. These documents would need to cover matters such as 'up time', guaranteed fix times, levels of and types of support to be provided to both archival staff and users, and would clearly define who holds responsibility for different aspects of the service, such as servers, routers, wires, software and data. Evaluation will be the key tool in monitoring how the service providers are fulfilling their obligations under the SLAs.

#### 6.3

The identity of the contractors would depend upon the precise network model adopted. A single network supplier might be interested in providing the central gateway, the communication lines, the network of regional servers in England, and the technical support and development functions; alternatively, it might be more straightforward to negotiate separately with individual institutions for the rental of server space and processor time for the regional nodes and the central gateway, and to rely on a network supplier simply for the communications lines and support functions. Choices in this area would be a matter for detailed consideration by the managing agency, and may be usefully informed by the proposed first stage implementation project, and the proposed research into data availability. The starting point might well have to involve a fairly fragmented management structure, with some elements perhaps being supplied directly by the management association; the pattern of provision could mature with the network

#### 6.4

Monitoring compliance is one of the key roles that the management association would have to carry out in the post-establishment phase. Other key activities would include support for users, the creation of connections to other networks and future enhancement and replacement strategies. It would liaise with the partner institutions themselves, and might take some responsibility for the delivery of content to the network, by working with or on behalf of partner institutions to secure grants for retrospective conversion and/or cataloguing projects. It would need to police the quality of data submitted to the network, to ensure that individual institutions were meeting their obligations to conform to the agreed data standards. Some of these

functions would involve the employment of staff, and might again be discharged through contracts. Activities such as the provision of support and maintenance could well be covered under the terms of the SLAs. Other aspects of support would only be possible if provided by the archives profession itself as they will relate to the data and not to the technical infrastructure. These might be provided by awarding a contract to act as 'data manager' for the network to an existing professional agency. The management association would retain a role as an arbitrator in the case of disagreement between partner institutions and the data manager about adherence to standards.

## **Funding a National Archives Network**

### **6.5**

Having arrived at an outline of what a National Archival Network might achieve, how it might be delivered in professional and technical terms, what it might cost and how it might be managed, it remains to address the issue of how it might be funded. The benefits of the network will be felt to varying degrees by all those involved in the provision and use of archive services, and indeed by Society at large, for the reasons set out above. Broadly speaking, it is felt that the costs of providing the network should be shared in a similar way. It has, however, been a long and bitterly defended principle of the archive profession that access to archives should be free to all comers at the point of delivery. Despite the absence of any legal prohibition on charging and the very considerable pressures on the budgets of archive services in all sectors of the profession, there have been very few instances of charges being introduced and most of these have been withdrawn after the immediate crisis has passed. The NCA has consistently supported this position, and notes with pleasure that the present Government has adopted in relation to museum admission charges a policy consistent with its belief that access to archives should not be conditional on ability to pay. Accordingly, the Council does not wish to recommend a method of funding the National Network that involves charges for searches levied on individual researchers.

### **Recommendation 10:**

**That searches in the National Archival Network should be free at the point of use.**

### **6.6**

The cost of constructing the infrastructure of a National Network has been estimated at between £6.1m and £9m, depending on the nature of the structure and the volume of traffic carried by the Network. These capital costs are well beyond what the archival community might fund by collective action, and they also fall outside the remit of the Heritage Fund, as set out in its forthcoming guidance to applicants. The NCA therefore recommends that individual repositories, perhaps acting collectively, should concentrate on the purchase of the necessary software for generating digital catalogues. It considers that the infrastructural costs of creating the Network will probably only be found by bringing together funding from a variety of sources. These might potentially include the Heritage Fund, European Union telematics programmes, the sort of sponsors which the Government believes would be willing to invest in networking public libraries, and the education sector. As has been noted above,

some of the infrastructural costs might be reduced or avoided if the Network formed part of the National Grid for Learning. The Government itself could play a valuable co-ordinating role in establishing a suitable funding package to enable the creation of the Network.

**Recommendation 11:**

**That the Government should co-ordinate the funding of a national archival network from a basket of sources, including the Heritage Fund, the European Union, the private sector, and the education sector, and that individual repositories should concentrate on funding the purchase of cataloguing software.**

6.7

The revenue costs of the National Archival Network have been estimated at between £2.6m and £7.7m a year, depending on the volume of network traffic and the structure of the network. Again, these costs are well beyond what individual repositories could afford to fund, although the NCA recommends that repositories should make some annual contribution to these costs, in recognition of the benefit they derive from the inclusion of their finding aids in the Network and the benefits their users derive from high-speed, online access to the network in the searchrooms of individual repositories. Repositories may be able to recover most if not all of these costs through increases in the volume of value-added services for which charges are already made, such as research and copying, brought about by the increased accessibility of their finding aids on the Network.

**Recommendation 12:**

**That individual archive repositories should contribute an annual membership fee to the National Archival Network**

6.8

The main beneficiaries of the creation of a National Archival Network are, of course, the existing and potential users of archives, but the determination to ensure that the network remains free at the point of use means that no system of access charges for individual users can be applied. The NCA feels that the costs of a National Archival Network should be spread broadly through Society, as the benefits will be felt very widely. Contributions towards the revenue costs of the network could be made from advertising revenue, from the New Opportunities Fund and the Heritage Fund, and from the education sector - tangible recognition of the dependence of the higher education community on research sources held in archive repositories both within and outside the higher education sector. Establishing a stable basis for the funding of the revenue costs remains very difficult, and is one of the tasks which the NCA would wish the proposed Resources Sub-Committee (see paragraph 7.4) to undertake.

**Recommendation 13:**

**That the establishment of a stable basis for the funding of the annual revenue costs of the National Archival Network, without the application of charges to end users, is referred to the proposed Resources Sub-Committee.**

6.9

The key importance of enabling archival institutions to convert their existing manual finding aids into digital form has been stressed several times in this report. On creating and sustaining the momentum for a programme of retrospective conversion of this kind depends the ultimate success of the National Archival Network in moving from being just a tool for resource discovery to becoming the backbone of the UK archival community. Work of this kind is clearly within the remit of the Heritage Fund, which has declared its willingness to fund projects for the retrospective conversion of finding aids to digital form. It is also an area that falls within the scope of the £50m of the New Opportunities Fund earmarked for content development for the National Grid for Learning and library networking. The NCA therefore recommends the Heritage Fund and the New Opportunities Fund work in partnership with repositories and the management association of the Network to maximise the volume of data available for the National Archival Network, and thus its value to the users of archives.

**Recommendation 14:**

**That the Heritage Fund and New Opportunities Fund support projects in the archival sector to achieve the retrospective conversion and improvement of manual finding aids for inclusion in the National Archival Network.**

6.10

The final element of the costs identified by this report concerns the training of archive staff, both in understanding and exploiting the technology of the National Archives Network and in the use and observance of the data standards described in section 4 above. Training programmes on data standards are already underway under the aegis of the Society of Archivists and the Universities of London and Liverpool, but will need to be delivered much more widely if all archivists in the main collecting institutions are to receive training. A comprehensive programme of ICT skills training will also be required if the full potential of the National Archival Network is to be properly exploited. The NCA therefore recommends that the Government should expand the programme of ict training for librarians and teachers being funded by the New Opportunities Fund to include the training that archive staff will require; this is estimated to cost £1.5m.

**Recommendation 15:**

**That the Government should extend the programme of ict training for librarians being funded by the New Opportunities Fund to include the ict and data standards training archivists will require.**

## 7 DEVELOPMENT AND IMPLEMENTATION PLAN

### 7.1

This report has sought to demonstrate the desirability of creating a National Archival Network if Britain's record offices are to fulfill their potential role in our nation's culture and in a national information strategy. It has identified that the construction of such a network is technically feasible, but that there are significant financial, technical and content development obstacles to be overcome. A National Archival Network will not happen by itself, except in the messy, inadequate form of each repository mounting some information about itself and its collections on the Internet. This section of the report lays out a plan and timetable for the development and implementation of a National Archival Network over the next three years. Actions to be undertaken by the National Council on Archives will be performed on its behalf by the Networking Policy Committee of the Council established in 1996, the remit of which was changed to this effect by the Council at its meeting in April 1998. It will in future be called the Network Implementation Committee (NIC).

### 7.2

Following the publication of this report, the National Council on Archives will lead a publicity and consultation initiative through its member organizations, designed to raise awareness and support for the proposals among the archival profession, archival institutions, the users of archives and related information specialists. A seminar will be held jointly with the Joint Information Systems Committee on 18 June 1998 to compare our proposals with other initiatives that are currently under development, and to assess what opportunities for collaboration there may be. The members of the NIC will seek opportunities to speak about the Network at professional and user group meetings around the country during the summer, and a number of invitations of this kind have already been received. Comments on the report are invited from all interested parties by 30 September 1998.

### 7.3

During the summer, the NIC will establish a Technical Sub-Committee to manage the First Stage Implementation Project (FSIP) outlined in section 5.15 above. This sub-committee will seek to draw in as members individuals from within or outside the archival profession who can contribute to the management or evaluation of the project. The FSIP will build upon the work of the National Networking Demonstrator Project and will model specific aspects of the network, including the searching of data on multiple platforms from a common gateway site, and the relative advantages of different models of grouping data from participating institutions. The Public Record Office and the Scottish Archives Network, based at the National Archives of Scotland, have agreed to participate in this first stage implementation, and will both be hosting data from local authority repositories as well as their own data. The Royal Commission on Historical Manuscripts will also participate, through the development of a National Name Authority File. These bodies will each fund their own involvement in this project themselves. The Higher Education sector is also expected to participate through the JISC Archives Sub-Committee, by providing a third node. It is envisaged that the FSIP will be commenced on 1 October 1998, and the results will inform the specification for the provision of a national network which will be prepared in parallel. The Technical Sub-Committee would also have responsibility for continuing to monitor technical developments and new initiatives involving archives, both in Britain and around the world.

#### 7.4

The NIC will also establish a Resources Sub-Committee, which will locate the content and resources necessary for the National Archival Network to be successfully created on an inclusive national basis. The Resources Sub-Committee will draw in as members individuals from within or outside the archival profession who can help with this task, in addition to members of the full committee. The sub-committee will carry out a number of tasks during the coming 18 months. It will seek funding for and manage a detailed survey of existing digital catalogues to assess the likely scale of digital catalogue data that could be included in the Network. This survey will be hosted by the Royal Commission on Historical Manuscripts on the Council's behalf. It will commence on 1 October 1998 and be completed by 31 March 1999. Alongside the conduct of this survey, the Sub-Committee will seek the 'in principle' support of archival institutions to participate in the network, on the basis of the model of the Network outlined in sections 4-6. The Sub-Committee's main task, however, will be to discuss with potential funding agencies their willingness to contribute to the costs of establishing and running a network. Some indication of the order of magnitude of these costs is already available, but the survey conducted by the Sub-Committee, and the findings emerging from the Technical Sub-Committee will make a realistic timetabling of these costs feasible. The Sub-Committee's target will be to deliver to the NIC by 30 September 1999 the 'in principle' support of a sufficient number of participants and funders to make the National Archival Network possible.

#### 7.5

The NIC itself will steer the work of its sub-committees and concentrate its attention on the establishment of the structures necessary to deliver a National Archival Network. Informed by the consultation work undertaken during the summer of 1998, and by the reports of its sub-committees, it will design and set up a Consortium of national institutions and professional bodies to own and manage the network, and in due course register this body as necessary to give it a legal identity. It is envisaged that the initial membership of the Consortium would probably be similar to that of the NIC itself, and that the NIC might seek to co-opt additional members to ensure this was so. With legal advice, which the Public Record Office has agreed to supply, the NIC will also prepare contract documentation for the Consortium to issue to potential network providers.

#### 7.6

When the work of its sub-committees is complete, in autumn 1999, the NIC will launch the Consortium, which will make full and formal applications for the funding necessary to deliver a National Archival Network. When this funding is in place, it will initiate a tendering process among potential network providers, with a view to the commencement of full network installation before the end of 2000.

#### 7.7

While these inevitably somewhat lengthy processes are taking place, it will be open to individual repositories to prepare for their participation in the National Archival Network by acquiring hardware and software for the creation of digital catalogue data and by seeking funding to undertake the retrospective conversion of their existing manual finding aids into digital form. In order to ensure that such developments are undertaken in ways that will be compatible with a future National Archival Network, the Society of Archivists has initiated a programme of training in data standards for British archivists, and the Technical Sub-Committee will act on the Council's behalf in vetting project proposals that repositories wish to submit to it for certification that their outputs will be compatible with the data requirements of a future National Archival Network. Through its Resources Sub-Committee, the NCA will endeavour to secure commitments in principle from key funding agencies to the support of retrospective

conversion projects that are certified as being compliant with the objectives of a National Archival Network.

#### 7.8

The United Kingdom's documentary heritage offers a unique window on our national past which both helps to shape our future and improves the public understanding of our culture. Without the resources identified as needed in this report, the archives community will be unable to extend access to that heritage, and the appreciation of our past will become increasingly fragmented. On behalf of the archives community, the NCA urges the importance of creating a National Archival Network, to ensure that future generations can share and enhance the insight that we enjoy into our national past.

## 8 SUMMARY OF RECOMMENDATIONS

8.1 The following recommendations are made in the body of this report, and are summarised here for convenience:

### **Recommendation 1:**

**That the national and international data standards recognised by the Society of Archivists and the other professional bodies are generally adopted by archive repositories in the UK.**

### **Recommendation 2:**

**That training opportunities are made available by the professional associations to ensure that UK archivists are ready and equipped to apply the national and international data standards.**

### **Recommendation 3:**

**That the UK National Archival Network aims for coverage down to item/piece level descriptions in the long term, but that collection and repository level descriptions should be acceptable initially.**

### **Recommendation 4:**

**That a substantial programme of retrospective data conversion should be part of the project for establishing a national archival network.**

### **Recommendation 5:**

**That the HMC should, with the participation of representatives from archival institutions with significant quantities of data in digital form, develop and implement an active National Name Authority File.**

### **Recommendation 6:**

**That a piecemeal approach to the creation of a Network is not recommended, but that if the profession is obliged by funding constraints to proceed on this basis, efforts are made to co-ordinate the design of archival web-sites to make them as easy to use as possible.**

### **Recommendation 7:**

**That a series of interlinked regional and national servers connected to a central gateway be adopted as the best structure for a national archival network at this time.**

### **Recommendation 8:**

**That the suggested model for a National Archival Network should be tested by a limited scale implementation project.**

### **Recommendation 9:**

**That research should be undertaken into the availability of digital catalogues and the readiness of repositories to participate in the project**

**Recommendation 10:**

**That searches in the National Archival Network should be free at the point of use.**

**Recommendation 11:**

**That the Government should co-ordinate the funding of a national archival network from a basket of sources, including the Heritage Fund, the European Union, the private sector, and the education sector, and that individual repositories should concentrate on funding the purchase of cataloguing software.**

**Recommendation 12:**

**That individual archive repositories should contribute an annual membership fee to the National Archival Network.**

**Recommendation 13:**

**That the establishment of a stable basis for the funding of the annual revenue costs of the National Archival Network, without the application of charges to end users, is referred to the proposed Resources Sub-Committee.**

**Recommendation 14:**

**That the Heritage Fund and New Opportunities Fund support projects in the archival sector to achieve the retrospective conversion and improvement of manual finding aids for inclusion in the National Archival Network.**

**Recommendation 15:**

**That the Government should extend the programme of ict training for librarians being funded by the New Opportunities Fund to include the ict and data standards training archivists will require.**

## APPENDIX 1

### MEMBERSHIP OF THE NETWORKING POLICY COMMITTEE

**Victor Gray**, Chairman of National Council on Archives; Head of Corporate Records & Archives, N.M. Rothschild & Sons Ltd. (Chairman)

**Ishbel Barnes**, Head of Private Records Department, National Archives of Scotland

**Heather Forbes**, Senior Archivist (Customer Services), Hampshire Record Office

**Nicholas Kingsley**, Secretary of National Council on Archives; Central Library Manager (Archives, Local Studies & History), Birmingham Central Library (Secretary and Compiler of the Report)

**Patricia Methven**, College Archivist, King's College London; Chairman of JISC Archives Sub-Committee

**Margaret Procter**, Records Manager, University of Liverpool

**Lesley Richmond**, Acting Director of Archive Services, Glasgow University Archives and Business Records Centre

**Seamus Ross**, Director of Humanities, Computing and Information Management, Glasgow University

**Dick Sargent**, Assistant Keeper, Royal Commission on Historical Manuscripts

**Keith Sweetmore**, Area Manager, West Yorkshire Archive Service; Archival Consultant, National Networking Demonstrator Project

**Virginia Teehan**, College Archivist, University College, Cork

**David Thomas**, Head of Information Records Department, Public Record Office

**John Watts-Williams**, Assistant Keeper, Department of Manuscripts and Records, National Library of Wales

## APPENDIX 2

### GLOSSARY

The following list of terms includes both acronyms and technical archival and ICT terms used in the text of this report. The definitions of technical terms are not intended to be technically rigorous but to provide the necessary understanding for the comprehension of this report. Terms in italics have a glossary entry of their own.

ACW	Archives Council Wales; a non-statutory body representing archives in Wales.
AHDS	Arts & Humanities Data Service; a body funded by <i>JISC</i> to collect, manage, preserve and promote the re-use of scholarly digital resources.
Bandwidth	A term used to describe how much data can be sent through a <i>network</i> connection, measured in megabits per second.
BL	The British Library
BLRIC	The British Library Research & Innovation Centre
Broadband	A transmission medium capable of supporting a wide range of frequencies, typically from audio up to video frequencies.
Browser	Software used to access information from the <i>World Wide Web</i>
CAIRS	A proprietary <i>database</i> system used by a number of <i>repositories</i> for digital cataloguing of archival collections. Marketed by Leatherhead Information Technology Ltd.
CALM 2000	A proprietary archival cataloguing system, used by a number of <i>repositories</i> and marketed by DS Ltd.
Catalogue	A descriptive list of an archival collection, usually including entries at more than one <i>level</i> ; typically at <i>collection level</i> , <i>series level</i> and <i>item level</i> .
Collection level description	The description of the characteristics of an entire archive, such as the records of a company, or the papers of a scientist.
Database	A collection of data held in a digital environment in a

structured way, so that it can readily be updated and searched.

Data elements	Components of an archival catalogue description capable of being separately identified, such as date, reference number etc., and often recorded in separate fields of an electronic <i>record</i> to facilitate the searching of that record.
DCMS	The Department of Culture, Media & Sport, formerly the <i>DNH</i>
Dial-up services	Services accessed by using telephone lines or <i>ISDN</i> networks to connect a computer to a network such as the Internet.
DNH	The Department of National Heritage, now <i>DCMS</i>
DTD	Document type definition; a subset of <i>SGML</i> defined to give consistency to the formatting of similar documents; <i>EAD</i> is a DTD for archival catalogues.
EAD	Encoded Archival Description: an <i>SGML DTD</i> widely used by American archives for the presentation of archival catalogues in a digital environment.
eLIB	Electronic Libraries Programme; a Higher Education sector initiative funded by <i>JISC</i> to improve the range and quality of library services in the electronic age.
Fields	A term used to describe individual components of a <i>record</i> . In archival systems, fields commonly contain information about a single <i>data element</i> .
Gateway	A network node equipped to function as an interface between two networks, e.g. between the National Archival Network and the Internet.
Guide	A high-level overview of the archival collections of a repository, typically combining information at <i>collection level</i> and <i>series level</i> , and presenting the holdings of the repository in a structured order.
HEFC(s)	The Higher Education Funding Council(s) for England (HEFCE), Wales and Scotland
HLF	The Heritage Fund: one of the National distributors

HMC	The Royal Commission on Historical Manuscripts (also referred to as RCHM)
HTML	Hypertext Markup Language; the software language currently used to create <i>WWW</i> documents.
Hypertext link	On <i>Web sites</i> , an instant way of going to another site with related content, usually indicated by underlined text or an on-screen button.
ICA	The International Council on Archives
ICT	Information & Communications Technology
IMPRoS	A proprietary archival cataloguing system, used by a number of repositories, and marketed by MFT Computer Systems Ltd.
Index(es)	In an archival context, secondary finding aids, arranging in (usually alphabetical) order references to persons, places, subjects, etc. appearing in archival <i>catalogues</i> .
The Internet	A worldwide interconnection of individual networks operated by governments, industry, higher education institutions and the private sector.
Internet service providers	Companies that provide connections to the <i>Internet</i> , so that individuals and businesses can use e-mail and other Internet-based services. Internet Service Providers also market their expertise in the creation and administration of <i>Web sites</i> , training and the administration of <i>intranets</i>
Intranet	An intranet uses <i>Internet</i> conventions and applications over an internal, password-controlled <i>network</i> , typically within a company or other institution.
ISAAR (CPF)	The International Standard Archival Authority Record for Corporate Bodies, Persons & Families, issued by the <i>ICA</i>
ISAD (G)	The General International Standard Archival Description, issued by the <i>ICA</i> , outlining a set of <i>data elements</i> for inclusion in <i>collection level descriptions</i> of archives.
ISDN	Integrated Services Digital Network: an internationally

standard communications network supporting a *broadband* range of applications based on voice, image, text, video and data.

Item level description	The description of one or more individual items (e.g. a volume) in an archival collection. Such descriptions are normally combined with descriptions at <i>collection level</i> and <i>series level</i> to form a <i>catalogue</i> .
JANET	Joint Academic Network: the <i>wide area network</i> which links the UK higher education institutions, offering <i>broadband</i> connections to the Internet and other services.
JISC	The Joint Information Systems Committee of the HEFCs.
Leased line	A dedicated communications line, maintaining a 24-hour connection to a particular network. A single subscription is paid and the cost does not depend on the number of times or the length of time for which the connection is used. It therefore offers better value to high volume users, and is always available when needed.
LIC	The Library & Information Commission.
Local area network	A data communications system that links together personal computers in a limited physical area, typically within a single building; it may be, but does not have to be, connected to a public network.
MAD 2	Manual of Archival Description, 2nd edn. Publication analysing the typical structure of British archival description, and proposing some rules and conventions for the content of such descriptions, which although not widely adopted in full, have been very influential on the thinking of British archivists concerned with the development of standards.
MARC	Machine Readable Catalogue. An American library cataloguing standard, adapted for the description of archives as MARC-AMC (Machine Readable Catalogue - Archives & Manuscripts Control). Widely used in the USA but not much adopted in the UK.
Multimedia	A term used to describe the processing and integrated presentation of information in a variety of forms, e.g. video, sound, animated graphics and text.
NCA	National Council on Archives. A non-statutory body

representing the interests of the owners, custodians and users of archives in the UK; the publisher of this report.

NCA Rules	The National Council on Archives' <i>Rules for the construction of personal, place and corporate names</i> . A series of rules for ensuring common standards between <i>repositories</i> in the way names are represented in archival <i>catalogues</i> and <i>indexes</i> .
NGL	National Grid for Learning. A term coined by the present Government to describe a network delivering <i>multimedia</i> resources to schools for teaching and learning purposes, which the Government has since decided to extend to public libraries.
NNAF	National name authority file. A series of agreed standardised versions of names (place-names, personal names, corporate names) for use by archivists in indexing catalogue descriptions. In a hypertext environment, entries in such a file can be directly linked to catalogue entries and can function directly as index entries.
NRA	National Register of Archives. The <i>HMC's</i> service, which collects copies of the <i>catalogues</i> of archival collections from repositories all over the UK and produces indexes of the main places, persons and corporate names featured in them. The index entries but not the catalogue data to which they relate, are now available on the Internet.
NRA(S)	National Register of Archives for Scotland. A similar service to above, recording archival collections in Scotland, maintained by the <i>SRO</i> .
OPAC	On-line Public Access Catalogues. PCs connected to a networked catalogue, equipped with software designed to facilitate read-only access to the catalogue, to be used by the public with minimal assistance by staff.
PC(s)	Personal computer(s).
PRO	Public Record Office. The principal national archival repository for the UK, holding the records of national government and some non-departmental public bodies and former nationalised industries, except for records relating specifically to Scotland (see <i>SRO</i> ) and Northern Ireland (see <i>PRONI</i> ).
PRONI	Public Record Office of Northern Ireland. Performs similar functions to the PRO in relation to the government of

	Northern Ireland, but also acts as almost the only archival repository in Northern Ireland for non-governmental papers.
Record	A term used to describe a single entry in a database; it may be further subdivided into <i>fields</i> .
Repository	A term used to describe the places where collections of archives are held.
SCAN	Scottish Archives Network. A proposed network connecting the SRO and many other Scottish repositories. An Internet-accessible central server at the SRO will hold <i>collection-level descriptions</i> of archives in all the participating repositories and incorporate hypertext links to full <i>catalogues</i> of those collections if they are available on the WWW.
SCRAN	Scottish Cultural Resources Access Network. A project funded by the Millennium Commission to make available digitised images of items from museums, libraries and archives in Scotland on the Internet.
SEAX	A proprietary archival cataloguing system, used by a number of <i>repositories</i> , marketed by Essex County Council.
Series (or class)level description	The description of a group of records within an archival collection having a common nature and purpose (e.g. the ledgers of a company). Series level descriptions are usually combined with <i>collection level</i> and <i>item level descriptions</i> to form a <i>catalogue</i> .
Server	A central computer which provides some service for other computers connected to it via a <i>network</i> . A Web server is such a computer running a Website.
SGML	Standard Generalised Mark-Up Language. A software language used for encoding text; <i>HTML</i> , <i>XML</i> and <i>EAD</i> are limited subsets of SGML used for specific purposes. Not actually a programming language at all but a way of tagging documents so that they can be displayed in a pre-determined way on screen.
SLA(s)	Service level agreement(s). Contracts for the provision of services which specify in detail the nature and extent of services to be provided.
SRO	Scottish Record Office (now known as the National Archives of Scotland). The repository performing a similar range of

functions to the *PRO*, but also responsible for providing the *NRA(S)* and *SCAN*, and (as with *PRONI*) possessing large and important holdings of non-governmental archives.

UKOLN	The United Kingdom Office for Library & Information Networking which plays a leading role in planning and supporting technological developments in the library and information sector.
WAN	Welsh Archives Network. A proposed network to link the National Library of Wales and the other archival repositories in Wales in a similar way to that proposed for <i>SCAN</i> . Confusingly, the same acronym is also used for <i>wide area networks</i> .
Websites	A collection of documents in digital form, usually related to a particular topic and provided for use by <i>WWW</i> users by an individual or organization.
Wide area network	An interconnection of local area networks over a large geographical area, such as the network serving a local authority or large company.
WWW	World Wide Web. A collection of <i>Websites</i> which are provided from all over the world to be available for viewing by anyone else around the world with a suitable <i>Internet</i> connection.
XML	Extensible Mark-Up Language. A subset of the <i>SGML</i> encoding language designed especially for Web documents and allowing the authors of Web documents access to a greater range of functions than <i>HTML</i> . For example, XML supports hypertext links that point to multiple documents. It is expected that XML will eventually replace HTML as the main software language for Web documents.
Z39.50	An 'interoperability protocol' allowing searches to be performed across data held on a number of servers, irrespective of the format in which the data is held on each server.