

FEASIBILITY STUDY FOR A NATIONAL UNION CATALOGUE

**by
Peter Stubbley, Rob Bull and Tony Kidd**

Final Report

25 April 2001

Prepared for the Joint Information Systems Committee of the Higher Education Funding Councils (JISC), the Research Support Libraries Programme (RSLP) and the British Library's Co-operation and Partnership Programme

CONTENTS

Executive summary	2
1. The Study and its organisation	6
2. Guidelines for developing a UK National Union Catalogue	6
3. Discussion, survey and consultation.....	8
3.1. Discussions with key players.....	8
3.2. Questionnaire surveys.....	8
3.2.1. Survey of academic staff, researchers and postgraduates.....	8
3.2.2. Survey of librarians.....	9
3.3. Concertation Day	10
4. Experience with union catalogues	10
4.1. International union catalogues.....	10
4.2. North American experience.....	11
4.3. European experience.....	12
4.4. National union catalogues and the UK.....	12
5. Serials issues.....	13
5.1. Serials data and quality.....	13
5.2. Testing of serials	15
6. Tests and comparisons of systems.....	15
6.1. The challenges of technological change.....	15
6.2. Z39.50 and testing.....	16
6.3. Comparative testing of virtual and physical union catalogues.....	17
6.4. Comparative costs of virtual and physical union catalogues.....	19
6.5. Outcomes from system tests and comparative costs.....	20
7. A model for the UK National Union Catalogue.....	21
7.1. Recommended architecture.....	21
7.2. Monographs and non-serials formats	21
7.2.1. Content of the monographs and non-serials catalogue.....	21
7.2.2. Architecture of the monographs and non-serials catalogue	23
7.3. Serials	26
7.3.1. Content of the serials catalogue.....	26
7.3.2. Architecture of the serials catalogue	26
7.4. Clumps research and development.....	28
7.5. Services.....	29
7.6. Criticisms of a National Union Catalogue	30
7.7. Cost and timescale.....	31
7.8. Towards a <i>National</i> Union Catalogue.....	33
7.8.1. Expanded coverage of HE libraries.....	33
7.8.2. Public libraries, special libraries and small libraries	34
7.8.3. Outside the National Union Catalogue.....	35
8. Conclusions	36
Annex A: membership of Feasibility Study Steering Group	38
Annex B: List of completed reports in the Feasibility Study.....	39
Annex C: List of people consulted	40
Annex D: Libraries expressing interest in participating in a UK National Union Catalogue.....	43
Annex E: Bibliography	44

EXECUTIVE SUMMARY

The aim of the Feasibility Study for a National Union Catalogue was to undertake a review of the key issues that impinge on the creation of such a resource for monographs and serials, with the primary focus being UK HE and national library catalogues, and make recommendations for a way forward.

Guidelines and requirements

The approach taken involved a wide consultation exercise to determine, as closely as possible, how a National Union Catalogue might be used and what functions it could perform and, from a review of other systems world-wide, which technologies were appropriate in a modern setting. As the work progressed, so a number of guidelines were developed which helped in defining the scope and perception of the resource. These suggested that the National Union Catalogue should:

- be based on cross-sectoral resources but only where pre-existing machine-readable OPAC records exist;
- primarily be a vehicle which supports research;
- be free at the point of use;
- not require authentication for resource discovery or known-item searching;
- require authentication for value-added services such as inter-library loans;
- be hospitable to technological developments;
- be a coherent, managed, extensible, robust resource sized to cope with demand;
- return reliable search results;
- provide quick responses to all users, and
- include bibliographic records of appropriate quality.

In short, if the catalogue is going to bear the name 'National', or aspire to this over a period of time, it must, from the start, be worthy of that name.

Associated with the guidelines, a conceptual model of the National Union Catalogue was developed which identified five generic ways in which the resource might be used by either academics and researchers or by librarians: resource discovery; known item search; bibliographic record supply; re-directed search (e.g. from another electronic resource); and updating. In the conceptual model, each of these requirements was analysed and a list created of search parameters through which the user requirements were defined and used to inform the testing programme. This list of fifteen parameters included well-used search points such as author, title, and subject, took in requirements associated with new additions to the catalogue – updating needs – and incorporated access via geographical location and format. Key parameters of different union catalogue architectures were also enumerated – for example, consistency of search results; handling of large result sets; de-duplication; sorting; performance – and incorporated into the testing process.

Consultation

In parallel with this work, the Study Team developed the survey, discussion and consultation processes to further round out the user requirements and key features of a National Union Catalogue. Two large-scale postal surveys were undertaken, one of academic staff, postgraduates and researchers, the other of librarians. For both, expert input was obtained to maximise reliability of the results: on the sampling framework from STATLAB at the University of Edinburgh and on the planning and design of the questionnaires from The Survey Team, also at

the University of Edinburgh. Questions were piloted with academics and librarians at various stages in the survey planning process.

In the survey of academic staff, researchers and postgraduate research students, a total of 846 questionnaires were returned, representing a response rate of 48%, very acceptable in these circumstances. In outlining the facilities they would like to see in a National Union Catalogue, these respondents said that it should comprise at least the British Library, libraries of the major research universities and libraries of 'traditional universities', be comprehensive in its coverage, have a user-friendly interface, and provide a good range of search access points. The majority of respondents said that they would prefer to obtain any materials that are identified through a search by inter-library loan, there being a marked reluctance to consider travelling outside the 'home' city. The option of buying items did not find favour. At the end of the questionnaire, a minimum of 81% of all respondents indicated that a National Union Catalogue would have a positive effect on their information searching: they saw the potential of the resource as a new large-scale database for undertaking existing information searches more efficiently and providing an opportunity for added-value services.

The survey of library staff was cross-sectoral, taking in higher education institutions, public libraries, the national libraries, Research Councils, and further education colleges. An overall response rate of 43% was achieved, from a total of 1,700 questionnaires dispatched. The potential make-up of a National Union Catalogue requested by librarians bore similarities to that requested by users: it should include the British Library, libraries of the major research universities, libraries of 'traditional universities' and public libraries, have comprehensive coverage, provide a good range of search access points and, in addition, have a facility for downloading high-quality bibliographic records. 98% of staff said that a National Union Catalogue would improve services for users, and would improve the efficiency of their own work. Provided what is seen as a practical way forward could be found, with some central funding, there was strong support from the library community.

Wider consultation included meetings with staff from 35 'key players' in the UK, continental Europe and North America: any commercial, academic or other organisation who owned or marketed a bibliographical database within the ambit of the Study, had experience relevant to the creation of large-scale physical catalogues, serials catalogues or had experimented with clumps, had library collections that could contribute to a national resource, and including professional organisations and other key-interest groups. In addition to providing invaluable insights and experiences and enthusiasms, this group of approximately 70 persons also participated in a Concertation Day in February 2001 at which the early recommendations of the Study Team were presented. This consultation process proved a demanding experience but one through which it was possible to view the recommendations in renewed light.

System tests and comparisons

Extensive system testing took in seven vendor systems commonly used in UK HE, physical union catalogues and the four eLib clumps projects. In any distributed framework for a National Union Catalogue, interoperability between vendor systems is especially important and the performance of their Z39.50 servers gives an indication of the viability for connectivity between individual systems and the work involved in creating clumps from these important building blocks. The tests indicated a tremendous variability in the implementations of the Z39.50 protocol, to the extent that one might begin to question whether this is an international standard at all. As expected, a wide variation of attribute combinations was found to be necessary to undertake reliable searching.

Moving from vendor systems to a comparison of physical and virtual catalogues, it was evident in all cases that the physical catalogue architecture offered a more reliable, faster and consistent response than any of the virtual systems tested. Comparison of identical searches confirmed the supremacy of the physical model at the current time, particularly in relation to the user requirements identified in both the conceptual model and the questionnaire survey: for all possible search points, the physical catalogue showed superior consistency and performance every time. The reliability achieved from a search is paramount to the success of any database system and it is believed that any large-scale implementation of virtual systems in their current state of development would undermine the high quality towards which a National Union Catalogue must aspire.

Comparative cost figures indicated that, contrary to expectations, the building of clumps is not significantly cheaper than building physical union catalogues. In fact, the costs of creating both are remarkably similar. The annual operating costs of both systems are also similar.

Supporters of distributed systems have pointed out that the clumps are projects and have not had time to mature into fully-developed services, and to the way in which upcoming initiatives such as the Bath Profile will improve search results. These are genuine comments which have received a sympathetic hearing, but the recommendations for a National Union Catalogue model have to be based on replicable, accurate, robust and responsive systems that are available with current technology, not based on promises for the future which might not materialise.

On this basis, the physical catalogue model offers a clear foundation. But as technology moves on, the National Union Catalogue must be able to adapt: it must not be seen as wedded to a vision of the past, to outmoded concepts, but be hospitable from the outset to new ways of doing, as long as these new ways can deliver the quality resources required. Accordingly, opportunities promised by virtual catalogues should not be dismissed but be investigated further and, if they come to fruition would point to a future hybrid architecture, an integrated resource of physical and virtual elements.

Testing of serials union catalogues revealed not so much differences between physical and virtual architectures, as the dubious quality of much of the underlying data from local library catalogues, including both the bibliographic record, and the volume/year holdings information. As a result of this, it is recommended that resources be assigned to improve data quality, through utilisation of the CONSER database of serials records to facilitate an intensive programme of upgrading at the national and local levels, including provision of standard summary holdings information. This programme will initially lead to a physical National Union Serials Catalogue, but as with the broader National Union Catalogue this must be hospitable to links with other existing catalogues, with the possibility of forming an integrated resource in the future.

A model for the UK National Union Catalogue

Aligning the outcomes of the system tests and cost comparisons with the National Union Catalogue guidelines of operation leads to the recommendation that a physical architecture be pursued in the first instance to create a foundation for a National Union Catalogue over an initial period of three years. The foundation will consist of the integration of COPAC, the British Library catalogues (including existing union catalogues such as the English Short Title Catalogue), the catalogues of twenty cross-sectoral libraries with collections of identified importance to research, and a National Union Serials Catalogue. It will offer the facility of bibliographic record supply, subject to agreements being reached with all relevant parties. During the initial period of three years, further research and development will be undertaken and, based on the results, decisions taken on how to build upon the foundation. Links to clumps will be

incorporated into the foundation architecture from the outset, as will the ability to access this from outside via Z39.50.

At this stage, it is recommended that the architecture utilises three linked systems, separate databases but with identical data schema/structures and internal operations, each holding the full data of a distinct and non-overlapping group of libraries. Such a process would remove limitations which could arise in a direct expansion of the COPAC model and has the additional advantage that it enables the continued operation of COPAC separate from the National Union Catalogue if that was considered desirable. It is further recommended that improved search and navigation facilities be provided for resource discovery, particularly in the areas of subject search, geographical/regional search and format search.

The real challenge for the National Union Catalogue is in agreeing a methodology for adding other libraries to the database in parallel with CURL and the British Library during the important foundation period. These issues have been discussed widely at Steering Group meetings and they were also a focus of the Concertation Day, at which the Study was perceived by some as being limited to HE but with a spurious 'National' badge. The recommendation from the Study Team is that the twenty new libraries should be representative of cross-sectoral interests and be chosen from the regional public library centres, the Research Councils, government and other special libraries, the other two national libraries and a range of HE institutions. The emphasis on research would still be maintained but such an approach would transmit the signal that the resource being planned was indeed National. Comprehensiveness – within a sector or subject area – will not be achieved in the foundation catalogue but the vision for the future must be.

The National Union Serials Catalogue would comprise records and holdings from the British Library, the National Libraries of Scotland and Wales, some large research university libraries, and some smaller more specialist institutions. This will be a physical catalogue of high quality bibliographic records, making use of customised software (the 'Serials Upgrade Application') to import serials records from local catalogues, match with existing CONSER records, upgrade where appropriate, load into the National Union Serials Catalogue with standardised holdings information, and copy back upgraded records to local catalogues. As records are upgraded it is possible that it will eventually become unnecessary to maintain a separate National Union Serials Catalogue – records will be available in the other components of the National Union Catalogue – but it is equally possible that added-value services may be developed which will enhance and maintain the value of a separate serials catalogue.

The National Union Catalogue for the UK, as recommended, would consist of 23 million bibliographic records from the CURL database, 9 million records from the British Library and an perhaps an additional 15 million items from the other cross-sectoral contributing libraries. The National Union Serials Catalogue would include around 750,000 locations from the British Library and the National Libraries of Scotland and Wales, and up to 450,000 other serials record locations. It has not been possible to consider in any detail the services that might flow from such a resource. However, there would appear to be real possibilities of linking a National Union Catalogue to OpenURL initiatives and to the DNER Join-Up developments. Services associated with interloans will be difficult to implement until a national authentication scheme is in use.

The creation of the foundation National Union Catalogue will take three years and cost around £1.5 million; thereafter, operating costs would be approximately £114,000 per annum. Both these costs are in addition to the current financial support for COPAC.

Should in principle funding be agreed, a full specification and operational requirement would need to be made and an open procurement exercise undertaken.

1. THE STUDY AND ITS ORGANISATION

The Feasibility Study for a National Union Catalogue for the UK (UKNUC) has been funded jointly by the Joint Information Systems Committee of the Higher Education Funding Councils (JISC), the Research Support Libraries Programme (RSLP) and the British Library's Co-operation and Partnership Programme.

The contract was granted in May 2000 to a consortium led by the University of Sheffield Library and including the University of Glasgow Library, Crossnet Systems and CURL (the Consortium of University Research Libraries). The consortium was chosen because of its wide range of experience with union catalogues and related issues. Peter Stubley, Project Manager, is Assistant Director of Library Services at the University of Sheffield and was Project Director for the RIDING virtual union catalogue 'clump' project funded by eLib (JISC). Tony Kidd, Head of Serials and Document Delivery at Glasgow University Library has extensive serials expertise. Rob Bull of Crossnet Systems has been involved in the review and development of many Z39.50 systems in the UK and continental Europe at the national and regional level. CURL was formed in 1982 to facilitate the development of common interests between the major research libraries in the UK and has developed, in association with JISC, the COPAC physical union catalogue. Dr Clive Field, Librarian and Director of Information Services at the University of Birmingham is Chair of CURL and the HE libraries representative on the Full Disclosure Implementation Group. Michael Hannon, Director of Library Services and University Librarian at the University of Sheffield is Treasurer of CURL and Project Director for the Feasibility Study. Thus, four out of the five members of the Study Team are directly responsible for day-to-day service provision and have first-hand understanding of the requirements and demands of library users. A Steering Group, chaired by Derek Law, University Librarian and Director of Information Strategy at the University of Strathclyde, and including individuals with a wide range of experience and interest in union catalogues, has met at approximately two-monthly intervals throughout the Study. Its membership is shown in Annex A.

The aim of the Feasibility Study was to undertake a review of key issues that impinge on the creation of a National Union Catalogue for monographs and serials for the UK, with the primary focus being UK HE and national library catalogues and collections. The work began with a literature and systems review, considered the requirements of users, librarians and key players, and undertook a technical analysis of systems and an overview of costs and potential service spin-offs. This accumulated information has been integrated into the recommended model presented in section 7 of this report. A list of completed reports of the various stages of the Study is presented in Annex B and an extensive bibliography is included at Annex E.

2. GUIDELINES FOR DEVELOPING A UK NATIONAL UNION CATALOGUE

Over a period of time, and through discussions in the Study Team, among the Steering Group and with some of the key players mentioned in section 3.1, a number of guidelines were developed which have helped to progress the thinking behind the National Union Catalogue. These cover:

Content:

- the UKNUC will be based on cross-sectoral resources but only where pre-existing machine-readable OPAC records exist.

The Study Team recognises that significant resources exist outside HE and the national libraries, particularly in the public library sector. However, the records of many of these collections, especially in public libraries, have not yet been retroconverted and there is an important Full Disclosure-type role in converting more manual records, regardless of sector.

Intended audience:

- the UKNUC should primarily be a vehicle which supports research.

In making this statement, this report does not recommend the creation of an elitist resource. Rather, it recognises that the National Union Catalogue will be built over a period of time and that an alignment at the early stages with the aims of the RSLP, the Distributed National Collection and Full Disclosure will provide guidance on its foundations. The focus on research also helped to determine the target audience for the questionnaire survey of potential users. But hospitality is also provided to other initiatives, especially at regional or sub-regional levels, which have a different or less exclusively research focus, and to library collections irrespective of their sectoral location.

Access:

- the UKNUC should be free at the point of use,
- not require authentication for resource discovery or known-item searching, but
- require authentication for value-added services such as inter-library loans.

The Study Team firmly believe that the catalogue should be free at the point of use and no-one in HE (or in the national libraries) has challenged this principle. It would also be recommended that JISC's lead in funding COPAC be followed so that individual universities do not have to subscribe to make available what should be a resource for the whole of HE (in spite of the use made by persons outside HE and from abroad). The lack of a requirement for authentication follows from this, though if value-added services are to be offered from a UKNUC, the need for a (future) national authentication scheme, probably associated with metering and secure electronic payment systems, cannot be side-stepped.

Architecturally:

- the UKNUC should be hospitable to technological developments.

With an apparently ever-increasing range of technical options becoming available for searching on the internet, for linking and for document delivery, and with new, developing, standards, it is important that future possibilities are not closed-off in whatever design is adopted.

Quality:

- the UKNUC should be a coherent, managed, extensible, robust resource sized to cope with demand,
- return reliable search results,
- provide quick responses to all users, and
- include bibliographic records of appropriate quality.

The [Chambers] dictionary definition of coherent, as 'consistent; easily followed, connected', could almost act as a definition of the catalogue itself, though good management (e.g. the delivery of the resource/service in accordance with a service level agreement) and a design that permits planned growth are equally important. Robustness and reliability are vital, not only in terms of hardware and software – extensive demand or using the system in an 'unplanned' way must not make it crash – but also in its response to searches. The reliability achieved from a search is paramount to the success of any database system and users will only keep returning if they have confidence in the results obtained.

In short, if the catalogue is going to bear the name 'National', or aspire to this over a period of time, it must, from the start, be worthy of that name.

3. DISCUSSION, SURVEY AND CONSULTATION

3.1. Discussions with key players

Communication has always been seen as the key to the success of the Feasibility Study. A development of this magnitude impinges on myriad aspects of information work while at the same time itself being shaped by the needs of its user community of academics and librarians, and the organisational and economic necessities of interested institutions. The views of users were gathered through two questionnaire surveys (section 3.2) but the detailed and considered responses to the proposal of a UKNUC from 'key players' were only ever going to be gained by personal visit, discussion and the exploration of ideas. 'Key player' was deliberately used in this context as replacement for the more common 'stakeholder', for to imply that certain individuals or organisations held stakes during the progress of the investigation might have undermined the independence with which the Study wished to be perceived. In the eyes of the Study Team, key players were any commercial, academic or other organisation who owned or marketed a bibliographical database within the ambit of the Study, had experience relevant to the creation of large-scale physical catalogues, serials catalogues or had experimented with clumps, or had library collections that could contribute to a national resource. Views were also valued from professional organisations and key-interest groups who, while not having data to contribute themselves, could offer an informed perspective on National Union Catalogue processes.

In October 2000 Peter Stublely and Tony Kidd visited North America to gain further insight into the organisation, operation and limitation of union catalogue systems. In addition, the opportunity was taken to discuss the current thinking on standards issues, in particular Z39.50 and the Bath Profile. The organisations visited were the National Library of Canada, the Committee on Institutional Cooperation (CIC), OCLC and OhioLINK.

The outcomes of some of these meetings are included in the following sections in this report and a list of people consulted is included at Annex C.

3.2. Questionnaire surveys

3.2.1. Survey of academic staff, researchers and postgraduates

The questionnaire survey of potential users of a UKNUC attempted to do three things: provide cross-sectional information about academics, postgraduates and researchers and their current use of library catalogues; gather data to inform the design parameters of a UKNUC; and gain reactions on the idea of a UKNUC and its potential usefulness.

The users targeted were chosen from academic staff, researchers and postgraduate research students as a committed and interested group with significant experience of library catalogues and library services. The sampling framework for the survey was worked through with Mr John Duffy of STATLAB at the Department of Mathematics and Statistics, University of Edinburgh and the questionnaire was designed in close co-operation with The Survey Team at the University of Edinburgh. A total of 846 questionnaires were returned, representing a response rate of 48%, recognised as a very significant percentage response for a postal survey.

The responses to the set of questions on existing catalogue use indicated that the design target audience had in fact been hit. With a minimum of 90% of respondents making use of the catalogues of their home institutions, around 85% valuing library services as highly as discussion with colleagues and use of the Web, and a similar percentage utilising journals databases and online full-text services, these are experienced information seekers and library users who have an understanding of how to use catalogues and what they can get out of them. Even so, existing

experience of union catalogues was low, apart from the 46% of arts and humanities academics in the research-led universities who used these.

If the UKNUC were built to satisfy the requirements of respondents, it would:

- comprise at least the British Library, libraries of the major research universities and libraries of 'traditional universities'
- be comprehensive in its coverage and have a user-friendly interface
- provide access via author; title; journal title; subject; and keywords
- enable the location of all journals in a particular subject
- permit the location of all books in a particular subject or by a particular author
- enable the pre-selection of libraries with strong collections in subject areas and in specific geographical areas and the subsequent searching of these.

A clear steer was also received for integrating search facilities with journals databases and, while primarily coming from the STM area, this did receive some support from social scientists. The strength of feeling suggests that this feature should be considered in future design, notwithstanding the fact that differences in specificity between the terminology used in books and journals may make its successful implementation difficult.

The majority of respondents would prefer to obtain any materials that are identified by inter-library loan, there being a marked reluctance to consider travelling outside the 'home' city. The option of purchasing items did not find favour, though arts and humanities and social sciences academics expressed more interest in this than scientists. Interestingly, a number of written comments from the STM area suggested that this was not the type of resource that would meet their needs but, in the final analysis, 76% of this group in the research-led universities were positive about the creation of a National Union Catalogue.

One might expect that, in the face of a new, unproven resource, users would be understandably reticent about making a commitment to a concept. In fact, the reverse was true: a minimum of 81% of all respondents indicated that a National Union Catalogue would have a positive effect on their information searching. This impressive figure was obtained at the end of a detailed questionnaire at a point when the issues surrounding the new resource would have been obvious to all. Clearly, academics, postgraduates and researchers across UK HE see its potential as a new large-scale database for undertaking existing information searches more efficiently and providing an opportunity for added-value services.

3.2.2. Survey of librarians

A similar, though different, survey was undertaken of library staff cross-sectorally: in HE, public libraries, the national libraries, Research Councils, and further education colleges. An overall response rate of 43% was achieved, from a total of 1,700 questionnaires dispatched, again a very respectable response from which it is possible to draw meaningful conclusions.

In their existing use of union catalogues, librarians pointed to the British Library OPAC as being by far the most popular, 70% of all respondents consulting it frequently or occasionally. HE librarians make as much use of the CURL database/COPAC, while public library staff rely on Unity and V3.Online as well as the BL. This current use translated into a potential requirements list for the National Union Catalogue so that, if the UKNUC were built to satisfy the requirements of librarians, it would:

- include the British Library and the libraries of the major research universities, and possibly catalogues of 'traditional universities' and public libraries

- be comprehensive in its coverage and have a user-friendly interface
- include serials as well as books, with grey literature also a strong candidate; there were mixed views on the practicality of including electronic resources
- have a facility for the downloading of high-quality bibliographic records
- provide access by title, author, keywords, journal title, ISBN and subject.

To sum up, 98% of staff said that a National Union Catalogue would a) improve services for users, and b) would improve the efficiency of their own work. Provided what is seen as a practical way forward can be found, with some central funding, the Catalogue was given high priority, and there was strong support from the library community (see also section 7.6).

3.3. Concertation Day

A Concertation Day, planned in the original Feasibility Study Proposal, was held on 28 February 2001 to allow key players an opportunity to comment on the provisional findings. Chaired by members of the Steering Group, the event attracted around 70 participants, representing most of the organisations and companies that had been consulted over the preceding months.

Following presentations by the Study Team, a range of issues was raised for clarification, comment and criticism. Some participants questioned the need for and level of support for a National Union Catalogue in spite of the results of the surveys of users and library staff. Others suggested that there should be more emphasis on the 'national' rather than just 'HE' nature of such an undertaking, including public and other library collections and accommodate regional agenda. The quality of serials records, serials cataloguing standards, serials holdings information, and links to journal article databases were all discussed.

Technical discussions showed varying support for physical architectures, distributed systems, and alternative approaches, with concern expressed about the extent of de-duplication possible under any system. There was also discussion of the particular approach taken to the physical model, both as a concept, and in terms of content. Some suggested that bringing together BL catalogues and COPAC was a useful, and perhaps sufficient, end in itself, while others were concerned that existing union catalogues such as ESTC should be included and expanded. Present work being carried out on collection strengths and descriptions was felt to be a useful enhancement.

Consultation with the UK library community proved a demanding experience and one from which it was possible to consider the proposed models in renewed light. Some responses to the critical comments that have been received about the National Union Catalogue – not just at the Concertation Day – are included in section 7.6.

4. EXPERIENCE WITH UNION CATALOGUES

4.1. International union catalogues

In the discussions with key players, three major international organisations were contacted to gain a broader picture: Research Libraries Group (RLG); OCLC Europe; and Pica. All operate large union catalogues: RLG with over 40 million unique MARC records and 30 million non-MARC records; OCLC with over 44 million unique bibliographic records; and Pica with 12 million records in the Netherlands and 15 million records in their sister German system. All were initiated to support record supply and shared cataloguing rather than for resource discovery, and the system designs and the economics support this particular perspective. However, all have now expanded services to include resource sharing, using their physical union catalogues as starting points. OCLC now feels that a single physical union catalogue is a barrier to global co-operation

and instead are considering the development of a series of regional hubs around the world: a European Union Catalogue, based on Pica resources is under consideration and is being perceived primarily as a resource sharing rather than a catalogue creation tool.

Pica differentiate between ‘shared cataloguing system’ and ‘union catalogue’, the former having a specific, and clear, purpose, while once a union catalogue had been created, added-value services can be readily incorporated: the Pica CBS – the Central Library (or Bibliographic) System – is offered as a ‘generic solution for the creation and maintenance of union catalogues and for controlled document ordering and delivery’. Through use of the same software – though not the same library management system – at all member libraries, connections to the well-developed, central system permit improved search functionality, payment mechanisms, direct user services and integration with journals databases and full-text. This architecture is also used in the Norwegian academic community. Because of the multiplicity of vendor systems within the UK and the independence exerted by individual institutions, it is not felt that this model, impressive though it is, could easily be adapted to the UK.

Another large-scale catalogue – though not a union catalogue – that has caused wide-spread interest is that of the internet bookseller, Amazon. Although the speed of retrieval and the flexibility of search are impressive, the catalogue, being oriented towards a commercial goal, has particular shortcomings for wider utilisation: out-of-print material appears only sporadically; there appear to be very few explicit edition statements; monographs from organisations and societies are not included; and, for an international company, the database is curiously bounded by country of publication and/or publisher: a book at Amazon.com or Amazon.de may not appear at Amazon.co.uk.

4.2. North American experience

During discussions with the National Library of Canada (NLC), the operation of AMICUS as a physical union catalogue was explained and its importance to and acceptance by the information community in Canada stressed. This acceptance was particularly contrasted with the more developmental Virtual Canadian Union Catalogue (vCuc), especially in the key areas of resource sharing and ILL, and it was clear that vCuc does not have widespread support – apart from a small enthusiastic following – even within the NLC. The support for, and success of, AMICUS are no doubt related to the costs being borne fully by the NLC. The original blueprint for organisation of vCuc, initiated as a technical, feasibility exercise in 1996, was based on provincial clumps but the small number of Z-targets available at the time meant that this co-ordinated approach was impractical. However, the distributed vCuc has shown the way forward in Canada and the model is now being investigated at the provincial level. In spite of this, future funding for addressing drawbacks and expanding vCuc is uncertain and there was a feeling that it could stand undeveloped, remaining simply as a NLC demonstrator.

The Committee on Institutional Cooperation (CIC), incorporating the 14 major research universities in different states throughout the Midwest, aim to provide through the Virtual Electronic Library – the VEL – ‘seamless access’ to a total of 55 million items. Based on CIC philosophy, this has been achieved in a non-centralised manner: no central gateway to the VEL has been created and instead each library develops its Z-client and Z-target according to the institutional priorities and resources available. Software clients thus return different result sets on the same search at different institutions. The development of strong intra-consortial lending is being investigated and a further aim is the provision of access to full text via the VEL. In spite of the decentralised development path, the importance of strong central co-ordination of union catalogue activities was emphasised and it was clear that, at CIC, the commitment is certainly present, and the central leadership is very willing and able, but the lack of core resources does

seem to be a problem. It would appear that the creation of the VEL has been largely a technical exercise. The catalogue has not placed major demands on any library system because, while moderately successful technically, local publicity and uptake have been limited. With no overall co-operative borrowing policy in place and the need for a more developed approach to interlending across the CIC, the espoused seamless access will continue to be a goal for the future rather than a present-day reality.

From these discussions it is clear that the development of distributed union catalogues has not progressed significantly further than work in the UK. All three communities – Canadian, US, UK – are well-informed and aware of the main issues facing Z-based systems, and staff at the National Library of Canada were particularly enthusiastic about the support provided by JISC, eLib and MODELS in investigating resource discovery and related issues. However, whereas both vCuc and the VEL had been considered as prime examples of distributed catalogues before the trip, it is clear that there are limitations in both implementations.

4.3. European experience

Awareness of the architectures used in continental Europe – in Norway, Italy, Germany, Sweden, Finland and Denmark – has been achieved through the involvement of Crossnet Systems in international projects such as ONE and ONE-2. In nearly every case, countries use long-established physical union catalogue systems because they:

- use established indexing techniques which provide consistent search and recall;
- take advantage of established database technology; only recently have suitable networking infrastructures become available to benefit distributed architectures;
- are not reliant on open standards to the same extent as distributed systems.

Traditionally, countries have not had a need for data interoperability with other countries and have developed their own flavours of MARC standards to suit national requirements. Only with the initiation of distributed systems and interoperability projects have the differences been exposed, resulting in the development of complex conversion systems. There is a tendency for countries to support only their own national MARC format and MARC21.

Research and development has seen the addition to national catalogues of extended operational services such as interlending and document supply, as indicated earlier for the international systems. This has allowed national catalogues to be seamlessly integrated with regional services in a country, thereby increasing their profile within user communities. Examples include national library systems and services in Germany and France.

Experimental work in distributed systems has attracted huge interest from national communities and might offer advantages in the future. However, to date, no country (as far as we are aware) has replaced its national physical catalogue fully by a distributed system. National systems are tending to take advantage of modern networking infrastructures by offering online access to other services of interest, e.g. providing links to OCLC. Here, the use of open standards for search and retrieval are working well, although there are many peripheral issues to resolve and implement such as the use of different national character sets.

4.4. National union catalogues and the UK

National, or other large-scale, union catalogues have invariably been established to satisfy a particular need: the availability and supply of good quality bibliographic records to the libraries of a country, a group or an organisation. This need for records – associated with their limited

availability in the early days of library automation – and the ability to link it to a satisfactory charging formula, forged a strong membership base which facilitated the introduction of business models and led to the gradual building up of a portfolio of related services.

In the UK, the library culture that emerged with automation, while similar to this in many respects, developed in different directions. The need for records, particularly in academic libraries, was largely satisfied by the library co-operatives such as BLCMP (now Talis Information) and, perhaps because there were a number of these emergent and moderately successful enterprises covering a relatively small country (in comparison to the US), there was never any suggestion of a *national* catalogue being needed. Until the substantial work surrounding the UK Library Database System (UKLDS) in the early 1980s where, finally, ‘a conflict emerged between the strong wish of the co-operatives to work with the British Library to set up UKLDS ... and their constitutional objective to serve first and foremost the interests of their members...’ (*Vine* 53, April 1984, pp. 45–46). There was also a lack willingness to make central funding available to move the initiative forward.

The creation of the National Lending Library for Science and Technology in the mid-1960s, its growth to become the British Library Document Supply Centre (BLDSC) and its continued success as an integral part of all library services in the UK has also impacted strongly on considerations of a UK National Union Catalogue. Most interloans requests were (and probably still are) dispatched to BLDSC with the expectation that in the vast majority of cases these will be satisfied, and satisfied quickly.

Over the last few years, the library interlending and bibliographic record environment in the UK has changed substantially as a result of fragmentation: an increase in suppliers of bibliographic records and less reliance on a single supplier of these; the introduction of electronic materials from a variety of sources including publishers; an increase in commercial document suppliers; a greater willingness of some groups of (university) libraries to consider wider interlending; and a revised pricing policy from the British Library reducing the cost of ‘non-returnables’ (serial articles) at the expense of book interloans. Over the same period, JISC have funded the development of the CURL database into the publicly-available COPAC and the eLib Programme supported four ‘clumps’ projects.

These changes have occurred at a time of opening-up of resources, the development of particular services directed at researchers, and an increased user interest in taking information searching into their own hands, whether on the internet or in library catalogues. While it is still possible to source bibliographic records from a union catalogue (depending on the model chosen) and, ultimately, to base services around it, the utility of a UK National Union Catalogue should perhaps be judged more on ‘usefulness’ to users.

5. SERIALS ISSUES

5.1. Serials data and quality

A prime motivation in establishing the Feasibility Study has been the question of access to serials information. Focus groups of academics organised by RSLP confirmed their keen interest in locating serials titles and holdings, while the results of the survey, outlined earlier, reaffirm the importance of serials for both the researcher and for the librarian.

The nature of the serials ‘package’ is of particular relevance. A serial title is typically made up of volumes, divided into issues. These constructs are important to users, with the serial title in particular providing both relevant subject information and an indication of the likely quality of

the research contained within it, but it is the individual articles making up these issues and volumes that are of vital concern to the researcher. One option would be to create a union catalogue of serial articles, but to do this in any traditional way would be both prohibitively expensive and unnecessary, given the existence of large bibliographic databases covering most – though not all – subject areas. What is required are methods of seamlessly searching a collection of databases, and then relating the results obtained both to local collections, whether print or electronic, and to remote collections, again both print and electronic.

Notwithstanding this emphasis on the article, access at title level is also important, especially for older material, and especially for arts and humanities researchers. It is in this area perhaps that resource discovery is particularly aided by the provision of subject access, whether at the individual title level (subject headings), or at the collection level (databases of collection descriptions/collection strengths).

A particular problem when considering the bringing together of serials records into a union catalogue is their variable quality. In many libraries, there has until recently been no tradition of creating or obtaining quality serials records, perhaps because these records were not so easily obtainable as for monographs, but also because a simple finding list for serials holdings was probably sufficient, given reliance on the services of BLDSC. However, as resource sharing demands have grown, so has the basic requirement to improve the standards of serials records, if only at the level of ISSN provision, and this is particularly relevant in the context of union catalogue provision.

As a result of these limitations with current data, ways of improving the quality of serials records are required and there are two main sources of these: CONSER and the ISSN International Register. The CONSER database consists of around one million high-quality serials records created in MARC21 format by the Library of Congress, the National Library of Canada, and about 25 other mainly North American university libraries, with the National Library of Wales having recently joined as an associate member. The British Library is in regular contact with CONSER, and the forthcoming moves towards MARC21 may in due course lead to closer involvement. CONSER records are mounted on OCLC WorldCat, are available directly from the Library of Congress, and are present on the RLG database.

The ISSN International Register is available from the ISSN International Centre in Paris. It is of similar size, just over one million records in UNIMARC format, with similar coverage (85–90% overlap with CONSER according to our tests), although it is likely to have a wider spread of records from around the world, containing records supplied by national ISSN centres. We believe however that the advantages of CONSER – more complete records, including subject headings and corporate body headings (both sometimes lacking from ISSN records) for resource discovery and record matching; and MARC21 as opposed to UNIMARC format – outweigh the slightly wider coverage of the ISSN Register. Acquisition costs for both databases are similar.

CONSER and the ISSN International Centre hold regular consultations, and any moves to bring the two databases closer together would be welcome, if this allowed UK libraries to benefit from an enhanced choice of high-quality records. Negotiations are well advanced, for example, to institute the concept of the International Standard Serials Title, to encourage harmonisation and therefore to be able to share records more widely.

Just as the bibliographic record for serials in many UK libraries has resisted standardisation, this is perhaps even more true for the holdings record i.e. the record of which volumes and issues are held at particular libraries. UKMARC in particular has not been hospitable to a standard format for this information, but the wider acceptance and adoption of MARC21 will lead to

improvements, although not solving the problem of 'legacy data'. We accept that in the medium term it is not possible for UK libraries to provide detailed holdings information for their serials, but recommend that summary holdings, with standard nomenclature, are made available in the 866 field as MARC21 is adopted.

The MARC Format for Holdings Data (MFHD) may be appropriate for some libraries to provide more detailed holdings, and any serials union catalogue should be hospitable to this format, while recognising that summary holdings will predominate for some time to come. The CONSER Publication Patterns Initiative, where the first significant tranche of detailed publication pattern data, from Harvard, has just been loaded into relevant CONSER records in OCLC, may prove a way of facilitating the introduction of more detailed holdings information. Vendor systems are now making available the appropriate hooks in their serials check-in and cataloguing modules.

5.2. Testing of serials

Serials testing revealed serious deficiencies in the systems examined. The deficiencies related to de-duplication and sorting, but also particularly to problems arising from the underlying quality of the records available from local catalogue systems – different versions of the same title were common, ISSNs were present on a small proportion of records, and holdings data varied widely.

This further emphasised the need for improvement in the quality of serials records at the local level and to aid this process, and to provide the basis of a central catalogue based on high quality data, the availability of sources of serials records was investigated, testing both the CONSER database and the ISSN International Register against samples of records from a local library and from the British Library. These tests showed both that full records are available for a reasonable proportion of holdings (at least at the local library level), and that there are significant numbers of titles where no record is present – this is particularly true for the British Library.

The testing also confirmed that there are titles – annuals, directories, etc. – which fall into a grey area, considered by some libraries as 'serials', by others as 'monographs'. It is clear that full consistency will not be achieved, and this is one reason to offer the facility of combining serials and monographs data records for record retrieval, although it is also the case that a National Union Catalogue should permit separate searching, either through a separate database, and/or by permitting appropriate search qualifiers to be applied.

6. TESTS AND COMPARISONS OF SYSTEMS

6.1. The challenges of technological change

A consideration of the wide-ranging technical and systems issues is vital for a full understanding of the ways in which the resource might be established. The technological world is not standing still, new possibilities are emerging at regular intervals and quite often these new developments are accompanied by criticisms of the out-of-dateness of existing techniques, even when these continue to deliver the best results. Such is the case with the possibilities for search and retrieval when tools used with apparent ease elsewhere on the internet – for example, search engines; Napster – might be assumed to be readily and easily transferable to bibliographic endeavours.

The constancy and unalterability of this change process impacted on the consideration of technical issues in this Study in two main ways. First, while attempts were made to take on board as many new technical developments brought to the attention of the Study Team as possible, some – because of their newness – might have slipped through the net. This is almost inevitable in the current circumstances, though it is believed that nothing that *substantially* effects the

outcomes – and which could not be introduced at a later date – has been omitted. Secondly, in all considerations, the guideline of a robust, coherent, extensible resource has been followed and while many interesting developments have come to light, it is not felt appropriate to recommend the building of a national resource on untested foundations. The difference between developmental, project work and the need to run realistic, real-time services has been recognised.

The Z39.50 protocol is a good example of a technology that falls on both sides of the fence described above, depending on the position of its supporters and detractors. On the one hand the protocol is seen as dinosaurian – far too large, self important and complicated to be able to provide a solution for the streamlined modern age – and on the other the light to lead us down the path to scalable, distributed systems and away from large, unwieldy physical union catalogues. Inevitably, both views have an element of truth but the differences of opinion – not to say misunderstandings – held on the protocol have required a substantial consideration of its *pros and cons*, gained through significant testing in a wide range of systems and implementations.

At the same time, Z39.50 and union catalogues do not stand in splendid isolation from the real world and any recommendations must take account of the potential use made of a large-scale national resource. As an aid to this process, a conceptual model was developed which identified five generic ways in which the Catalogue might be used: resource discovery; known item search; bibliographic record supply; re-directed search; and updating. In the conceptual model, each of these requirements was analysed and a list created of search parameters through which the user requirements were defined and used to inform the testing programme. This list of fifteen parameters included well-used search points such as author, title, and subject, took in requirements associated with new additions to the catalogue – updating needs – and incorporated access via geographical location and format. Key parameters of different union catalogue architectures were also enumerated – for example, consistency of search results; handling of large result sets; de-duplication; sorting; performance – and incorporated into the testing process.

Extensive system testing took in seven vendor systems commonly used in UK HE, physical union catalogues and the four eLib clumps projects. In any distributed framework for a National Union Catalogue, interoperability between vendor systems is especially important and the performance of their Z-servers gives an indication of the viability for connectivity between individual systems and the work involved in creating clumps from these important building blocks. Similarly, the performance of the existing clumps against a range of search criteria indicates how they would fit into a national resource and, with tests on the physical catalogue systems, provides comparative data. The special nature of serials was accorded particular attention with its own testing programme, outlined in section 5.2.

6.2. Z39.50 and testing

The Z39.50 tests of vendor systems were undertaken on installed, working, library catalogues, not on special test rigs, and they therefore represent real-life responses. They indicated a tremendous variability in the implementations of the protocol, to the extent that one might begin to question whether this is an international standard at all. Out of the seven systems tested, three still supported only the outdated version 2 of the protocol, five supported the ‘Scan’ service and only three supported ‘Sort’. Scan and Sort could provide additional utility in distributed systems if implemented – Scan permits searching following the selection of terms from an index; Sort provides sorting capabilities from within the protocol rather than leaving this as a subsequent gateway operation – but with only a 40–70% take-up in vendor systems their effectiveness is always going to be compromised. If not all the catalogues in a distributed system support particular services, the user can be presented with a confusing lack of consistency and the gateway designer left with significant software overheads to face.

By contrast, Z39.50 'Search and present', what might be considered the cornerstone of search requirements, was supported by all systems tested. In spite of this, a wide variation of attribute combinations was found to be necessary to undertake reliable searching. As an example, one of the most popular access points into catalogues – author – is interpreted quite differently between systems, being represented by a minimum of three options: author keyword, normalised name, and non-normalised name. The effect of these variations in interpretations of key search parameters has particular impact on search results when a range of systems from different vendors is combined together, to such an extent that the user can never be sure of the accuracy of the results. Although measures can be taken to optimise search performance across a group of targets, the results can still be disappointing.

From discussions with staff closely involved with Z39.50 issues, it is clear that resolution to these variations in search response is expected to be found through the Bath Profile. With the status of an Internationally Registered Profile, it has been developed – and is still being fine-tuned – with the express intention of confronting interoperability issues head-on. It is lauded for these aims. But, without wishing to be pessimistic, Bath is a promise for the future rather than a remedy today. None of the systems tested support the Profile exactly as it currently stands and those vendors who responded to correspondence from the Study Team – at a less than 50% written response rate – claimed that they will support Level 0 by the end of 2001 with, in some cases, Level 1 following one or two years later. Some vendors appear to be employing a 'wait and see' attitude, ensuring that the Profile will be followed widely before committing resources. One supplier wrote, 'I believe it is 'dangerous' to build any model for the UK Union Catalogue on the assumption that all library vendors will implement the Bath Profile to a 100% compliance'. While recognising the real advances made by the Bath Profile and the fact that it has, in the words of Carrol Lunau of the National Library of Canada, 'reduced the wiggle room' for system suppliers, the Study Team are sensitive to this last comment and would recommend the adoption of a similar wait and see attitude towards the actual appearance of Bath in vendor systems.

Mirroring the current problems with Z39.50 search are those associated with holdings and circulation data. In the last two years, the Z39.50 implementors' community has recognised the need to improve holdings information and a Holdings Schema was published in January 2000. While important, this has – understandably – received less publicity than the Bath Profile and the number of implementations worldwide is very small: certainly none of the systems tested showed evidence of supporting this. What the tests did show is tremendous variability, both in the location of holdings information within the MARC records and the detail provided. Circulation data – though coming within the ambit of the Holdings Schema – showed even less standardisation. Practices that might be considered satisfactory at the local or even regional level become difficult to manage when used in a national service and there would appear to be little agreement on the timescale in which these holdings issues will be resolved, though holdings elements do also form a part of the Bath Profile.

6.3. Comparative testing of virtual and physical union catalogues

To compare the results across virtual and physical union catalogues, tests intended to represent the wide range of search approaches that might be taken by different users were carried out on the eLib clumps gateways and COPAC.

When using the search terms of author and title, real discrepancies were apparent in the results returned by the clumps. No test succeeded in retrieving all available items from the databases searched (for example, on one search 3 out of 5 hits were returned from one clump, 4 out of 6 from another). One clump produced some particularly disappointing responses with 0 hits (out of a possible 5) when the author's first name was reduced to an initial (in combination with title)

and in all tests failing to retrieve an item at one particular location unless the precise title was used. In all tests that produced discrepancies on the distributed systems, COPAC performed quickly, consistently and with a level of precision and predictability unequalled by the clumps.

Comparisons were also made from other starting points: subject/keyword; publisher; date of publication; ISSN; and ISBN. Subject and keyword are the starting points for resource discovery but the searches offer particular challenges in any library catalogue, local or union, virtual or physical. They are dependent not only on the indexing policies implemented in vendor systems but also on local procedures, such as the application of subject headings and the classification scheme in use. For these reasons, detailed testing was not possible and it was simply confirmed that both the clumps and COPAC provide subject search facilities. The use of collection descriptions as an aid to resource discovery have not been implemented to any great depth in any current system tested. M25Link enables the choice of broad subject area as a way of reducing the number of targets searched.

Publisher can be extremely useful in focusing on 'known items' when other elements are unclear and in an early communication from an academic about UKNUC, specific publisher access was requested. The clumps do not support publisher and while they offer ISBN, ISSN and date searching, all produced extremely unreliable results. All these entry points work well in COPAC.

No system currently permits the selection of geographical area (e.g. city or region) in association with other entry points. COPAC enables a selection of parameters – author, title, etc. – in combination with a university selected from a pop-up list, and some of the clumps provide a two-stage process whereby libraries in a certain region are chosen *before* the search terms are entered. Other than serials, no format (e.g. film; CD) searching is possible on either architecture.

In terms of performance, there are two measures: response time; and robustness, as indicated by system load. COPAC was significantly quicker in all comparisons of response time, performing on average four times faster than any of the clumps. The system load on a National Union Catalogue – denoted by the number of user searches performed – would have to at least equal that on COPAC, at around 17,000 per day, and BLPC, at around 25,000 per day. Very little data is available on clump use but, as an example, the RIDING web gateway averages 1,000 hits per *month*. Comparisons with the number of searches undertaken on the OPACs of three research-led universities – 17,000–25,000 per day – suggests that a virtual architecture would result in a doubling of system load on the library system of each member in a clump (more so for smaller institutions). Presenting these figures to one system vendor resulted in the remark, 'Our concern would be that if you allow a large amount of UKNUC searches to descend on any single institution you are likely to submerge it'. Utilising the physical architecture allows the large loading to be designed into the system from the beginning.

The importance of robustness was emphasised in the questionnaire survey of academics where sixteen respondents emphasised the importance of reliability, stability, the need for platform independent web access, open protocols and access from on and off campus. For example:

- I think this is a very good idea but it would need to be properly resourced so that it is up to date and the electronic service is very fast. During term time the load on such a service is likely to be very high and vary greatly in load. It would be important to specify system requirements at the high rather than average level.

In summary, the testing of virtual union catalogues showed them to be very much at an emergent stage. The eLib clumps projects have taken individual approaches to the creation of gateways and consequently their look-and-feel and method of presentation and handling of results are quite

different. Their performance can be affected quite drastically by targets being unavailable or unresponsive and in extreme cases the complete system can be brought to a halt. In some gateways, feedback mechanisms have been implemented to indicate the progress of a search to the user. The handling of large result sets varies from the non-existent to fair, the provision of sorting and de-duplication facilities is sporadic, holdings information is extremely variable and circulation details are not displayed. Quality of bibliographic records is variable, the display of MARC fields variable and in many cases quite limited, and the operation of organised and controlled bibliographic record supply from this architecture would not be easy. To act as interactive and integrated components in a UKNUC, virtual union catalogues must themselves provide an entry point for bibliographical queries, a Z39.50 'point of presence'. This introduces further complications in identifying the source of search results and the presenting of these to the user in an understandable manner, and none of the projects have resolved these issues.

The testing of physical union catalogues showed them to be stable systems that responded consistently to a wide range of search requests. Their speed of response is impressive and they can be designed to handle anticipated user load. The effectiveness of de-duplication at the pre-processing stage was tested for COPAC and indicated that, in spite of the sophistication of the algorithms used, significant duplication still exists and needs addressing, though an operation to improve this is currently under way at one CURL member. Equally, the handling of large result sets is poor in COPAC and not particularly helpful for more than limited resource discovery. Quality of bibliographic records is good and the architecture enables controlled bibliographic record supply to be de-coupled from end user searching. The linking from COPAC to local systems for the retrieval of circulation data works very well indeed and shows how dynamically-changing data can be incorporated into a physical union catalogue.

6.4. Comparative costs of virtual and physical union catalogues

Cost data which allows accurate comparison to be made between physical and virtual union catalogues is difficult to obtain, particularly as it relates to future estimates in areas which are themselves continuing to develop. All efforts have been made to ensure that the data used is representative and as accurate as possible.

Figures for the virtual union catalogues have been based on conversations and data provided by the three regional clumps: CAIRNS, M25Link and RIDING. These figures are rather diverse and represent the individual nature in which the clumps have been created, the solutions taken, and their expectations for the future. The role of software developers and the utilisation of commercial, custom-designed products also impact on the costings, together with the organisational issues to be resolved in making the clump cohere into a successful unit. Any new clumps will not be stand-alone and must be software linked in a co-ordinated manner to other elements of a National Union Catalogue and a nominal figure has been included for this.

For the physical model, indicative costs have been based on information provided by Manchester Computing on the building of a catalogue based on CURL/COPAC design but not taken directly from existing figures. The reasons for this are that the make-up of any new catalogue is more likely to be a mix of large and smaller libraries rather than the large research libraries of CURL, and that certain software developments not in the current system will be required to translate this into a component of a National Union Catalogue. The costs assume that the existing infrastructure and expertise remains in place but otherwise includes all additional costs required to establish a new physical union catalogue.

Figures are presented in the table for comparative set-up and annual operating costs under five headings: staffing; hardware; software development; travel & subsistence; and publicity. The first

part of the table presents the set-up costs, though the timescales are a little different. A two-year set-up for a virtual catalogue being built from the bottom up would broadly reflect the experiences of the clumps but, as none delivered their current level of operation in that time, allows for a faster development process and the use of existing expertise. In the physical model, it is estimated that a fully-operational service for a catalogue of twenty institutions would begin halfway into year three although the first libraries would be on-stream within eighteen months. The annual operating costs for all systems are similar apart from RIDING which has taken the decision to dispense with dedicated clump staff and work through a Service Level Agreement for technical support. This figure may have to be reviewed in the light of experience.

Comparative set-up costs

System type	Staffing	Hardware	Software development	T&S	Publicity	Total
COPAC-like system	£167,500	£29,000	£0	£6,000	£5,000	£207,500
CAIRNS-type clump	£155,000	£16,000	£5,000	£6,000	£5,000	£187,000
M25Link-type clump	£190,000	£30,000	£5,000	£6,000	£5,000	£236,000
RIDING-type clump	£120,000	£18,000	£48,000	£6,000	£5,000	£197,000

Comparative annual operating costs

System type	Staffing	Hardware m/tenance	Software support	T&S	Publicity	Total
COPAC-like system	£35,000	£1,500	£0	£1,000	£1,000	£38,500
CAIRNS-type clump	£30,000	£1,000	£0	£1,000	£1,000	£33,000
M25Link-type clump	£35,000	£1,000	£0	£1,000	£1,000	£38,000
RIDING-type clump	£2,000	£1,000	£11,000	£1,000	£1,000	£16,000

While one might be encouraged to derive ‘per library’ set-up costs from these figures, the conclusions reached would not necessarily be comparable due to the large number of variables involved. For a start, the sizes of the clumps are quite different, with CAIRNS having twenty targets (though not all are library catalogues), RIDING ten, and M25Link starting with six in the original project but with plans to increase to forty. Based on these figures, any scaling-up of the way in which the UK could be covered by clumps would lead to quite different results, and results which would not take into account – as did the original projects – the quite different local, regional and organisational factors at work in bringing groups of libraries together for this purpose. A true like-for-like comparison is further complicated by the fact that the final products are different: the physical union model would be able to offer bibliographic record supply along the lines of the CURL/COPAC design (subject to agreement), and its utility and performance would be superior to the virtual.

6.5. Outcomes from system tests and comparative costs

In all cases, it was evident that the physical catalogue architecture offered a more reliable, faster

and consistent response than any of the virtual systems tested. The reliability achieved from a search is paramount to the success of any database system and it is believed that any large-scale implementation of virtual systems in their current state of development would undermine the high quality towards which a National Union Catalogue must aspire.

The comparative cost figures indicate that, contrary to expectations, the building of virtual systems is not significantly cheaper than building physical union catalogues. In fact, the costs of creating both are remarkably similar. The annual operating costs of both systems are also similar and until further experience is obtained, conclusions should not be drawn about the low upkeep costs from one clump in particular.

Supporters of distributed systems point to the fact that the clumps are projects and have not had time to mature into fully-developed services, and to the way in which upcoming initiatives such as the Bath Profile will improve search results. These are genuine comments which have received a sympathetic hearing, but the recommendations for a UKNUC model have to be based on replicable, accurate, robust and responsive systems that are available with current technology, not based on promises for the future which might not materialise.

On this basis, the physical architecture offers a clear foundation for a National Union Catalogue. But as technology moves on, the foundation must be able to adapt: it must not be seen as wedded to a vision of the past, to outmoded concepts, but be hospitable from the outset to new ways of doing, as long as these new ways can deliver the quality resources required. Accordingly, opportunities promised by virtual catalogues should not be dismissed but be investigated further and, if they come to fruition would point to a future hybrid architecture, an integrated resource of physical and virtual elements. The further work recommended is described in section 7.4.

7. A MODEL FOR THE UK NATIONAL UNION CATALOGUE

7.1. Recommended architecture

Aligning the outcomes of the system tests and cost comparisons with the National Union Catalogue guidelines of operation leads to the recommendation that a physical architecture be pursued in the first instance. This section details how that can be achieved to create a foundation for a UK National Union Catalogue over an initial period of three years. The foundation will consist of the integration of COPAC, the British Library catalogues, the catalogues of twenty cross-sectoral libraries with collections of identified importance to research, and a National Union Serials Catalogue. It will offer the facility of bibliographic record supply, subject to agreements being reached with all relevant parties. During the initial period of three years, further research and development will be undertaken on virtual systems and, based on the results, decisions taken on how to build upon the foundation. Links to potential virtual union catalogue elements will be incorporated into the foundation architecture from the outset, as will the ability to access this from outside via Z39.50. Ways of expanding the foundation to true 'National' status will be finalised at this time, though issues for consideration are included in section 7.8.

7.2. Monographs and non-serials formats

7.2.1. Content of the monographs and non-serials catalogue

The responses to the questionnaire survey from academic staff, postgraduates and researchers – and supported by HE librarians – indicated that the two groups of catalogues most preferred in a National Union Catalogue are those of the British Library and major research universities. Clearly, the BL and COPAC will not provide the full answer, but the importance of bringing these two major resources together in this way must not be underestimated, nor must the impact it

would have on the academic community in the UK and the wider international scholarly community. These two major building blocks would then be enhanced by other key collections which could be drawn cross-sectorally from around the UK.

COPAC is the consolidated and de-duplicated catalogue of non-MARC records created from the CURL database for items held in CURL libraries. Twenty UK university research libraries contribute their data and the size of the (un-de-duplicated) CURL database, including British Library and Library of Congress records, is around 23 million records. The size of COPAC is approximately 8 million records.

The British Library component of the National Union Catalogue would basically be the Corporate Bibliographic Database (CBD). This includes two main elements:

- the BL Current Catalogue comprising Humanities and Social Sciences (2 million records), Music (260,000 records), Maps (23,000 records), Science, technology and business (290,000) and the Document Supply Centre (1.1 million monographs; 500,000 serials); and
- the BL Retrospective Catalogue comprising pre-1975 printed books (4 million records), Science, technology and business books pre-75 (100,000 records), and Maps pre-75 (200,000 records).

British Library and Co-operative Bibliographic Resources – the British National Bibliography; the UK ISSN file; the Register of Preservation Microforms; and SIGLE (System for Indexing Grey Literature in Europe) – which are also components of the CBD, would in general not form part of the UKNUC. However, two key elements – the ESTC (English Short Title Catalogue) and the ISTC (Incunabula Short Title Catalogue) – could be incorporated, even though, as union catalogues already, their treatment might be different from other BL files. The ESTC, for example, contains approximately 600,000 records of early English books and serials to 1800, created to international bibliographic standards. It contains data on the locations of several million copies, based on the holdings of hundreds of collections in the UK and overseas. The total British Library component of the National Union Catalogue would be approximately 9 million records, though expected to be reduced substantially through their de-duplication programme.

COPAC and the British Library catalogues are obvious contenders for membership of a UK National Union Catalogue, particularly from the perspective of users. The real challenge comes in agreeing a methodology for further populating the database during the important foundation period. To give out signals that the resource being created is more than simply a combination of ‘obvious parts’, care must be taken to ensure that the libraries and collections bring real additional value at an early stage to academics, librarians and others who will use it. Failure to do so may bring accusations of stifling (or force feeding!) the infant at birth, while a considered approach to the foundation Catalogue will facilitate a flowering to its full potential.

These issues have been discussed widely at Steering Group meetings and they were also a focus of the Concertation Day, at which the Study was perceived by some as being limited to HE but with a spurious ‘National’ badge. Of course, the process is easier to talk about than achieve but, if twenty libraries are to be added in the three year foundation period, then the recommendation from the Study Team is that these should be representative of cross-sectoral interests. Some of the issues associated with the wider involvement of public libraries and special libraries are considered in section 7.8 but it would be entirely possible, and advisable, to include some of these at an early stage, while still maintaining the emphasis on research use. The regional public library centres (e.g. Birmingham, Manchester, Newcastle) are a self-selected group with specialist holdings, and the Research Councils, government and other related libraries have major

untapped resources. Inclusion of these categories with a range of HE institutions, together with the National Libraries of Scotland and Wales would transmit the signal that the resource being planned was indeed National, at the same time as increasing the 'uniqueness' of the content.

In building up the foundation Catalogue in this way, it will not be possible to provide comprehensiveness in any library type or subject area but rather a representative selection. Even so, not all subject areas will be covered and broad choices might result in favouring some subjects and disciplines – e.g. arts and humanities – over others. In the current absence of an objective measure of collection strengths, the HE libraries might be chosen from the non-CURL RSLP Access Survey libraries, a list of 15 institutions including, for example, Bristol, SOAS, Manchester Metropolitan, Cardiff, and Leicester. Changes in CURL membership (and criteria for membership) would also impact on the process. Further soundings need to be taken to reach agreement on the composition of the foundation National Union Catalogue. It is, however, recommended that *complete* catalogues be added, not just special collections which might be considered to be of prime interest. It is estimated that an additional 15 million items might be added in the ways suggested, equivalent to perhaps 6 million unique bibliographic records.

Formats other than monographs, including for example, maps, music, microforms, audio-visual materials and electronic resources (although – by-and-large – not archives and manuscripts), would be included in the foundation Catalogue, largely as they are added from contributing library catalogues; 24% of the CURL database, for example, is non-monograph material. Specialised databases of non-monograph formats have not been investigated in depth in the Study but there is no reason why the more important of these should not be identified and added.

7.2.2. *Architecture of the monographs and non-serials catalogue*

Once it was clear that a physical architecture should play a major role in a National Union Catalogue, extensive discussions were held with Manchester Computing to ensure that any proposed model was practical and based on existing expertise and experience, although no assumptions have been made on the eventual hosting of the system.

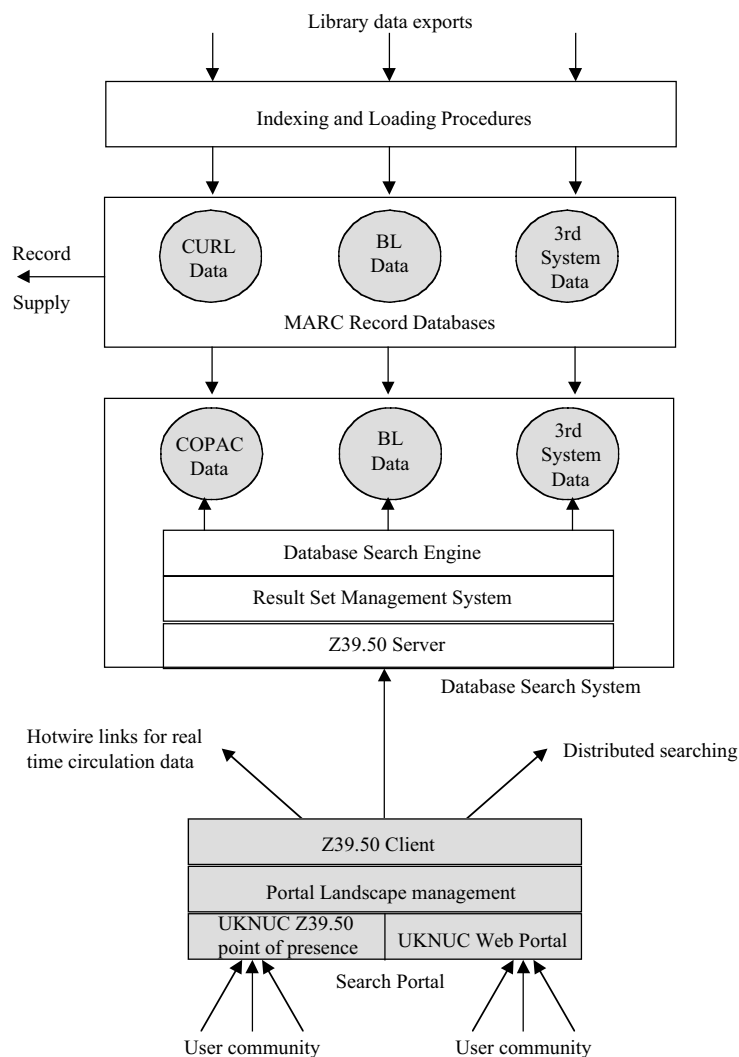
The most obvious starting point is a straightforward expansion of COPAC, exporting data from the British Library and the other libraries and creating a single physical union catalogue/database. This would offer good quality data with a consistency similar to COPAC and require simple networking and interface creation. The downside lies in the updating. At present, a single updating stream is created from all contributing libraries and, though the limit on this process has not yet been reached, bottlenecks could form as the resource expands.

A process whereby the update streams could be handled in parallel rather than in series would lay the foundation for the future by removing these limitations. By creating any number (up to a total of perhaps 10) of identical, linked databases, searching can still possess the accuracy and reliability of a single database, but the arrangement can offer real benefits in system management. Internal networking and interface issues – portal design and operation – become more complex when compared to the single physical system but are, according to Manchester Computing, well within known bounds. For example, de-duplication of records produced from the individual systems must take place 'on-the-fly' but at this stage the software would be handling only a fraction of the records in the complete database: no noticeable effect would be seen by the user.

Another option for combining COPAC and the British Library catalogues would be to leave these in place and link via Z39.50. However, there are real difficulties in doing this, particularly associated with de-duplication in presenting results to the user. COPAC does not itself produce MARC records, whereas the British Library catalogues do and so one would require a mechanism

that de-duplicated different data formats. Even trying to de-duplicate the CURL database against the British Library would be difficult – a typical resource discovery user would be retrieving moderate-sized result sets and these would take real time and substantial processing power to resolve, particularly given the load placed on these systems. This is not a practical alternative.

As the foundation of the National Union Catalogue will be based on three quite different elements, it is recommended that three linked systems be established initially: as described above, separate databases but with identical data schema/structures and internal operations, each holding the full data of a distinct and non-overlapping group of libraries. One system would hold the data of CURL libraries, one system would hold British Library data, and the third would hold the data of the new contributing libraries. This model has the additional advantage that it enables the continued operation of COPAC separate from a UKNUC if that was considered desirable. In other words, the UKNUC can easily be developed independently of the future directions of CURL itself. The schematic diagram indicates the make-up of the complete system:



Data from the individual contributing library catalogues will each require different load programs and will be held in two databases (like CURL/COPAC): MARC format un-de-duplicated and display format de-duplicated. The MARC format database will be available for bibliographic record supply as long as agreements on this can be obtained between the owners of the records, contributing libraries and those libraries wishing to participate in this service. CURL members

contribute to this from their CURL subscriptions and licensing/charging arrangements would need to be established for other contributors to this model.

The UKNUC search infrastructure comprises two main components separated by a local area network using the Z39.50 protocol:

- the database search system, and
- the search portal.

The database search system comprises:

- the database search engine that performs the actual search on the data;
- a result set management system, responsible for data harmonisation and de-duplication of items from the three databases; and
- the Z39.50 server to provide the open standard interface to the data.

The search portal comprises the following components:

- a Z39.50 client to perform the searching of the UKNUC database search system and to also perform the hotwire searches to the member library management systems to ascertain circulation data. This Z39.50 client would also be used to search other possible Z39.50 data sources, e.g. clumps.
- the management of user sessions via the portal landscape management module. This module is responsible for user session management including selection of databases to search on and handling of the data from the Z39.50 client.
- the user interface web portal allowing users with internet browsers to search the UKNUC; and,
- a Z39.50 server providing a Z39.50 point of presence to the UKNUC data.

It should be noted that the diagram shows the portal subsystem separated from the database subsystem. Although this is not a mandatory separation it introduces a future-proofing capability for system expansion, and provides additional merits of security, rigorous adherence to open standards and processing distribution. Thus, the UKNUC data could be accessed via the Z39.50 server (as part of the database search system) or the Z39.50 server in the search portal. However, searching the search system with Z39.50 would not provide access to the circulation data or to other distributed data sources.

Utilisation of the Z39.50 client in the Search Portal ensures that the system is further hospitable to future distributed developments. It is important that the open standards utilised in both the distributed and physical elements are implemented to the same and compatible high quality and building up a specification to facilitate this would be an objective of the proposed clumps R&D.

It is expected that the technology pioneered by Manchester Computing in providing circulation status directly from local systems for items on COPAC could be extended to all libraries.

Most search terms identified in the user requirements are available in COPAC and it would be expected that these would all be provided in the National Union Catalogue. However, a significant increase in functionality for resource discovery would be achieved if a subject search facility was developed in detail and if geographical/regional searching and format searching were provided. As collection description and collection strength systems are developed, these too should be integrated into the user interface.

7.3. Serials

7.3.1. *Content of the serials catalogue*

The importance of the British Library serials collections make them key components of a serials union catalogue. A very small sample test of recent serial additions to the BL suggests an overlap with CONSER records of less than 25%, although it is not possible to estimate how accurate a figure this would be over BL holdings as a whole. The BL is presently engaged on an exercise creating a Serials Title Register (STR), combining serials data from separate files covering BLDSC, Humanities and Social Sciences, SRIS, and Legal Deposit Office holdings. This Register will offer a substantially de-duplicated file of BL holdings, combining upwards of 600,000 holdings records, but will not contain high-quality records. Nevertheless, in order to provide access to BL records, prior to records becoming available through the BL Corporate Bibliographic Database, it may be necessary, depending on the serials union catalogue timescale, to copy the STR records to the serials union catalogue.

The serials holdings of the National Library of Scotland and the National Library of Wales are, in general, catalogued to a high standard and both already use the MARC21 format. It is likely that their 150,000 or so records would be early additions. There are also continuing discussions among the national libraries, and the other copyright libraries, on co-operative cataloguing of serials, which may become important in the union catalogue context.

The criteria for choosing the next layer of participating libraries will vary, though it is hoped that the libraries will include some representation of cross-sectoral specialisms. As indicated for monographs, further soundings need to be taken to reach agreement on the full composition of the Serials National Union Catalogue. An element of self-selection might also apply, by libraries wishing for example to improve the quality of their local serials records. Up to 450,000 record locations could be added from these libraries over a three year period.

Local serials information from libraries wishing to participate may be lacking in two distinct ways. First, holdings information may be inaccurate, even at the level of summary holdings. A significant proportion of (the relatively small number of) supply failures by Lamda libraries, for example, are the result of faulty local catalogue holdings data. While 100% accuracy is not a realistic goal, a sample survey will need to be undertaken comparing holdings with issues on the shelf at each potential library participant. If an appropriate level of accuracy is not achieved, then the local library will have to improve data quality.

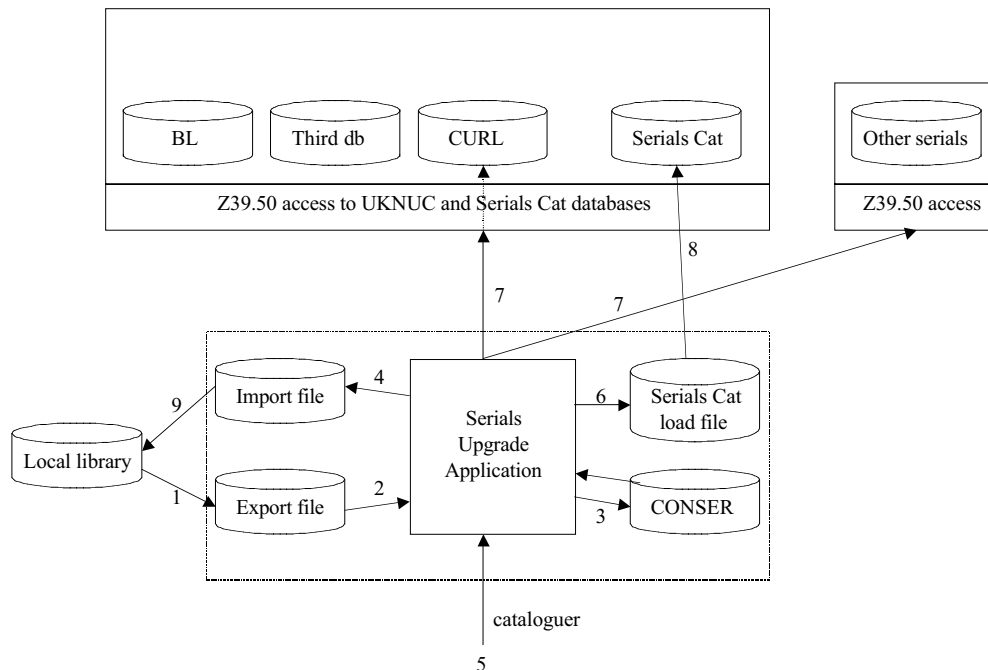
Secondly, the bibliographic records themselves will require considerable upgrading. As an example, tests at Glasgow University Library suggest that approximately one third of records will have ISSNs which can be matched against the CONSER database, a further third are present as CONSER records but not matched – usually due to lack of ISSN on the local record – and the final third represent non-CONSER titles. The procedures detailed in the next section will permit not only the creation of a National Union Serials Catalogue of good quality records but also upgrading of local records to the same standard.

7.3.2. *Architecture of the serials catalogue*

As for monographs and other formats, the National Union Serials Catalogue will be based on the physical model, though in this case it will consist of just a single system. Such an approach will ensure the creation of a database of good quality records from the beginning, using CONSER as a baseline. It will also mirror existing experience in the UK (University of London Union List of Serials, SALSER) and abroad (e.g. Zeitschriftendatenbank [Germany], NOSP [Scandinavia]). While searchable as a single database for high relevance of serials data, it will also be possible for

users to undertake a parallel search alongside the rest of the physical hub. It is conceivable that the need for a separate serials catalogue may decline or disappear altogether, as local catalogue records are upgraded and subsequently added to the physical hub in the manner described below. On the other hand, by the time that has occurred on a sufficient scale, other extended services may have been built on the serials union catalogue rendering its continued existence of value.

The recommended architecture for serials is shown in the diagram below. It allows: the extraction of bibliographic and holdings information from the local catalogue; the matching of this against the CONSER database; uploading to the serials union catalogue; and copying of the upgraded bibliographic record back to the local catalogue. It assumes the building of customised software, the Serials Upgrade Application.



The procedures are broadly as follows:

- The library exports its serials records to an export file (1) in MARC ISO-2709 format.
- The Serials Upgrade Application is loaded with specific rules for matching records in the library export file against CONSER; it performs the batch process reading in the export file records (2) and tries to match against CONSER records (3).
- Any records not identified by the batch process rules will be marked for manual process by a cataloguer. The cataloguer (5) may identify the record in CONSER, or may use other sources (7), including the already existing serials union catalogue. This identification process can use the Z39.50 protocol.
- The cataloguer will complete the records upgrade process, and all records (both those of the batch process and the manual upgrade process) will be placed in two files. One of these files (6) will contain records copied from CONSER (or other source), with the library location details added to field 852, and the library holdings/volume details added to field 866. The other import file (4) will contain a merger of the CONSER (or other source) record with the existing local record details.

- The appropriate file will be uploaded (8) to the serials catalogue file, which will then contain the relevant bibliographic and local holdings information.
- The local library will import the serials records from the import file (9) to improve the quality of its own local records, while protecting local fields.

It should be noted that this is a simplified sequence of events. For example, the Serials Upgrade Application should employ its search and match algorithm against both the CONSER database and the serials union catalogue to discover appropriate matches. There will also be a procedure for receiving updates from the local library contributors to the serials union catalogue, via the upgrade application, to include changes in holdings, deletions (which will require a record deletion from the serials catalogue if and only if that library is the sole location), and new additions (which will go through the matching procedures carried out when adding a new library's data to the catalogue). It is also an assumption that both CONSER and serials catalogue records will be freely available to (at least) participating libraries, and mechanisms will be required to allow search, retrieval and download of records.

The development of the architecture as illustrated requires the following elements:

- management of the file exported by the library to ensure records are structurally correct;
- loading and indexing of CONSER database;
- component to allow management of matching/identification rules;
- component to perform the batch run and save to the import file and the serials catalogue load file;
- cataloguer interface and Z-clients to allow records not processed by the batch system to be manually upgraded;
- user interfaces for rules management and system integration.

It is theoretically possible not to insist that upgraded bibliographic information be loaded back to local catalogues. However, the improvement of local records is at the heart of improving services, and would bring a number of advantages: the simplification of updating; the supply of an improved and consistent record at local level; an aid to on-the-fly de-duplication during a search of, for example, COPAC and the serials catalogue; ensure that local records are in a fit state to be hospitable to extended services as they are developed. The free availability of CONSER and the serials catalogue as a record reservoir should encourage the use of quality records by participating libraries.

Links to existing serials union catalogues – for example, the University of London Union List of Serials, SALSER, Lamda or MagNet – will need to be considered in the light of the research and development into distributed catalogues outlined below. The implementation of serials title indexes offered by the library systems suppliers (although this may come at a price of up to £5000 per library), and the application of the recently agreed extension of Bath Profile attributes to include serials titles, are also important steps to encourage the practical dissemination of serials information nationally via the clumps and other distributed catalogues.

7.4. Clumps research and development

To move towards a National Union Catalogue that has linked physical and distributed elements, two recommendations are made with reference to virtual union catalogues. The first is that a 'wait and see' attitude is adopted to implementation by vendors of the Bath Profile and the associated Holdings Schema. Secondly, further, focused, research should be commissioned into the following areas:

- the creation of Z39.50 points of presence into existing clumps
- a workable, preferably standards-based, internationally accepted, solution into the recognition of individual targets, when clumps are searched
- load-testing of clumps to establish any stress points placed on individual library systems, and the effects of these, as a result of substantial UKNUC activity
- the relationship between the different approaches to searching and presentation of results taken by the clumps projects and a consideration for any harmonisation in the context of a UKNUC, possibly through the establishment of a pilot system
- methods for handling results retrieved from physical and virtual catalogue elements and presenting these to users.

The relevance of this further work to the National Union Catalogue is in the *co-ordination* of clump activity rather than the individual approaches that have been taken hitherto (notwithstanding the co-operation that is known to have taken place). Ways are required of ensuring that the National Union Catalogue looks coherent and operates in a coherent manner and it is recommended that all four eLib clumps, and possibly the public libraries Co-East clump, should participate in this further research. Should the outcomes of this further work prove successful, consideration should be given to establishing a UK Z39.50 Office so that expertise can be retained and the costs of clump creation, itemised in section 6.4 above, possibly be reduced. No allowance has been made for this in the set-up costs for a National Union Catalogue.

7.5. Services

It has not been possible to consider, in depth, services that might flow from a National Union Catalogue. These were identified in the conceptual model as: physical access by researchers to other libraries; inter-library loan; full-text access; record supply; SDI services; re-directed search; generation of statistical reports; and ordering. While these may well not be available initially, it is likely that the presence of a well-regarded union catalogue will in itself facilitate the development of services such as these – and others as yet undefined – in the future.

One recent development which will ease the provision of services within the union catalogue framework is the OpenURL. In the serials context, for example, users at a library with an OpenURL resolver will be able to move from the local OPAC to the National Union Serials Catalogue when searching for a journal title not available locally. Similarly, when searching a database of journal articles, it should be possible to offer the National Union Serials Catalogue as a target in addition to the more usual local sources. Further, an OpenURL resolver could be provided for the National Union Catalogue itself, so that links *from* the catalogue could be provided as well as links *to* the catalogue. The links made via the OpenURL are sensitive to good quality bibliographic data and so the emphasis on this at the creation stage will repay in benefits from offshoot services.

The Join-Up family of projects under the auspices of the DNER is aiming to smooth the path of the user seeking to obtain the fulltext of a journal article, linking together the processes of discover, locate, request and deliver. A National Union Serials Catalogue will form a vital part of this chain as a major source of location information linked through either an OpenURL resolver as described above, or as a Z39.50 target at which the ZBLSA Join-Up locating service can point.

The ability to tap into a National Union Catalogue for user-initiated inter-library loans (whether for monographs or serials) sounds a distinct possibility but presents a number of operational challenges. As the interloan request would only be approved at the local level, there would be a requirement for links between national and local software, together with authorisation facilities. And, while users must have a password/username combination which confers eligibility to place

such requests, this does not necessarily grant the right to have the interloan progressed. In other words, any national authentication database must not only recognise that an individual belongs to a particular HE institution (say) but that they fall within an eligible ILL group whose criteria will differ – sometimes widely – between different institutions. This is a major issue and clearly outside the scope of UKNUC but it should be mentioned that RIDING – which has the software to place ILLs from items found in a search – has so far failed to resolve this. The Librarians of the institutions were not prepared to offer this service as it raises user expectations that might have to be dashed when the request is mediated and rejected further down the line.

There are considerably fewer problems in utilising the National Union Catalogue as a source of locations by interloans librarians as this could operate through dedicated software though this would need to be developed or adapted from existing products.

7.6. Criticisms of a National Union Catalogue

A number of negative comments were received during the course of the Feasibility Study, some expressed at the Concertation Day. Several of these questioned the purpose of a UK National Union Catalogue, some suggested alternative technical solutions, and others that funding could be better spent on different services; one correspondent linked all three strands together.

The difficulty in answering questions about ‘purpose’ is that they often assume one particular mind-set – the one being proposed – applies to all situations: the academic will only search in this way, must be constrained to look at local resources first, will never travel to explore other collections, will always request an inter-library loan, etc. etc. In contrast, when considering the building of a UK National Union Catalogue the intention of the Study Team has been to keep as many options open as possible, to not restrict the software design at an early stage when the full potential of the resource is unexplored. This was the reason for creating the conceptual model and considering the range of ways in which the National Union Catalogue might be used.

Of course, ‘use’ is not necessarily the same as ‘purpose’ and for the UKNUC this has already been covered at the beginning of this report (section 2) as ‘primarily a vehicle which supports research’. It will do this most particularly by creating a one-stop-supermarket for existence, availability and citation checking of monographs, serials and other formats. The time-savings to academics, researchers, postgraduates, undergraduates, and librarians in being able to search, at the same time, forty major, research-heavy, library collections in the UK *and* the British Library collections, instead of doing this one-by-one, will be substantial. The (unsolicited) comment received from an academic in the questionnaire survey well indicates the potential advantages:

- An indication of the great benefits that are at stake here is the improvement in academic research that has resulted from COPAC. I actually find it more useful now than I do [my own local] OPAC (which used to be my prime source of information). Is it possible for UKNUC to be an extended COPAC?

Alternative technology criticisms often rest on the assumptions that a) an internet search engine, harvesting or similar approach can be used, or b) given the rate of progress, something easier will come along and make the creation of expensive large-scale databases redundant. As pointed out earlier in this report, no *robust and tested* technologies are currently available which handle the complexities of bibliographic records, will allow a range of services to be offered by the UKNUC, and which will provide significant time-savings to the user. Ironically, some ‘lightweight’ solutions that have been suggested as alternatives actually need a large scale database such as the one proposed in this report to themselves operate effectively.

Better uses of funding is a criticism raised at all new (and some existing) services and quite often depends on the perspective of the critic, who frequently does not see advantages passing to them on this occasion. The Feasibility Study was not asked to compare and rate information services for researchers and, accordingly, finds it difficult to make any comments here. However, when asked about relative costs and expenditure in the survey, librarians were generally supportive of a UKNUC and were certainly not dismissive in ways that have been suggested in some quarters. The level of funding required to establish the foundation National Union Catalogue for the UK over a three year period is modest, as are the upkeep costs, and feedback, and a weather eye on technological developments will ensure that future expansion takes place in accordance with users' needs.

In summary, both the Steering Group and the Study Team considered at length the range of important issues, positive and negative, which arose from the Concertation Day and other discussions with key players. Most important of all, however, is the very strong support for the National Union Catalogue indicated in the questionnaire surveys of both academics and librarians right across the UK: there is convincing evidence here of the major benefits of the resource to the whole UK research community and beyond. Many very supportive comments have also been received by the Study Team from librarians and academics throughout the UK and abroad showing general enthusiasm, requesting that their collections be included in the catalogue, and suggesting that such a resource is long overdue:

- this is excellent news! The dream of old fogeys like me coming into the profession nearly 40 years ago are now being realised, just at the time I am at last thinking of giving it all up, and fading away into slippered retirement. I wish your project every success – it will bring huge benefits if it succeeds, which I profoundly hope it does.

7.7. Cost and timescale

On the basis of the recommended model, it will take three years to create the foundation for a National Union Catalogue for the UK and undertake necessary research and development work into distributed systems. However, the programme as costed will begin the integration of the British Library and the libraries of twenty further institutions with COPAC, coming on stream after the first 18 months and being fully available after 30 months. The Serials National Union Catalogue system will be operational within the first six months and the work of populating the bibliographic records with holdings data will start immediately afterwards. Three years have been chosen quite deliberately as the time to create the foundation Catalogue of the specified size and develop the UKNUC Portal but also to undertake the research and development needed to determine the position of the clumps.

The outline costings have been built up in association with Manchester Computing, based on COPAC experience, and discussions that have taken place between Manchester Computing and the British Library on work associated with integrating the BL catalogues and COPAC. Alternative, provisional, figures for hosting elements of the recommended model at the British Library have also been received; they are greater than those presented below. The clumps R&D costs are broadly indicative.

It is stressed that the costings presented are indicative only and make no assumption that the hosting of an integrated National Union Catalogue would necessarily be through these organizations. Should in principle funding be agreed, a full specification and operational requirement would need to be made and an open procurement exercise undertaken.

UKNUC costs over first three years (excluding existing COPAC costs)

Year 1						
System type	Staffing	Hardware	Software support	T&S	Publicity	Total
Monographs etc.	£130,000	£43,000	£0	£3,000	£1,000	£177,000
Serials	£200,000	£21,000	£50,000	£1,500	£500	£273,000
Clump R&D	£50,000	£5,000	£30,000	£2,000	£0	£87,000
Total Year 1	£380,000	£69,000	£80,000	£6,500	£1,500	£537,000
Year 2						
System type	Staffing	Hardware m/tenance	Software support	T&S	Publicity	Total
Monographs etc.	£130,000	£3,000	£0	£3,000	£2,000	£138,000
Serials	£310,000	£1,500	£5,000	£1,500	£1,000	£319,000
Clump R&D	£50,000	£0	£0	£2,000	£0	£52,000
Total Year 2	£490,000	£4,500	£5,000	£6,500	£3,000	£509,000
Year 3						
System type	Staffing	Hardware m/tenance	Software support	T&S	Publicity	Total
Monographs etc.	£100,000	£3,000	£0	£3,000	£2,000	£108,000
Serials	£280,000	£1,500	£5,000	£1,500	£1,000	£289,000
Clump R&D	£30,000	£0	£0	£0	£0	£30,000
Total Year 3	£410,000	£4,500	£5,000	£4,500	£3,000	£427,000

Thus, the cost, over a three year period, of establishing the foundation National Union Catalogue for the UK is £1,473,000, *in addition to current COPAC support*. The serials costs are higher than those for monographs because of the emphasis on the upgrading of records. Although aided by the provision of the CONSER database, this is still an inevitably staff-intensive process.

For this estimated cost of £1,473,000 the foundation National Union Catalogue would comprise the British Library, twenty-two of the major research universities of CURL, collections of high research interest included in a further twenty cross-sectoral libraries in the UK and a strong building block of a National Union Serials Catalogue, including around 750,000 records from the British Library and the other two National Libraries, and up to 450,000 other serials record locations (backed by the availability of around 1,000,000 CONSER records). For comparative purposes, CURL libraries hold about 570,000 serials titles – but it is not suggested that the foundation National Union Serials Catalogue includes all CURL libraries – and the Scottish participants in SALSER possess about 245,000 titles (of which around 100,000 are owned by the National Library of Scotland).

The annual running cost, after the initial three years, would be approximately £114,000, although this would be increased if there were a continuing programme of serials record upgrading. At the end of the three year foundation term, results of the tests on distributed systems and the linking of these with the physical elements would be delivered, together with early user feedback on the performance of the systems, allowing one of three directions to be taken:

- pursue the combined physical-virtual development, but only on the basis that all the current questions on virtual catalogues are answered satisfactorily and that the two elements – physical and virtual – can be integrated for users in their presentation and operation;

- accept that the physical-virtual model is still not mature enough to be the core of a national resource and start to build further physical systems; or
- decide that further development of the UK National Union Catalogue beyond the initial three-year programme will not be pursued.

Should the last decision be taken, a resource of lasting and major use to researchers will still have been created, at a modest operating cost.

Where regional initiatives link to national resources, a view would need to be taken on the level of Funding Council support. This is particularly germane to the relationship of clumps and physical union catalogues where it might be decided that, as clumps are established to benefit primarily the local and regional communities, funding for their creation and operation should not be derived from the national purse. Although the eLib clumps were established with JISC money, this was for investigatory purposes, to test the viability of Z39.50, and it was never the intention that funds would be made available either for new clumps or to continue the upkeep of the original four. While at the current time there might be few incentives for further clump creation, regional initiatives might change the political imperative and economics associated with this, transferring the cost to groups of local institutions and away from the Funding Councils. However, where costs were associated with linking the clump – which must be built to outline ‘national’ technical specifications – to the National Union Catalogue, these costs, and possibly some annual running charge, might be supported from the centre.

7.8. Towards a *National Union Catalogue*

As currently presented, the recommendation falls short of being a fully comprehensive National Union Catalogue for the UK. The Study Team believe that the recommended model, whether this is finally a totally physical architecture or develops to become a mixture of the physical and virtual, can be expanded to encompass a large cross-section of libraries in the UK. The speed at which this occurs will depend on the pace of technological and political change, and associated financial and organisational issues, but the groups of libraries that must be considered in a truly *National Union Catalogue* are discussed below. The section concludes with some general comments on the necessity for links to resources outside the National Union Catalogue.

7.8.1. Expanded coverage of HE libraries

Within HE alone there are over 150 institutions, varying markedly in size and research income from small, specialised colleges at one extreme to the research-led multi-disciplinary institutions at the other. Operationally, there would be little difficulty in expanding the physical union catalogue to encompass all of these (though some might fall into the category of ‘small libraries’ considered in the next section). The additional set-up cost to add all those HE libraries not included in the foundation Catalogue would be approximately £1,100,000, with an additional annual operating cost of around £190,000. Implications for the Serials Union Catalogue would have to be considered on a library-by-library basis but would involve additional investment: proportional to the serials costings in section 7.7 the figures will not be so high, both because the absolute number of serial records to upgrade per library will fall quite fast moving on from the large research-intensive libraries, and because the number of unique titles not already present in the National Union Serials Catalogue is likely to fall even faster.

Until the research and development has been carried out, no recommendations will be made about the inclusion of these libraries in any additional (or existing) clumps.

7.8.2. *Public libraries, special libraries and small libraries*

Public libraries and special libraries must be embraced as the *National Union Catalogue* expands, although, as mentioned earlier, it is recommended that they are represented in the foundation of the resource. Small libraries may bring with them special difficulties and are considered briefly at the end of this section.

Public libraries are an important national resource, and a key element in the expanding regional agendas. Although they did not rate highly in comparison to other libraries among the academics surveyed, they do have valuable collections of research importance. And their role in lifelong learning and as delivery points for the People's Network is recognised by the government. The alternative to inviting chosen – and by inference a few – public libraries to participate in the *National Union Catalogue* is to reach agreement for the two major databases in this sector to be linked to the initiative. These are V3.Online and Unity.

V3.Online is an established database of 4 million titles (approximately 40 million locations) though in addition to records of some 100 LASER libraries, this includes BNB, BLDS records and library suppliers data; relatively few serials records are available. As it only includes one academic library it would form a very useful component of a UKNUC without duplicating a large number of other resources and would help fill out resources in the area of London and the South East as well as contributing specialist and prestigious materials. In tests, the system reflected a high quality and consistent user interface (though operable only through a specialist browser 'plug-in'), and in its current form it does not have a Z39.50 point of presence. Incorporation of the data into a UKNUC would need to be either by data supply from LASER to UKNUC, or through LASER creating a Z39.50 server to search. However, towards the end of the Feasibility Study, the demise of LASER was announced and, though positive statements were issued, the future of V3.Online was uncertain. By the time the UKNUC is reviewed during the three year foundation stage, the future of V3.Online should be clearer and it is recommended that discussions be held to consider if financial and technical agreements can be reached to permit inter-working of the two systems. It is important that V3.Online be kept on the agenda for a *National Union Catalogue*.

Unity, the Combined Regions Database, provides bibliographic citation and location data for a wide range of local, regional and national library materials, as well as BL records. The Regional Library Systems charge the libraries in their area for co-ordinating bibliographic records and then contribute these to the central database which contains over 10 million catalogue records representing 35 million locations (though, as with LASER, few serials records are included). With the demise of Libris, the company who wrote the original software and established the system, Talis Information have joined the Combined Regions as technical and commercial partner and are actively developing a networked version, UnityWeb. At the time of writing, only a demonstration version of UnityWeb was available for viewing. Records of the current holdings of books and other materials are maintained but, created as an inter-library loans tool, the database is not comprehensive and non-loanable items will, in general, not appear. Figures on the holdings of academic libraries in Unity show large discrepancies, with fractions of collections recorded in many instances, but Talis and Unity are actively pursuing ways to fill these and other gaps and improve the sparse and inconsistent bibliographic and holdings data. If Unity data were loaded to the UKNUC, then this could either be through a data supply to UKNUC, or through Unity opening up the UnityWeb Z39.50 server to search.

It is clear, from discussions, that the Unity Management Board see UnityWeb itself growing towards a 'National Union Catalogue', though on the basis of the tests of the Study Team, there are real concerns about its utility for academic institutions in general and researchers in

particular. Concerns about quality of bibliographic data, comprehensiveness of holdings from individual contributing libraries, and the availability of the wide number of search points required would all need to be seriously addressed. A number of discussions have been held to consider ways of linking UnityWeb and UKNUC but these have failed to reach agreement, largely because both systems are at such an early stage of development. In particular, the Unity Management Board have developed a business model based on the purchase of an agreed number of concurrent accesses and, despite alternatives being broached, they consider that their financial viability will be undermined if a different model is adopted. Such a model is so opposed to the UKNUC principle of 'free at the point of access' – and potentially limits free access from any location – that further explorations were stalled. However, there does exist the possibility of the two systems coming closer together when both have become established and it is recommended that possibilities for closer working are explored with exchange of experience meetings taking place at six-monthly intervals.

'Special libraries' could form a unique component in a UK National Union Catalogue, providing access to data (and collections) not available in universities. A meeting has been held with a number of special librarians – e.g. Wellcome Institute; British Medical Association; National Institute for Medical Research; Institute of Advanced Legal Studies; London Business School; V&A Museum; British Film Institute; Imperial War Museum; Natural History Museum – who expressed the strong view that a UKNUC would significantly benefit from their (combined) specialist collections. A clear message was also given that some of these libraries wished to exploit their collections more widely. But there was genuine concern about the impact of a UKNUC: while librarians were enthusiastic, their libraries were generally small, serving just their own institutions and they were not sure that they could cope with any extra demand, in person or via interloans, that might ensue. On more than one occasion it was stated that funding to bring these specialist libraries into a UKNUC should be provided by HE and there was clearly a feeling of injustice that, for example, the RSLP was not prepared to financially support institutions outside HE while espousing the principle of cross-sectoral co-operation. Any such arrangement associated with a UKNUC will be difficult to achieve, though the development of Service Level Agreements might be one way of ensuring that, for a modest financial outlay, HE obtains a medium-to-long-term access to otherwise inaccessible collections. Several other special libraries contacted the Team during the course of the Study and, to give an idea of the interest, some of these are included in Annex D.

Small libraries have been identified, largely through the Music Libraries Online clump, as bringing with them a number of particular features, notably staff that often cannot provide the level of technical support necessary, and library management systems that do not have associated Z39.50 targets – making it difficult for them to participate in virtual clumps – and do not provide output in MARC-exchange format – making it difficult for them to participate in physical catalogues. However, there would appear to be no reason why a number of small libraries should not be included in the broader physical National Union Catalogue as long as the additional overhead of resolving unusual file formats was accepted. Whether these libraries contribute to bibliographic record supply would have to be decided on a library-by-library basis based on the overall benefit that would accrue from this and the quality of the records.

7.8.3. Outside the National Union Catalogue

The focus of this Feasibility Study has been a National Union Catalogue for the UK but research is not undertaken on just a national basis and neither is teaching and library co-operation with, for example, the franchising of courses, developments such as the e-university, and links between CURL and Research Libraries Group. Time has not permitted the exploration of international links into and from a UKNUC but there are a large number of international organisations and

resources with which the UKNUC – once established – could form co-operative agreements for mutual benefit. In the questionnaire survey of academics, twenty-one respondents emphasised the importance of looking outside the UK, linking to library catalogues across the world, or to other resources. One academic caught this overall mood very well:

- It would be of great importance if a National Union Catalogue for the UK was then connected/networked with other EU countries' NUCs. Has anybody thought that we do not live alone in this world? There are very important scientific developments taking place in other EU countries.

Links to related national initiatives such as Digital Scotland must also be considered. And an associated link might be made to Amazon.com, notwithstanding its limitations for researchers and the lack of interest shown in the questionnaire survey by academics for purchasing books.

8. CONCLUSIONS

The conclusions reached at the end of the Feasibility Study for a National Union Catalogue are:

- after a detailed technical assessment and cost comparison of physical and virtual union catalogues, the physical model currently offers to users significant coherence and reliability at no cost disadvantage to funders.
- virtual union catalogues, as typified by the eLib clumps, are currently at an embryonic phase and, while promising much through protocols such as Z39.50, cannot yet provide the level of accuracy, speed of response or robustness required of a national system.
- as the technology develops, and if international agreements such as the Bath Profile are embraced by library system vendors, the drawbacks of virtual systems may be reduced and their place in a national system can be re-evaluated.
- it is recommended that a foundation National Union Catalogue be established initially as a physical model but that the architecture incorporates links from the outset to potential virtual systems.
- serials bibliographic and holdings data inadequacies seriously inhibit the development of a National Union Serials Catalogue, and it is recommended that a programme of record improvement employing the CONSER serials database as a major record source be initiated.
- the content of the foundation National Union Catalogue will consist of the twenty-two major research libraries of COPAC, the British Library catalogues (including existing union catalogues such as the English Short Title Catalogue), the catalogues of twenty cross-sectoral libraries with collections of identified importance to research, and a National Union Serials Catalogue.
- the National Serials Union Catalogue will include the holdings of the British Library, the National Libraries of Scotland and Wales, major university research libraries, and a number of specialist libraries, including up to 1,200,000 record locations.
- the foundation catalogue will offer the facility of bibliographic record supply, subject to agreements being reached with all relevant parties.
- improved search and navigation facilities be provided for resource discovery, particularly in the areas of subject search, geographical/regional search and format search.
- it is further recommended that a clumps research and development programme be undertaken to investigate dedicated issues as they relate to a national rather than purely regional context.

- the creation of the foundation National Union Catalogue will take three years and cost around £1.5 million; thereafter, operating costs would be approximately £114,000 per annum. Both these costs are in addition to the current financial support for COPAC.
- at the end of the first three years, and knowing the results of the clumps research and development, a decision will be taken on the future directions of a National Union Catalogue.
- the provision of services by linking with OpenURL initiatives and the DNER Join-Up developments would appear realistic; services associated with (particularly user-originated) interloans require a national authentication scheme.
- many libraries in the UK will be excluded from the foundation National Union Catalogue, and every attempt should be made to embrace these at the earliest opportunity after the initial three years. The addition of all other HE libraries would cost an additional £1,100,000, with an additional annual operating cost of around £190,000. Public libraries have extensive and important resources, though the two main databases containing these are undergoing substantial change which prevents their immediate involvement. The wide range of special libraries in the UK have expressed particular interest in the National Union Catalogue and, having many unique and specialised collections would form an important component in the whole. Discussions should be enabled with public libraries and special libraries in particular to improve dialogue and involvement in future National Union Catalogue developments.

ANNEX A: MEMBERSHIP OF FEASIBILITY STUDY STEERING GROUP

Mr Derek Law, University of Strathclyde (Chair)
Professor Peter Brophy, CERLIM
Ms Gill Davenport, RSLP
Mr Lorcan Dempsey, JISC
Mr Fred Guy, National Library of Scotland
Mr Mike Hosking, Cambridgeshire County Libraries
Dr Paul Miller, UKOLN
Mr Ronald Milne, RSLP
Mr Bernard Naylor, University of Southampton
Mr Chris Rusbridge, University of Glasgow
Mr Geoff Smith, British Library
Mr Robert Smith, British Library
Ms Jean Sykes, LSE

Study Team:

Mr Rob Bull, Crossnet Systems
Dr Clive Field, University of Birmingham
Mr Michael Hannon, University of Sheffield (Project Director)
Mr Tony Kidd, University of Glasgow
Mr Peter Stublely, University of Sheffield (Project Manager)

ANNEX B: LIST OF COMPLETED REPORTS IN THE FEASIBILITY STUDY

None of the reports listed below are publicly available in their current form. They have been limited to members of the Study Team and the Steering Group.

Peter Stubbley, Tony Kidd and Rob Bull, *Union Catalogue trends, existing systems and technology: WP1 (Background Analysis) Report*, September 2000. Ref: NUCrev.doc. 45 pages.

Peter Stubbley, *Survey of academic staff, postgraduates and researchers as users of a National Union Catalogue: analysis of results*, January 2001. Ref: NUCSurU.doc. 25 pages.

Tony Kidd, *Survey of library staff as users of a National Union Catalogue: analysis of results*, January 2001. Ref: NUCSurL.doc. 26 pages.

Tony Kidd and Peter Stubbley, *Liaison with key players: North American visit report; separate Report for WP2b (Key player requirements)*, November 2000. Ref: NUCkeyUS.doc. 18 pages.

Peter Stubbley, Tony Kidd and Rob Bull, *Liaison with key players: first visit reports: WP2b (Key player requirements)*, January 2001. Ref: NUCkey.doc. 53 pages.

Peter Stubbley, Tony Kidd and Rob Bull, *Union Catalogue trends, existing systems and technology: WP3 (System Analysis) Report: Book 1 – methodology, testing and conclusions*, April 2001. Ref: NUCsysB1.doc. 76 pages.

Rob Bull and Tony Kidd, *Union Catalogue trends, existing systems and technology: WP3 (System Analysis) Report: Book 2 – test results and supporting information*, April 2001. Ref: NUCsysB2.doc. 98 pages.

Tony Kidd and Rob Bull, *A Serials Union Catalogue for the UK: a compilation of relevant sections from other UKNUC Feasibility Study reports*, April 2001. Ref: NUCser.doc. 25 pages.

ANNEX C: LIST OF PEOPLE CONSULTED

Chris Anderson, New Opportunities Fund
Bob Aspey, British Library
Mary Auckland, London Institute
Chris Bailey, CURL
Toby Bainton, SCONUL
David Balatti, National Library of Canada
Chris Batt, Re:source
Neil Beagrie, Joint Information Systems Committee (JISC)
Anne Bell, Lamda and King's College London
Marion Birch, Institute of Advanced Legal Studies
Kerry Blinco, Infostructure Consulting Services
Julie Booth, ExLibris
David Bradbury, British Library
Rune Brandshaug, BIBSYS
Sue Brown, Library Association
Per-Morten Bryhn, Oslo University
Peter Burnhill, EDINA
John Byford, British Library
Mark Carden, Innovative Interfaces Inc
Julia Chruszcz, Manchester University Computing
Neil Clements, Talis Information
Brian Clifford, University of Leeds
Ric Collins, Manchester University Computing
Vanessa Conte, Higher Education Funding Council for England (HEFCE)
Anita Cook, OhioLINK
Verne Coppi, Endeavor Information Systems Inc
Dr Shirley Cousins, Manchester University Computing
Marc Daubach, ExLibris
Laurie Davidson, Innovative Interfaces Inc
Mike Day, UMIST
Ray Delahunty, Endeavor Information Systems Inc
Lorcan Dempsey, DNER
Sue Detillio, OCLC
Matthew Dovey, University of Oxford
Simon Dunlavey, DS Limited
Sean Dunne, Manchester University Computing
Gordon Dunsire, Napier University
Laurie Edmonds, Endeavor Information Systems Inc
Helen Edwards, London Business School
Neil Edwards, British Library
Nancy Elkington, Research Libraries Group
James Elliot, British Library
Rob Froud, Chair, Unity Management Board and SWRLS
Christine Fyfe, University of Warwick Library
Robert Gent, Unity Management Board and Derbyshire County
Peter Gethin, Sirsi Limited
John Gilby, M25 Link
Anne Gilliland, OhioLINK
Moirra Goff, British Library
Rhidian Griffiths, National Library of Wales
John Gray, Linen Hall Library

Stephen Gray, Talis Information
Andrew Green, National Library of Wales
John Hall, Durham University and CURL
Tony Hall, Talis Information
Dave Harris M25 Link and Queen Mary and Westfield College
Joanna Harrison, Joint Information Systems Committee (JISC)
Mary Heaney, University of Wolverhampton
Chris Heaton, Cambridgeshire County Libraries
Frances Hendrix, LASER
Janet Higinbotham, National Library of Scotland
Jean Hirons, CONSER
Tim Hobbs, Leicester University
Marion Hogg, Music Libraries Online
Kate Holliday, Unity Management Board and Yorkshire and Humberside Region
Mike Hosking, Cambridgeshire County Libraries
Kathy Hughes, National Library of Wales
Paul Jacobs, CAIRS Limited
Graham Jefcoate, British Library
Stephen Jordan, Endeavor Information Systems Inc
Ellen Katic, National Library of Canada
Cathy Kellum, OCLC
Stephanie Kenna, British Library
John Killingbeck, British Library
Andrew King, DS Limited
Janet Lees, OCLC Europe
Ray Lester, Natural History Museum
Eileen Lim, National Library of Canada
Michael Long, Unity Management Board and Information North
Carrol Lunau, National Library of Canada
Dr. Liz Lyon, UKOLN
Nigel Macartney, University of Ulster
Bob McKee, Library Association
Tony McSean, British Medical Association
Ian Manson, Geac Computer Corporation Limited
Ann Matheson, National Library of Scotland
Anne Mealia, Cambridge University
Paul Miller, UKOLN
Mary Morley, Loughborough University
Paul Moulard, ELiAS (Extended Library Access Solutions)
Brian Murphy, M25 Link and Queen Mary and Westfield College
Robin Murray, Fretwell-Downing Informatics
Myrtle Myers, OCLC
Bernard Naylor, University of Southampton
Cate Newton, National Library of Scotland
Dennis Nicholson, CAIRNS
Fraser Nicolaidis, M25 Link
Frank Norman, National Institute for Medical Research
Bill Oldroyd, British Library
Glenn Patton, OCLC
David Pearson, Wellcome Trust
Françoise Pelle, ISSN International Centre
Janet Peters, University of Wales College Newport
Tom Peters, Committee on Institutional Cooperation
Sean Phillips, University College Dublin

Dennis Pilling, British Library
Stephen Pinfield, University of Nottingham
Andy Powell, RDN
Bridget Powell, Unity Management Board and SWRLS
Stuart Rawson, The Queen's University of Belfast
Norman Russell, Queen's University Belfast
Deborah Ryan, Unity Management Board and NWRLS
Harry Samuels, Northwestern University
Howard Shoemark, British Library
Debby Shorley, University of Sussex Library
Barbara Shuh, National Library of Canada
Bill Simpson, Trinity College Dublin
Kate Sloss, Music Libraries Online and BLPES
Neil Smith, Fretwell-Downing Informatics
Peter Smith, LASER
Robert Smith, British Library
John Sparrow, CAIRS Limited
Jean Sykes, M25 Link and BLPES
Hugh Taylor, Cambridge University Library
Ray Templeton, British Film Institute
Bridget Towler, University of Hull
Marguerite Tubby-Hille, EDINA
Paul Turner, Unity Management Board and Hampshire County
Martin van Muyen, Pica
Jenny Walker, Ex Libris
Michel Wessling, Pica
Geraldine White, V&A Museum
David Whitehurst, UMIST
Ann Whittle, Talis Information
Dave Wilkie, Unity Management Board and BL Bibliographic Services
Mary Wilkinson, Imperial War Museum
Terry Willan, Talis Information
Neil Wilson, British Library
Susi Woodhouse, Re:source
Robin Yeates, LASER

ANNEX D: LIBRARIES EXPRESSING INTEREST IN PARTICIPATING IN A UK NATIONAL UNION CATALOGUE

All (edited) comments received by e-mail.

British Universities Film & Video Council: I would be most grateful if you would put us on your list of potential contributors with regard to the cataloguing of moving image media in the feasibility study recently announced.

Specialist collections: the SPRU Library – Science and Technology Policy Research. There are two such libraries at the University of Sussex, the other being the Institute of Development Studies, with no affiliation to the main University Library. Our opinions, needs, and contribution will be missed unless we are contacted directly. Presumably this sort of situation prevails at other institutions.

The Union List of Art, Architecture and Design Serials, hosted by the National Art Library for the Art Libraries Society, ARLIS/UK & Ireland <http://www.nal.vam.ac.uk/ulist/>. Been awarded funding from the BL's Co-operation and Partnership Programme, to develop the project further.

Research Council libraries: these libraries often get sidelined in such projects for various reasons – principally the fact that funding programmes such as RSLP and the BL Cooperation and Partnership do not explicitly refer to them in their briefs. Ignoring these important collections would be to do the research community a disservice. Collaboration between HE and RC researchers is widespread, and RC libraries are involved in providing services to HE researchers either directly or via arrangements with local academic libraries. Recognition of the importance of these collections was part of the Anderson Report.

Share The Vision, a partnership of voluntary sector agencies and the British Library, Library Association, SLIC, SCONUL and SCL, which is attempting to enhance access to library and information services for visually impaired people. The National Union Catalogue of Alternative Formats is an essential building block in developing accessible library services, now renamed REVEAL: The National Database of Resources in Accessible Formats and already available via V3.Online and Unity.

Research institutes affiliated to the Research Councils: at Horticulture Research International there is an excellent collection within specialised areas of horticultural practice, particularly relating to fruit and woodland crops. Plant physiology, storage and plant science are covered and the collection includes books, serials and a collection of pamphlets 50K+ in size of international interest. It is a collection I would like to make accessible to a wider community of researchers.

Fire Service College: I think that the principle of this is excellent in theory. We have a unique collection of materials on fire engineering technology and related fields. Have you any thoughts about adding specialist collections to the catalogue?

Royal College of Physicians: this email is a plea to not forget the Medical Royal Colleges and their unique specialist collections when considering your final report on the feasibility of a UKNUC. Our manuscript collection is extensive – it includes such items as the doctors notes on the last fatal illness and post-mortem of James I (or VI), as well as non-medical material such as the exquisite Wilton Psalter from 1240, and one of only 18 extant copies of Caxton's first book *The history of Troy*. The other Medical Royal Colleges have similar collections which meld the old with the new. However, your study did not appear to include us. Is it possible to be surveyed at this late stage? Or for us (collectively) to provide you with an addendum?

ANNEX E: BIBLIOGRAPHY

All the following works were consulted and used to varying degrees during the Feasibility Study.

- Allen, Barbara McFadden (1999). The CIC: using collaboration to advance high technology initiatives. *D-Lib Magazine*, vol. 5, no. 2, (February) <<http://www.dlib.org/dlib/february99/02allen.html>>.
- Allen, Barbara McFadden (1998). CIC and OCLC transform interlibrary loan services with new agreement. *OCLC newsletter* (November/December), pp. 8–9.
- Anderson, B. (1997). CONSER on the Internet: facilitating access to serials information. *Serials Librarian*, vol. 32, no. 1/2, pp. 77–94.
- Arms, W. Y. (1973). Duplication in union catalogues. *Journal of Documentation*, vol. 29, no. 4, pp. 373–379.
- Ayres, F. D. *et al.* (1996). USBC: its origin and evolution. *Journal of Librarianship and Information Science*, vol. 28, no. 2, (June), pp. 83–91.
- Bakker, Trix (1998). Resource sharing in Dutch academic libraries. *Resource sharing and information networks*, vol. 13, no. 2, pp. 15–27.
- Bath Group (2000). The Bath Profile: an international Z39.50 specification for library applications and resource discovery <<http://www.ukoln.ac.uk/interop-focus/bath/1.1/>>.
- Birch, Katie and Pettman, Ian (2000). Linking distance learners with the UNiverse. In *Libraries without walls 3: the delivery of library services to distant users*. London, Library Association Publishing. pp. 166–174.
- Blue Angel Technologies Inc. (n.d.). *An evaluation of Z39.50 within the SILO project*. <<http://server.silo.lib.ia.us/bluang.html>>
- Brack, Verity (1999). Service developments at the RIDING Z39.50 Gateway. *New Review Of Information And Library Research*, vol. 5, pp. 135–144.
- Bryant, Philip (1997). *Making the most of our libraries: the report of two studies on the retrospective conversion of library catalogues in the United Kingdom, and the need for a national strategy*. Boston Spa: British Library. British Library Research and Innovation Report no. 53.
- Cameron, Jasmine (1992). National bibliographic control of serials in Australia, 1980–1990. *Serials Librarian*, vol. 22, no. 3/4, pp. 291–296.
- Chapman, Ann; Kingsley, Nicholas and Dempsey, Lorcan (1999). *Full disclosure: releasing the value of library and archive collections*. Library and Information Commission Report no. 10.
- CIPFA (1999). *Public library statistics 1998–99 actuals*. London: Chartered Institute of Public Finance and Accountancy.
- Coffman, Steve (2000). Big 12 Plus Libraries and OCLC join forces on web-based interlibrary loan system. *Newsbreaks & conference reports* <<http://www.infotoday.com/newsbreaks/nb000207-2.htm>>.
- Co-operative Automation Group (1984a). Progress towards a UK Library Database System. *Vine* no. 53, (April), pp. 45–46.
- Co-operative Automation Group (1984b). A statement and a standard. *Vine* no. 57, (December), pp. 34–47.
- Cousins, Shirley Anne (1998). Duplicate detection and record consolidation in large bibliographic databases: the COPAC database experience. *Journal of Information Science*, vol. 24, no. 4, pp. 231–240.
- Coyle, Karen (2000). The virtual union catalogue: a comparative study. *D-Lib Magazine*, vol. 6, no. 3, (March) <<http://www.dlib.org/dlib/march00/coyle/03coyle.html>>.
- CURL and Crossnet Systems (1998). *CURL Z39.50 feasibility study*. <<http://www.curl.ac.uk/projects/z3950.html> >

- Delsey, Tom (2000). *The library catalogue in a networked environment*. Presented at the LC Bicentennial Conference on Bibliographic Control for the New Millennium, November 15–17, 2000. <http://lcweb.loc.gov/catdir/bibcontrol/delsey_paper.html>.
- Dempsey, Lorcan and Russell, Rosemary (1997). Clumps or... organised access to printed scholarly material: outcomes from the third MODELS Workshop. *Program*, vol. 31, no. 3, (July), pp. 239–249.
- Dempsey, Lorcan, Russell, Rosemary and Murray, Robin (1999). A Utopian place of criticism: brokering access to network information. *Journal of Documentation*, vol. 55, no. 1, (January), pp. 33–70.
- Dovey, Matthew (2000). So you want to build a union catalogue? *Ariadne*, no. 23, <<http://www.ariadne.ac.uk/issue23/dovey/>>.
- Dye, Juliet and Harrington, Jane (1999), Clumps in the real world: what do users need? *Ariadne*, no. 20. <<http://www.ariadne.ac.uk/issue20/clumps-workshop/>>.
- East, John W. (1999). Requiem for the National Bibliography? The implications of internet access to National Library catalogues. *AARL*, (March), pp. 1–10.
- Favret, Leo and McSean, Tony (eds.) (1983). *UK Library Database System and union catalogues: proceedings of a seminar organised by the Cataloguing and Indexing Group on 12 January 1983*. London: Library Association Publishing.
- Field, Clive D. (1999). CURL and resource description and discovery. *New Review of Academic Librarianship*. vol. 5, pp.125–142.
- Gillis, Helena (1999). Guiding the user to relevant information within a clump, with particular focus on an early version of the CAIRNS dynamic clumping service. *New Review Of Information And Library Research*, vol. 5, pp. 99–105.
- Gould, Sara (1999). The IFLA World Directory of National Union Catalogues. *Serials*, vol. 12, no. 1, (March), pp. 51–53.
- Gould, Sara (1999). From cards to clumps: a look at developments in the world of union catalogues. *Interlending & Document Supply*, vol. 27, no. 3, pp. 116–121.
- Hinnebusch Mark (1998). *Report to the CIC on the State of Z39.50 within the Consortium*. <<http://www.cic.uiuc.edu/cli/z39-50report.htm>>.
- Hirons, Jean and Schottlaender, Brian (1997). The CONSER/PCC evolution. *Serials Librarian*, vol. 32, no. 1/2, pp. 95–105.
- Interlibrary Loan Application Standards Maintenance Agency (1997). *An introduction to the ISO InterLibrary Loan Application Standards*. <<http://www.nlc-bnc.ca/iso/ill/standard.htm>>
- Joint Information Systems Committee (2001). *JISC Strategy 2001–05 executive summary*. <http://www.jisc.ac.uk/pub01/strat_01_05/exec.html>
- Lewis, Peter R. (1987). The future of the National Bibliography. *Library Association Record*, vol. 89, no. 10, pp. 516–520.
- LIBER (1992). *Library bibliographic networks in Europe: a LIBER directory*, second edition. The Hague: NBLC.
- Library and Information Commission (2000). *Empowering the learning community: report of the Education and Libraries Task Group to the Secretaries of State for Culture, Media & Sport and for Education & Employment*. <<http://www.lic.gov.uk/publications/policyreports/empower/index.html>>
- Library and Information Commission (2000). *Libraries: the essence of inclusion*. <<http://www.lic.gov.uk/publications/policyreports/inclusion.html>>
- Line, M. B. (1985). Planning union catalogues of books in developing countries. *IFLA Journal*, vol. 11, no. 1, pp. 27–35.
- Lunau, Carrol D. (1998). The need for an international Z39.50 profile for searching virtual catalogues <<http://www.nlc-bnc.ca/resource/vcuc/profiss.pdf>>.
- Lynch, Clifford A. (1996). *Defining and maintaining attribute sets for use with the Z39.50 protocol – a discussion paper*. <http://lcweb.loc.gov/z3950/agency/attrarch/archive/cliff4.txt>

- Lynch, Clifford A. (1997). Building the infrastructure of resource sharing: union catalogs, distributed search, and cross-database linkage. *Library Trends*, vol. 45, no. 3, (Winter), pp. 448–461.
- M25Link Project (1999). *Serials study report*.
http://www.m25lib.ac.uk/M25link/documentation/serials_report/serials_report6_issue1.html
- Madsen, Mona (2000). The National Bibliography in the future: new recommendations. *Alexandria*, vol. 12, no. 1, pp. 45–50.
- Miller, P. (1999). Z39.50 for all. *Ariadne*, no. 21 <<http://www.ariadne.ac.uk/issue21/z3950/>>.
- MODELS (1996). 3rd. Workshop: *National resource discovery: organising access to printed scholarly material* <http://www.ukoln.ac.uk/dlis/models/models3.html>.
- Moen, William E. (2000). *Resource discovery using Z39.50: promise and reality*. Presented at the LC Bicentennial Conference on Bibliographic Control for the New Millennium, November 15–17, 2000. <http://lcweb.loc.gov/catdir/bibcontrol/moen_paper.html>.
- Naylor, Bernard (1983). *Inter-library loans, union catalogues and the national database*. MARC-Users' Group Newsletter, no. 1, (February), pp. 20–29.
- Nicholson, Dennis (1999). Clumping towards a UK National Catalogue? *Ariadne*, no. 22, <http://www.ariadne.ac.uk/issue22/distributed/distukcat.html>.
- NISO (1995). *Information Retrieval (Z39.50); Application Service Definition and Protocol Specification, ANSI/NISO Z39.50–1995*.
- NISO (1999). NISO Circulation Interchange Protocol (NCIP) Standard: principles and guidelines for development <http://www.niso.org/drafts/Z3982v1.html>.
- Ober, John (1999). The California Digital Library. *D-Lib Magazine*, vol. 5, no. 3, (March) <<http://www.dlib.org/dlib/march99/03ober.html>>.
- OCLC, Pica begin discussions on collaboration (1999). *Advanced technology*, vol. 28, no. 5, (May), p. 1, 10–11.
- Powell, Andy; Heaney, Michael and Dempsey, Lorcan (2000). RSLP collection description. *D-Lib Magazine*, vol. 6, no. 9. <<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/september00/powell/09powell.html>>
- Ridley, Mick (1999). Practical clumping. *Ariadne*, no. 20. <<http://www.ariadne.ac.uk/issue20/bopac/>>.
- Russell, Rosemary, ed. (2000). *Making sense of standards and technologies for serials management: a guide to practice and future developments for librarians, publishers and system developers*. London: Library Association in association with UKOLN.
- Ryan, Deborah (1999). Unity Combined Regions database. *The electronic library*, vol. 17, no. 4, (August), pp. 219–225.
- Slack, Frances and Rowley, Jennifer (1999). Pathways to knowledge: a perspective on information and knowledge delivery in Australia. *Journal of Librarianship and Information Science*, vol. 31, no. 4, (December), pp. 197–203.
- Stubley, Peter (1999). Clumps as catalogues: virtual success or failure? *Ariadne*, no. 22, <<http://www.ariadne.ac.uk/issue22/distributed/distukcat2.html>>.
- Stubley, Peter (2000). What have the clumps ever done for us? *Ariadne*, no. 23, <<http://www.ariadne.ac.uk/issue23/stubley/>>.
- Turner, Fay (1995/1997). An overview of the Z39.50 Information Retrieval Standard <<http://www.nlc-bnc.ca/ifla/VI/5/op/udtop3.htm>>.
- Van de Sompel, Herbert; Hochstenbach, Patrick and Beit-Arie, Oren (2000). *OpenURL syntax description*. <<http://sfx1.exlibris-usa.com/openurl/openurl.html>>
- Walker, Jenny (2001). SFX – the context-sensitive linking system for libraries. *Serials*, vol. 14, no. 1, pp. 71–72.