

**CREATING A
PERMANENT DIGITAL
ARCHIVE OF LOCAL
MATERIALS**

A SLIC funded project

Project Manager: Don Martin

Researcher: Stephen Winch

© East Dunbartonshire Council &
Scottish Library & Information Council
2002

ISBN 904966 53 4

Available from

East Dunbartonshire Information & Lifelong Learning
The William Patrick Library, 2 West High Street, Kirkintilloch
G66 1AD

Contents

Introduction and Project Aims	4
Additional Requirements	6
Summary of Findings	8
Chapter 1. Continuing Development: Imaging	10
1a) Options for digitising images.	10
1b) Options for storage of images.	20
Chapter 2. Continuing Development: Metadata	22
2a) Options for the creation and handling of metadata.	22
2b) Using CORC to catalogue images.	29
Chapter 3. Continuing Development: Web Hosting	34
Chapter 4. Permanent Conservation.	36
Conclusion	41
Bibliography	43
Appendix I	50
Appendix II	51
Appendix III	52
Appendix IV	53
References	54
Acknowledgements	57
Glossary	58

Introduction and Project Aims

East Dunbartonshire Council is a medium sized local authority with a population around 110,000 situated to the north of Glasgow. The library service comprises eight community libraries, with its headquarters at the William Patrick Library in Kirkintilloch. The Reference Library has built up a large collection photographs (around 30,000), documentation and ephemera relating to the area and its history. Some of this material is available on open access, but a large proportion cannot be viewed because of the limitations of space and the fragility of the items.

East Dunbartonshire Libraries has been committed to the digitisation of its collections for a number of years, with a view to increasing public awareness of them and making them more widely available. They have participated in a number of digitisation schemes and with this project hope to develop a means of addressing the commitment in a sustainable way.

Project aims

The project seeks to establish viable routines and procedures for local “content creation” in small and medium-sized library authorities in Scotland. Guidelines are being developed by large institutions, but these seem inappropriate for smaller library authorities. The report looks beyond the nof-digitise program to ways of digitising content on a continual basis rather than the more limited short-term projects that are currently on offer. Two areas of policy are addressed within the report:

1. **Options for continuing development beyond nof-digitise.** Like many other library services East Dunbartonshire has had involvement in SCRAN and nof-digitise projects and/or proposals. However, these have been strictly subject-based, within very narrow limits, to meet SCRAN and nof criteria. It is seen as vital that a policy covering all areas of the local collection is developed. Options for future development must be addressed, to ensure achievement of the highest technical standards, whether by use of in-house digitisation facilities or external contractors.

2. **Permanent conservation.** Much of the digitisation in libraries at the present time is achieved without full consideration of future implications. With local studies collections, in particular, it is vital that long-term security of digitised images is achieved. Alternative strategies require to be compared, including the creation of conservation microfilm or photographic negatives (as surrogates) and the development of a programme of routine regeneration of digitised images.

Looking at these areas the project seeks to find best practice which is financially sustainable for libraries such as East Dunbartonshire.

The research project was carried out by a literature review, consultation with experts in the field and through the digitisation of 60 images as exemplars. These images represented a sample of the collections that East Dunbartonshire holds. Twenty-five collections were selected as being representative and two items from each collection were photographed (occasional extra items were included for items in more than one part). The items selected were chosen as representative of the different types of material held by the library, including photographs, pamphlets, minutes and other ephemera. The resulting negatives and colour slides were digitised onto Kodak Photo CD in accordance with current library procedures. They were then catalogued and placed onto the Internet. A researcher was employed to assist in the completion of the project.

It is not possible to deal with all aspects of the digitisation process within this report and therefore issues of content selection and the planning of a digitisation project are not addressed. Equally, this report does not provide a full description of the digitisation process and requires some previous knowledge of digitisation. There are many useful introductions to digitisation available, including Stuart Lee's *Digital imaging a practical handbook*¹ and Kenney & Rieger's *Moving theory into practice*² as well as the Getty's introduction to imaging website³. The Internet journals *Ariadne*⁴, *D-Lib magazine*⁵ and *RLG diginews*⁶ all provide up to date information on current digital projects and best practice.

Additional Requirements

In carrying out this research SLIC⁷ advised East Dunbartonshire to take on board the use of established library procedures for the creation and distribution of metadata, adherence to the UK government's eGif guidelines and best practice in the cataloguing and indexing of web content. SLIC is currently preparing guidelines in this area and it was considered wise for this project to try to reflect their likely requirements:

Use of established library procedures, including MARC, AACR2 and Z39.50 for the creation of metadata. MARC⁸ and AACR2 together form the standard mode of metadata entry in the library community. Increasingly libraries are using MARC to catalogue Internet resources which are then placed alongside traditional resources in library catalogues. The introduction of OPACs (Online Public Access Catalogues) and the 856 field into MARC, which allows a hyperlink to be placed into a catalogue entry, have facilitated the expansion of MARC into the digital realm.

In examining this option it was suggested that we seek advice from the Centre for Digital Library Research (CDLR) who along with SLIC are advocating the use of MARC in cataloguing digital collections. They believe that to create interoperability between collections a detailed record format is needed. These records should be completed to AACR2 standards and use recognised subject description schemes. When creating web resources the global nature of the Internet should be remembered, as opportunities to share information with other institutions and across domains will only be possible through co-ordinated methodologies for content creation.

Z39.50⁹ is a technical standard which allows computers to exchange common types of data. It has been used in the library community to allow cross searching of multiple library catalogues. In Scotland it has been used to create the CAIRNS service¹⁰, which allows users to search all University libraries' and the National Library of Scotland's catalogue in a single search. The majority of library servers are Z39.50 compatible and capable of joining such schemes. There have been some technical difficulties in bringing libraries into line, because of different specifications of the standard, but in 1999 the Bath Profile¹¹ was created to give an agreed interpretation of Z39.50, with a view to achieving more commonality within the library sector.

Ways of meeting the Electronic Government Interoperability Framework (eGIF). The eGif¹² is a mandatory requirement for all government departments including local government and public libraries. It requires all departments to create web content in a similar way. This is partly to enhance social inclusion by limiting the range of file types used by departments to those that are commonly supported by browsers, and partly to create interoperability between government departments by ensuring similar formats are used across the board. The eGif has implications for digitisation projects in the requirements for accepted image file types and metadata. Dublin Core is the chosen metadata format and this must be encoded in XML web pages.

XML is the successor to HTML and has a far greater functionality. HTML allows web pages to be structured by enabling basic formatting of documents, such as creating paragraphs, line breaks, text size and other basic functions. As the Web has developed users have wanted increased control over how they present information. XML allows this by enabling users to create their own tags. This has a great potential for creating schemes of tags for particular user groups. For example, there is an XML version of MARC records¹³. The advantage XML gives is that the record also carries information about how the tags should be interpreted; this information is carried by the DTD (Document Type Definition):

The DTD defines tags and their relationships and the structure of the document, thus allowing structural searching as well as enhanced content searching.¹⁴

This means that your computer does not need special software to interpret the XML file; the “key” is sent with the data. XML has a great potential for overcoming interoperability problems. At the present, however, although widely used commercially, it has yet to make a significant impact on the library world.

Summary of Findings

Chapter 1. Continuing Development: Imaging

- Continue out sourcing digitisation process.
- Scan at high resolution using HEDS guidelines as minimum standards.
- Photo CD adequately reproduces photographs, but not text documents with fine detail.
- Text documents should be scanned directly from the source document.
- Photo CD should be phased out in favour of scanning to TIFF files to meet eGif requirements.

Chapter 2. Continuing Development: Metadata

- Metadata should be prepared in MARC 21 format incorporating AACR2, meeting SLIC requirements.
- Library of Congress Subject Headings to be employed.
- CORC to be used to create metadata.
- MARC records to be exported into library catalogue.
- Dublin Core RDF records to be exported into Council website to meet eGIF requirements.

Chapter 3. Web Hosting

- JPEG files to be used for images on Internet, meeting eGif requirements
- Long term: develop XML database on council website to house images and metadata, to meet eGif requirements.
- Short term: create simple flat file HTML structure to house images until database is available.
- Utilise CORC's pathfinders to create link pages for image archive.

Chapter 4. Permanent Conservation.

- Refresh images every two years.
- Examine necessity for migration of images every two years.
- Maintain practice of creating 35mm negative archival surrogates.

- Create microfilm for documents that require preservation.

Chapter 1:

Continuing Development - Imaging

1a) Options for digitising images

In-house and third party digitisation. The previous digitisation projects carried out by East Dunbartonshire were done through an external vendor. This was on the advice of SCRAN, who advised external scanning as the most cost effective option. In approaching the present research the relative merits of using external vendors and in house scanning were re-examined.

The decision between in-house and external scanning is dependent on the type of materials that are to be digitised. The range of different types of document, their condition and their size, will affect the approach to digitisation. For example, capturing a 35mm negative requires a much higher resolution, being a reduction, than a full size document. Equally, the intended use of the digitised image will also effect how it is scanned. The level of detail required for an archival master will be different from an image that is to be viewed over the Internet or one that is to be printed.

The largest cost when digitising is likely to be staffing. Equipment needs to be manned and items have to be prepared before they are scanned. The Higher Education Digitisation Service (HEDS)¹⁵ suggest that preparation, such as project planning, preparing materials, moving materials, assigning identifiers, clearing copyright and checking the return of materials can account for up to 30% of a project's total cost. If digitisation is being carried out in-house, the training and staffing of equipment also becomes a cost factor. Alternatively, if digitisation is being carried out externally some preparation work, such as cleaning transparencies, may be offered by the bureau, although this is likely to be subject to additional charges.

HEDS¹⁶ split the Technical cost of digitisation into two elements:

- The cost of handling or otherwise moving the original material through the scanning process.
- The cost of writing an output file to the required resolution, bit depth and quality.

The cost of handling is dependent on what degree of automation can be built into the scanning process. Materials that are easier to batch process, such as A4 sheets or 35mm mounted slides, can be processed automatically with special scanning equipment, whereas a bound book or a glass slide will need to be manually positioned for scanning and will therefore incur greater labour costs. This price difference will affect both in house and bureau scanning, but is more significant for in-house scanning as the cost of scanners with high speed automatic feeders are significantly more expensive than more basic models. HEDS note that:

The drum scanners used commercially to produce up to 600 dpi black and white scans have scan speeds of less than a 3 seconds per sheet with autofeed for several hundred sheets and thus may cost over £25,000. Microfilm scanners start at £60,000 for basic automated versions with prices rapidly going through the £100, 000 mark for full automation and greyscale capability.¹⁷

So although prices of scanning equipment are continually reducing, prices for hardware that allow a fast output remain high. If a scanner is being purchased its expected lifespan will have to be taken into account. Many libraries have used bureaus because they are able to keep their equipment up to date and are likely to have hardware that has a far higher specification than most libraries can afford¹⁸. Large institutions such as the national libraries and the large universities around the world are buying scanning equipment. Often they are embarking on large scale scanning projects that can justify the costs involved. For smaller institutions the purchase of scanning equipment is harder to justify. The cost of training staff and having to man and maintain scanning equipment adds significantly to the cost of scanning.

In the case of East Dunbartonshire, there is a range of document types identified for capture. Although the large collection of 35mm negatives is one of the principal candidates for digitisation, there are also large amounts of documents and ephemera that merit consideration. For the project sampler a range of different types of material was selected. This was to assess the range of scanning solutions that the library might require. As scanning requirements change with document type it is essential to prepare items in batches that have similar characteristics. If a scanner has to be adjusted to take account of a different size of document or because of different levels of detail the scanning process will be more time consuming and therefore expensive. Equally, the process can be automated for large groups of similar objects such as 35mm slides or loose leaf A4 paper and this can lead to significant cost savings.

It was decided that the costs involved in purchasing scanning equipment were likely to be too high for an authority like East Dunbartonshire and so use of an external bureau would continue. This policy could change in the future if the cost of high quality scanning equipment and digital cameras continues to decrease. A 35mm scanner could be purchased for the bulk processing of photographic negatives, but this would still require other items being digitised in an alternative way or through the creation of surrogates. If the cost of high quality digital cameras continues to reduce it may become an option, as a portable medium it may provide a simple way of capturing a range of resources. The cost of training, staffing and maintaining the scanner would also increase the price of this option considerably.

Using an external bureau does carry a security risk as the items are being handed over to people outside of the library's direct control. There is a great potential for items to be lost or damaged in transit or during the scanning process. Some material may be considered too valuable to be entrusted to a third party. In such circumstances digitising from surrogates may be preferable. The credentials of the organisation carrying out third party scanning is important to assess and instructions should be given to them on how materials should be handled. In supplying materials to external companies it also important to create a mechanism for checking items out and in, to prevent disputes about loss or damage.

Scanning requirements. When deciding on scanning requirements decisions have to be made about what quality of image is required. As future requirements are uncertain Stuart Lee suggests that an item be digitised to the highest standard that can be

afforded, remembering that increased quality also increases file sizes and scanning times and therefore cost¹⁹.

Resolution levels, measured in dots per inch (dpi), are often discussed when benchmarking digitisation projects. As the resolution is increased, a greater level of detail from the document is captured (the higher the dpi, the more pixels from each inch of the source item are captured). Increasing resolution will also increase the time it takes to carry out a scan and decrease workflow and ultimately raise the cost. It will also increase the file size of an image, which can become very large. The measure of resolution is affected by the size of the original document, particularly if it is a format that is a reduction of the original such as 35mm negatives, slides or microfilm.

The following table shows the minimum scanning resolutions recommended by various institutions:

Table 1: Comparison of resolution specifications from various institutions

Institution	Document size	Resolution
Printed text		
California Digital Library, 2001 ²⁰	8.5" x 11" printed text	600dpi
U.S National Archives and Records Administration, 1988 ²¹	Text documents up to 11" x17"	300dpi
Colorado Digitization Project, 1999 ²²	Text	600dpi
Cornell: Digital Preservation Working Group, 2001 ²³	Printed text	600dpi
Digital Library Federation ²⁴	Printed texts	600dpi
35mm Negatives		
California Digital Library, 2001.	35mm negatives	2800dpi
JIDI: JISC Image Digitisation Initiative, 1998 ²⁵	35mm negatives	2100dpi
Cornell: Digital Preservation Working Group, 2001	Transparencies	4000-5000 dpi
Photographs		
California Digital Library, 2001.	8.5" x 11" Photographs	600dpi
U.S National Archives, 1988	Photographs	3,000 pixels across longest dimension
Colorado Digitization Project, 1999	Photographs	3000-5000 longest dimension
JIDI: JISC Image Digitisation Initiative, 1998	Photographs	300dpi
Cornell: Digital Preservation Working Group, 2001	Photographs	400dpi
Digital Library Federation, 2001.	Illustrated texts	400dpi

The table shows a range of recommended archival scanning benchmark figures. Printed texts are mostly benchmarked at 600dpi (except for NARA), negatives are between 2100dpi to 2800dpi, while photographs range from 300dpi to 600dpi. Given the range in figures it is difficult to draw firm conclusions.

NOF technical guidelines²⁶ do not provide any benchmark standards for scanning resolution level. They recommend examining guidance from other institutions including the following benchmark recommendations:

Table 2: HEDS guidelines for image capture

Institution	Document type/size	Resolution	Bit-depth
HEDs (Higher Education Digitisation Service) ²⁷	Photographic prints (colour)	300	24
	Photographic Prints (monochrome)	300	8 greyscale
	Printed material (colour)	300	24
	Printed material (monochrome)	300	8 greyscale
	35mm colour slide/negative	2400	24
	35mm monochrome slide/negative	2400	8 greyscale

The HEDs minimum requirements are lower than those expressed by the institutions listed above. Anne Kenney in an article for the Joint RLG and NPO Preservation Conference²⁸ argues that institutions should look for guidance rather than guidelines. As each project will require different standards depending on the size of materials and the amount of information that is required to be captured. Equally the type of scanning equipment used will affect the quality of the scan. Kenny and Reiger²⁹ suggest benchmarking as a way of finding the correct level capture for a digitisation project. Here, the items to be digitised are assessed for the smallest level of detail that needs to be captured. Then test scanning is carried out to find appropriate settings to reproduce the digitised documents as required. Benchmarking is more easily carried

out if scanning is carried out internally and the equipment is available to run tests. For institutions using external digitisation vendors, there is a certain amount of trust given to those vendors, who having expertise in that area should be able to provide advice on the level of scanning required. Despite this an institution can state the level of detail that needs to be reproduced and what the image is likely to be used for.

The HEDs guidelines can be taken as minimum requirement, but this has to remain flexible to the needs of the particular collections being digitised. Cost may also prove to be a factor as if the desired scanning resolution costs more than the institution can afford decisions on scanning quality may have to be compromised.

As resolution offered by bureaus is often high, a standard file size is often used to dictate the level of resolution scanned at. Once a level is selected the resolution is adjusted to produce a file of the requisite size. Again the size of a document will affect this with a lower resolution being used for larger documents. This highlights the importance of grouping materials into batches of similar items, which can have their own requirements. The file size will also be affected by the file format used and the type of compression algorithm used by the scanning software.

It is also useful to consider what the images are going to be used for when assessing the resolution needed. An archival image will need to be captured at a higher level than an access image. For access images thought needs to be given to how the items are to be displayed. Although, at present, monitors commonly display at 72 dpi, higher resolution monitors are becoming available, which may show up deficiencies that are not apparent on current hardware. If an image is to be used to produce high quality prints it will need to have a resolution good enough to meet the resolution level of printers. In previous years the printers that were generally available scanned at 300dpi, today 600dpi printers are commonplace³⁰. This highlights the fact that scanning requirements change over time. For this reason, it is sensible to create a master image that is beyond current needs to prevent the need to recapture data.

Given that the images that we are scanning are not principally archive masters, they need not be scanned at the higher benchmark figures. Despite this the price of scanning within these ranges is not significantly altered, so there is justification for scanning at higher resolutions to increase the longevity of the images. In approaching the Scottish Newspapers Microfilming Unit (SNMU)³¹ for scanning advice, we were staff would select an appropriate resolution based on the level of detail to be captured,

based on readings from the digitisation equipment. This was a result of the fact that the collection format varied greatly and so could not be uniformly digitised.

As resolution offered by bureaus can be high, a standard file size is often used to dictate the level of resolution scanned at. Once a level is selected the resolution is adjusted to produce a file of the requisite size. Again the size of a document will affect this, with a lower resolution being used for larger documents. This highlights the importance of grouping materials into batches of similar items. The file size will also be affected by the file format used and the type of compression algorithm used by the scanning software.

The resolutions selected provide an archive quality scan. This is a high level of capture that attempts to reproduce all detail in a document, including the texture and type of material that a document or image is printed or written upon. Scanning at high levels is also a means of future proofing digitised images. Although, at present, monitors commonly display at 72 dpi, higher resolution monitors are becoming available, which may show up deficiencies that are not apparent on current hardware. If an image is to be used to produce high quality prints it will need to have a resolution similar to a photograph. As technology improves libraries may require to produce still higher resolutions, for example the recent introduction of digital television has given viewers a higher resolution television picture that is certain to become a new standard.

For East Dunbartonshire the choice of resolution level and other scanning standards is related to the intended use of the images and the preservation strategy in use. When a new photograph is added to the libraries collection a 35mm copy negative is taken. This is kept in archival conditions and is considered the archival copy (see Chapter 4). As the digital image is not the primary archival master, there is less need for it to be a completely accurate copy of the original as is the case with many university based projects. Instead, the reproduction of the major content of the resource is important rather than to capture finest detail possible. To reproduce to the finest detail requires an increased control over the scanning process and is also likely to increase the amount of extraneous information captured, such as the grain of the paper a document is written on. A large file size will be required.

The bit-depth of a scan also affects the quality of digital images. Bit-depth can be described as the number of colours or shades available in a computer's palette when recreating an image during the scanning process. A 1-bit image has two possible

colours: black or white; a picture reproduced as a 1-bit image will be two tone. An 8-bit image is made up of 256 shades of grey or colours and a 24-bit image contains over 16 million combinations of colour and shade³².

The bit-depth has implication for the level of colour recorded in an image and consequently file-size. Scanning a two-tone line drawing as a 24-bit image is clearly capturing detail that is unnecessary, recording colours and shades reflected from the page. To prevent recording unnecessary detail the bit-depth should match the type of material being scanned:

Printed texts: 1-bit

Black white photographs: 8-bit

Colour photographs: 24-bit

Scanning with Photo CD. East Dunbartonshire's previous experience with the scanning process has been through the use of Kodak Photo CD master discs (PCD). As well as being a file format PCD is also Kodak's own scanning process, which provides scans at five different resolutions, the maximum being 2048x3072 for 35mm film (2200dpi). Slides or transparencies are scanned using a Kodak PCD Imaging Workstation (PIW)³³ with Kodak's own colour management and compression techniques. The 50 sample images for the project were photographed by project staff and by a professional photographer who is employed by the local authority. The resulting 35mm negatives and colour slides were scanned onto Photo CD at a local digitisation bureau.

In 1998 the Department of Preservation and Conservation at Cornell University produced a report assessing Kodak Photo CD usefulness for libraries and archives³⁴. They found that although PCD technology gave good results for many types of material those items that had fine detail such as text were rated as unacceptable by evaluators. This was particularly the case for large documents that had fine detail. Photographs, halftones and handwritten manuscripts were all judged to be acceptable surrogates for the original, even if there was a noticeable shift in colour representation. The research highlights the fact that slight colour shifts were not considered detrimental in providing a usable surrogate for an original item. The results are a reflection of the fact that the PCD format was developed as a way of

reproducing consumer photography. As a result large items with fine detail fall beyond the system's original design.

The research from Cornell confirms the results gained from East Dunbartonshire's sample images. The photographs produced good quality digital surrogates. Typescript and hand written documents were also reproduced satisfactorily, except where the type was particularly small or where there was relatively small detail in a item that was larger than A4 when photographed. As with Cornell's research there was some change in the reproduction of colours, but this was slight and was not thought to be a significant problem for images that were intended primarily to provide access over the Internet. This was also mitigated by the fact that if a copy of photograph was requested the library would always provide a copy print from the original negative (rather than a digital offprint).

One of the problems of using the PCD format is that it requires a surrogate to be made of the original object, unless transparencies already exist. Each time a document is reproduced in whatever form there is a loss of quality:

The tonal characteristics of the image file were necessarily transformed at three different points in the workflow: when the film was exposed, when the film was scanned, and when the viewing file was derived from the Photo CD master file.³⁵

To minimise this, the best possible photography should be used and therefore it is sensible to rely on a professional photographer. With our samples, some items were photographed by a professional while others were taken by the research team. Although reasonable results were obtained by research staff, a far greater consistency and clarity were obtained by the professional photographer.

From our research we have concluded that the Photo CD scanning method does provide a cost effective means of digitisation. However, items with fine detail, particularly those above an A4 size are not accurately reproduced and therefore material needs to be carefully selected before scanning. An advantage of this method is that in addition to the digital version it provides a 35mm negative surrogate that can be archived.

1b) Options for storage of images

A decision also has to be made regarding the file format in which the digital image will be saved. When images are saved they are compressed to reduce file space. The different file formats employ different file compression algorithms, many of which remove extraneous information from the image. Those that do this are referred to as lossy, as the information that is removed cannot be recovered once the process is complete. Libraries and archives creating archival masters, that are expected to outlast the original, do not recommend using lossy compression methods. Again the purpose of the digital image is important in deciding which file format is used. Commonly a high quality digital master is created for archiving and storage purposes. From this lower quality access copies can be derived for daily use whether on CD-ROM or through the Internet, where a large file size gives long loading times. The Internet is becoming faster through an increased take up of broadband connections, but for the majority this is still not available. Keeping a high quality digital master allows the option of upgrading Internet versions of images as and when faster connections become more widespread. The same is also true for printers, which were generally available at 300dpi whereas now 600dpi printers are commonplace³⁶.

TIFF files. The TIFF format is probably the most widely used file format in digital preservation. It uses lossless compression and is therefore ideal for archival use, providing an accurate representation of the information scanned. It also allows high quality images (up to 24bit) and conversion to other formats is straightforward. The format is non-proprietary and has commonly been used for creating a master image for archiving and to create further compressed derivatives. The e-Gif framework recommends this format for creating images that will not tolerate information loss.

JPEG files. Along with GIF files, JPEGs are the most common image file formats used on the web. They are lossy in their compression but the resultant small file size makes them ideal for transmitting images over the Internet. JPEGs have 24bit colour depth allowing for high quality image reproduction and are a non-proprietary

standard. The e-Gif framework recommends JPEGs as an appropriate file format for the delivery of image files over the Internet.

GIF files. Like JPEGs, GIF files are a common file format used on the web. The format allows for an 8bit image and consequently is used for line art and greyscale images. The compression used is lossy and is recommended by the e-Gif framework.

Kodak Photo CD. Photo CDs are provided with images saved in Kodak's own file format, PCD. The format uses what Kodak calls "visually lossless" compression, meaning that although some information is lost through compression this is not discernable to the naked eye. So although not strictly a lossless file format it does provide better results than most other lossy methods of file saving.

The fact that Photo CD is a proprietary format could create problems in the future. As it is not an open standard, fewer commercial image-editing packages support the format than with the open standards. There is also a danger that Kodak could alter the nature of the format to improve compression or other features whenever it wishes. This could leave institutions with differing versions of format, which may not be supported by all software. The PCD format is also not browser compatible, so for web delivery it often has to be converted to a more commonly used exchange format, such as TIFF or JPEG. The majority of digitisation projects worldwide use the TIFF file format to store digital masters³⁷. There are two reasons for this. Firstly, it is an independent file format that is not linked to any particular company. This means that it is more widely supported in software and is likely to have a longer lifespan than a proprietary standard that can be modified or replaced at any time. Secondly, the TIFF format allows lossless compression, so when an image is opened all the information that was captured during the digitisation is retained. This is not the case with the majority of image formats, including Photo CD, JPEG and GIF. So although Photo CD provides reasonable quality for digitisation of photographs, it does not deliver true archival quality.

East Dunbartonshire currently uses Photo CD. Although this fits in well with current local practice it is recommended that a move to using TIFF format is made. This will bring the service into line with the eGif and provide a better quality master file with a longer lifespan than Kodak Photo CD.

Chapter 2:

Continuing Development: Metadata

2a) Options for the creation and handling of metadata

Creating metadata for a digitisation project can be one of the most time consuming parts of the project but is vital to facilitate proper access to the images. One of the problems encountered by this project was the sheer number of ways used to create metadata and the lack of standardisation throughout the library community. In selecting a method for creating metadata we were advised by SLIC to liaise with The Centre for Digital Library Research³⁸ and to examine the use of established library procedures for creating metadata, namely, MARC and AACR2. On top of this, as a local authority East Dunbartonshire is required to follow the recently produced eGif guidelines, which state how government related Internet material should be catalogued, by making use of Dublin Core.

Requirements for metadata. It is useful to begin by stating what was wanted from the metadata that we were creating:

- To allow a high specificity of retrieval.
- To provide adequate description of our resources.
- To establish a standard that would have longevity.
- To allow interoperability with other systems.

Interoperability. The term interoperability is often used when discussing digital libraries. Paul Miller in an article in *Ariadne* states that to be interoperable:

One should actively be engaged in the ongoing process of ensuring that the systems, procedures and culture of an organisation are managed in

such a way as to maximise opportunities for exchange and re-use of information, whether internally or externally³⁹.

This definition encompasses more than strict technical interoperability, i.e. allowing computer systems to exchange information with each other, but also encompasses semantic and inter-community interoperability. To maximise the value of a digital collection or Internet resource it needs to be accessible to as many people and in the most straightforward way possible. As the Internet has grown collections of various types of material have developed in various institutions using different hardware, software, cataloguing and semantics. Although many of these collections would compliment each other, there are often large problems in trying to link them together. Often this is because of how they have been catalogued and the semantics and thesaurus used to describe the resources. The Internet provides a great potential for these resources together, but without common standards, consistently applied extra expense and often duplication of work is required to bring the various standards in line with one another.

Scotland provides a good example of a community where the sharing of information resources could become greater than the sum of its parts. As a fairly small country there is an opportunity to link together historical and cultural resources relating to Scotland from a range of institutions. This has already started to take shape through the CAIRNS⁴⁰ project (the Co-operative Academic Information Retrieval Network for Scotland), which has linked all SCURL (Scottish Confederation of University Research Libraries) members' library catalogues together, allowing them to be searched from a single web interface. Making use of Z39.50 servers the service can search MARC records held by the individual institutions. There is a desire to expand this further into the public library community; although the fact that the majority of public libraries do not use the MARC record format may hamper this move, it should be noted that the CAIRNS requirement is for libraries to output, rather than store, metadata in MARC format, and that Aberdeen University is a fully functioning member of CAIRNS even though it is not a MARC site. East Dunbartonshire, Edinburgh City Libraries and Glasgow are likely to join the scheme in the near future.

Another project that has grown out of the experience with CAIRNS is the SCONE⁴¹ service. SCONE is a database of special collections and where they are held. Users can limit searches by area or by topic, find where in Scotland collections

are held and then link to online resources or catalogues. The service is still in its early stages, but as more information about libraries and archives collections are added it will become a powerful research tool.

These projects are very much driven by SLIC and the Centre for Digital Library Research (CDLR). Dennis Nicholson, the director of the CDLR, sees these projects as a stepping-stone towards a National Electronic Library for Scotland that would:

Enable Scottish citizens (and others) to discover any significant educational, research, public information, or recreational resource anywhere in the country, regardless of source (libraries, museums, archives, electronic services, or elsewhere) or format (electronic or hardcopy – it is the process of discovery that is necessarily electronic)... The assumption behind the vision is that user requirements do not fit neatly within the boundaries of a sector (eg HE libraries, or FE libraries or public libraries) or domain (eg libraries or museums or archives)⁴².

Moves within Scotland towards this idea of a National Electronic Library have, up until now, focussed mainly on the large University libraries, the FE sector and the National Library. Despite this there is little co-ordination in creating digital content and still less in the broader community of archives and museums. Public libraries need to be aware of trends in the wider library community, that could help enhance their services and bring their collections to a wider audience. It is for this reason that East Dunbartonshire in planning a digitisation strategy is looking to be interoperable with other institutions within Scotland. The methods used will also need to be manageable for a small Library service like East Dunbartonshire.

Dublin Core. Dublin Core⁴³ has become an increasingly popular format for recording metadata. It consists of a set of 15 elements which can be use to describe electronic documents. The elements are:

Title	Publisher
Creator	Contributor
Subject	Date
Description	Type

Format	Relation
Identifier	Coverage
Source	Rights
Language	

As well as the 15 base elements, there is an additional set of qualifiers that can be added as sub-elements to provide more specific entry points. The format is also user extensible, in that local sub-categories can be added to any of the elements. So, for example, if you wished to distinguish between an author and a programmer, you could create the elements “creator.author” or “creator.programmer”. The design is interoperable as a computer program that recognises Dublin Core would not need to know locally created elements such as “creator.programmer”, but would just read the parent level of “creator”. Although this gives flexibility to extend Dublin Core for local circumstances, any extensions need to remain within the meaning of the source element so that when it is truncated it will still have meaning. This is the limiting factor of Dublin Core in that having so few entry points many different types of data are lumped together. Even though this can be overcome to an extent with qualifiers, common schemes of qualifiers have to be used to create full interoperability. As a result many digitisation projects (such as SCRAN), which use the scheme, also add extra fields (in addition to the fifteen) to allow a greater specificity of metadata, but there is no agreed standard for doing this in a consistent or interoperable manner.

Dublin Core is simply a range of elements which are often recorded in HTML and imbedded into a web resource. The format can also be used in various different encodings such as with XML or any other means of saving data.

Devon Libraries. Devon Public Libraries have digitised a range of local history photographs using the Dublin Core based metadata standard. They have used the extensibility of the format to create what they call “Devon Core”⁴⁴ which remains within guidelines of Dublin Core, in that any extension to an element can be reduced to its parent element. In completing the records a local history subject thesaurus is being developed, principally to provide a hierarchy of search terms. Being on a smaller scale it is easier for the library to maintain a consistency in the application of subject terminology. The Devon system provides a good example of an application of Dublin Core, which also makes use of a formal subject scheme. However, although

there is potential for linking these resources with other Dublin Core based systems (the records are recorded in HTML), the difference in subject classification schemes used in other projects is likely to cause problems in a large scale database.

SCRAN. The Scottish Cultural Resource Access Network (SCRAN)⁴⁵ was a Millennium Commission funded project that provided East Dunbartonshire's first experience of digitisation. Metadata for the project was created from a Microsoft Access template that was supplied to participating institutions. The format of metadata was Dublin Core-based, with several fields added for extra specificity. The service is cross-domain, combining resources from museums, archives and libraries. Although the scheme was successful in bringing together images from a range of communities the metadata presents some problems. Firstly, it was difficult to enforce a subject classification scheme because of the range of materials involved in the collection. Secondly, although there was some modification to records being submitted it was difficult to create a uniform method for completing the records, as each institution filled out their own records, often with little experience of cataloguing. This has meant that as the number of records has increased in the database it has become harder to track down the items that can be classified under a number of different subject terms.

The SCRAN scheme is also a closed scheme in that an Internet search engine would not be able to find items contained on SCRAN; instead the SCRAN interface has to be used. As the scheme is subscription based users can only enjoy full access to the collections through a library or other subscribing institution. For East Dunbartonshire to make its SCRAN collections available to a wider audience we would have to host the images at an additional location. If there was no subscription and the images had a static URL there would be a possibility of creating our own catalogue records that link directly to the SCRAN resource. This is also a problem in linking other schemes such as CAIRNS to SCRAN. The records in SCRAN are converted into index+ software and saved in GRS-1, which is a basic data-saving format. For it to link to other systems this syntax has to be readable by any such systems.

Other Scottish projects. Glasgow City Libraries have two digital image collections available: the Virtual Mitchell⁴⁶ and the Lipton Archive⁴⁷. The images and metadata

are managed on a database from Ibase⁴⁸. The metadata was migrated from an early format and so is not Dublin Core compatible. This means that it will be difficult to integrate this system into any other wider scheme. Aberdeen University, Dundee University and St Andrews University are involved in a collective digitisation project funded by the Research Support Libraries Group⁴⁹. They are linking together their digitised collections so that they will all be searchable from a single website. The project uses Talis/Ibase's inVisage software⁵⁰, which includes Dublin Core as the default format, with optionally configurable local fields added by Ibase. The collections will be searchable on common Dublin Core fields. They use a local subject scheme, which is a modified form of the Library of Congress Thesaurus for Graphic Materials. Other libraries could potentially link up with this scheme by using a Dublin Core-based system. To create proper interoperability the same subject classification scheme would also have to be used.

Although there is some uniformity in the use of Dublin Core there is no consistency in the application of cataloguing rules and subject thesauri. Although these and other systems could be joined together the disparity in resource description will hinder access to the collections.

MARC 21. MARC records have been used in University libraries around the world since the early eighties. MARC 21⁵¹ is the latest incarnation of the standard and is an amalgamation of USMARC, UKMARC and AUSMARC. The use of MARC has facilitated record sharing around the world. To improve the interoperability the various versions of MARC have been consolidated into MARC 21, which is based on USMARC. The MARC format is probably the most complex format for metadata, allowing for great specificity in its construction. Its use is not so common in the public library sector, a fact often attributed to the complexity of the format.

MARC has been used to describe electronic resources, but its use for digitisation projects has not hitherto been as widespread as Dublin Core. Increasingly, however, the MARC format is being used to catalogue web resources such as online journals and websites. The advantage of this is that these resources can be placed alongside traditional library resources in the library catalogue.

Importantly, MARC records are filled out to AACR2 standards, which provide a control for how data is entered into the record. This is an important advantage in creating interoperability. In fact the primary problem with large-scale databases is not

the format that is used, but rather the differences in how the records are completed. This was well illustrated in the findings from the CAIRNS project⁵². Technical problems with the linking of libraries through Z39.50 occurred, but they were more easily fixed than difference in the ways records were filled out. Although most member institutions were using UKMARC there were considerable differences in cataloguing practices. This was apparent in the application of AACR2, differences in the use of authorities for names, and in the variety of subject classification schemes in use. Because of these problems CAIRNS proposed the HILT project⁵³ that is looking at possibilities for mapping together the major subject classification schemes (DDC, LCSH, UNESCO, AAT and possible UDC)⁵⁴. This could provide a way of further improving subject interoperability within CAIRNS and other projects.

Given the range of problems found by CAIRNS it can be suggested that it is too difficult to create a uniform means of cataloguing within Scotland. However, the benefits from creating a unified approach are large. The SCONE model may provide a more feasible way of linking resources by treating them at collection level rather than at an item level as in CAIRNS. It is to all libraries' advantage to remain as flexible as possible in creating metadata so that opportunities for exchanging data are kept open. It is also important because of the changing nature of metadata. As new formats become available it is essential that we think of our ability to migrate records to new formats to avoid the possibility of our records becoming redundant. To this end more restrictions on how metadata is to be prepared are likely to be dictated by the library community and other agencies.

Using MARC 21 in East Dunbartonshire. Using MARC has several advantages for East Dunbartonshire. It will allow records for digital resources to be contained in the main library catalogue. The specificity of the record format should also allow complex searching of the resources that are held. Using MARC will also allow records to from part of external schemes such as CAIRNS.

The use of MARC is also seen as having a safe future. It is the most common record format in the library community and as such is likely to supported by software and other services. There is currently a great deal of discussion about the potential for XML in delivery of the next generation of library catalogues. As most records are held in MARC it should be a straightforward process to convert to an approved XML

version once this becomes available. This might not be so straightforward if we took a Dublin Core format and extended it to fit our circumstances.

Cataloguing in MARC is seen as a challenge for East Dunbartonshire as currently the library uses a Dynix system that does not employ MARC records and so there is no local experience of MARC cataloguing. As a member of CAIRNS, however, the library is committed to converting to the format in the near future. This will allow other institutions to search the East Dunbartonshire catalogue and also enable the library to take advantage of record sharing through OCLC or the British Library. However, to add records created by the present project additional indexes will be required and will need to be purchased from Dynix.

At the outset of the project we were recommended to follow established library practice in cataloguing, making use of MARC and AACR2, and to follow the eGif guidelines. The eGif requires that metadata should be created in Dublin Core and be presented in XML web pages. Given that the two methods were contradictory we looked for a way of accommodating both sets of guidelines.

2b) Using CORC to catalogue images

CORC. The Co-operative Online Resource Cataloging (CORC)⁵⁵ service is provided by OCLC and is an online cataloguing service for Internet resources. Cataloguing is done online and all records entered are added to OCLC's Worldcat. From there they can be downloaded by any member institution and added to local collections. The Worldcat is going to be made open to the public in the near future and will allow users to search all the records for online Web resources catalogued by member institutions.

Cataloguing can be done in either MARC21 or Dublin Core and once a record is entered it can be converted between the two formats. The service also provides access to the Library of Congress authority files, allowing control over personal, corporate and geographical names. CORC is a subscription service with participants earning credits for records that they add to the database. This is of particular interest to libraries adding digital collections as their records will not have been catalogued before. SLIC have purchased a trial consortia subscription to CORC which they are

offering to smaller libraries such as East Dunbartonshire to allow the cataloguing of Internet resources.

Using CORC in East Dunbartonshire. CORC was selected for creating the metadata for the project because it allowed the creation of MARC records as well as Dublin Core. We decided to create the main records in MARC because it would provide us with high quality, detailed records, which could be converted to other formats. Although records can be switched between the two formats better results are gained from going from a complex format to a simpler one. A conversion from a MARC record to a Dublin Core record will always be accurate, but this is not necessarily so the other way round. For example, MARC distinguishes between authors, editors and other types of creator. In Dublin Core the creator field can have many interpretations, but it will be converted directly into a author field when transferred to MARC.

With the assistance of the CDLR a template was drawn up for a MARC record that could accommodate the detail we needed to record. A sample record is included in Appendix I; although it looks large much of the data is repetitive and is kept on a template in CORC known as “constant data”. Appendix II shows the constant data template used for creating the records. As well as including repetitive data such as the library’s address and the availability of the resource, each part of the record that is to be filled in is described and highlighted with a “*” symbol. When a record is to be completed the highlighted parts of the record are replaced with detail about the resource. This keeps the form simple to fill out and ensures consistency, accuracy and completeness of the data. The only part of the form which could not be dealt with in this way was the subject fields as different MARC tags will be needed depending on the subject matter of the resource. The format enables someone with little knowledge of MARC to be able to complete the form, as each field is described within the template. Subject categorisation does require knowledge of MARC and knowledge of using non-filing characters for the title is also needed. However the format can be reduced to a simple table (Appendix III), which library staff can use to complete the record. This can then be cut and pasted into CORC.

The CORC service provides an easy-to-use web interface for cataloguing. When a website is catalogued the screen is split into frames to enable the resource to be displayed while the item is catalogued. The service also provides assistance with the

understanding of MARC fields. Each of the MARC fields represented on the record is also a hyperlink that takes the cataloguer to a description of the field and contains OCLC CORC guidelines for completing the entry. This is a useful feature, particularly for those who are new to cataloguing Internet resources. There is also access to Library of Congress authority files, which at present have few Scottish place names entered. The CDLR have set up an informal CORC users group for Scotland and it is suggested that the CORC service could be used as a way of developing authority control for place names. CORC allows authorised users to submit terms for authority approval; for Scottish placenames and other headings the authorised user is the National Library of Scotland. As the CORC user base continues to grow in Scotland libraries could use this facility to submit terms as way of developing an approved list of Scottish place names. This would make cataloguing more straightforward as well as enabling greater cross searching of databases.

Design of MARC records. The MARC records are designed to provide a comprehensive description of the resource and allow easy retrieval. To achieve this we imposed various guidelines on how certain parts of the record should be completed. The Title field involves a composite title made up of the following elements:

Main subject/the medium or aspect of the resource/description of the resource/date.

e.g.

“Petticoat Dan, full length portrait of Dan holding shovel, while resting on wheelbarrow”

or

“Bearsden and Milngavie District council, first page of minutes from the first meeting of the District council, 16th May 1974”

Although this method does not always provide a flowing title, it keeps the main subject to the beginning for sorting purposes and gives a brief description of the entire resource. Not all items fit neatly into this categorisation, so it has to be taken as a guide rather than an absolute. One of the reasons for giving this information in the

title is that the MARC record represents the digital image rather than the original item and therefore the date of publication will be the date the image was placed onto the Internet. Equally, the medium of the image is a JPEG file rather than its original context. By including this information in the title it also provides a miniature description of the whole item, which when appearing in a brief description or title list should provide all the information a potential user will need in deciding whether they wish to view the resource.

In the SCRAN scheme the metadata contained several paragraphs of information about the background to the resource. This was a feature of SCRAN that East Dunbartonshire appreciated and wanted to continue as part of in the current project. In MARC the summary field was used to provide a biographical and/or historical context for the images. It was considered that this added value to the resource. The downside was that it was very time consuming, so the level of detail included for the sample images is unlikely to be sustainable in an ongoing process of digitisation.

Subject classification has proved difficult to achieve. East Dunbartonshire does not use an official subject scheme and although a provisional decision was made to make use of Library of Congress subject headings this has yet to be achieved. The library possesses its own subject classification system that was developed to catalogue the local history photographs. This was not felt to be sufficient for the current project and so other systems were examined. The need to use a subject classification scheme, to facilitate interoperability, led the project to examine the established library subject classification schemes. This has not been the trend in digitisation projects, with projects such as the St. Andrews University, SCRAN and the Mitchell using home-grown schemes. The majority of schemes being carried out tend to look for a scheme that will fit the type of material that is being catalogued or devise their own because the subject matter of a collection is particularly narrow. From a local history point of view a dominant local history thesaurus could be created to serve the whole of Scotland, but there is not one available at present. The main disadvantage of developing a local thesaurus is that once it is placed on the Internet the terms used may be limited in a global context. For example in the East Dunbartonshire photographic subject headings the term "Canal" is used to represent the Forth & Clyde Canal which runs through the district. This term is sufficient for local use in the library, as there is only one canal in the area, but the term is meaningless in a database that is available worldwide. By using LCSH subject terminology we will have more

general subject terms that will be more appropriate for the global reach that the Internet provides.

CORC metadata solution. To meet both eGif and SLIC conditions it was decided to have two sets of metadata. A MARC version of the record would be downloaded into the library catalogue, giving direct access to the images through the library OPACS. To meet eGif standards a Dublin Core RDF version of the MARC record would be downloaded into a specially constructed XML database on the council website. The files on the website would also incorporate the images into the XML pages. This would provide means of accessing the images and would meet both criteria for metadata construction.

Chapter 3:

Continuing Development: Web Hosting

File preparation

Before the images could be placed onto the web they had to be edited and converted to JPEGs. Using Jasc's Paint Shop Pro package images were resized, cropped and saved as JPEGs. As some images lacked clarity from the scanning process the brightness was altered and sharpening filters were applied. This was mainly for the text-based documents where fine text had become slightly hazy. In saving the images the compression was adjusted to produce a file size below 100k. This was to give an image that would not take too long to download over a standard Internet connection. Unfortunately some of the text images containing fine detail could not be reduced below 100k to retain legibility. Although some of the results were not as clear as we would have wished, all provided an access copy that was legible.

Each photograph in the East Dunbartonshire collection has a unique negative which is given a sequential number based on when it was acquired. As this provided a unique identifier and because it could be traced back to the original negative, each digital image was given the negative number as its filename. So a photograph with the negative number "N2757" would be saved as "n2757.jpg". To catalogue the images on CORC the images needed to be online before they could be catalogued. The East Dunbartonshire website was not able to house the images during the project and so SLIC allowed us to place the images temporarily onto the SLAINTE website. When they are moved to the local website it is intended that they will have a file extension that is easily extendable. The project has suggested that the images are grouped into batches according to when they were digitised, so a file extension could be rendered in the following way: www.edunbarton.gov/library/batch1/n2757. This will enable the files from each batch to be kept together, but is also a scheme that is extensible.

Image and metadata hosting

The images are currently housed as raw JPEGs on the SLAINTE website. When they are moved to the East Dunbartonshire website they will be placed in a HTML page which will provide a corporate look to the collection and allow a title to be displayed with the images. This change will also necessitate a minor change in the cataloguing entries because they are no longer JPEG files but HTML. This is a temporary solution as the intention is to develop a XML database in line with the eGif guidelines.

It is proposed that further external funding be sought to create a database. In looking at image databases the commercial companies such as Ibase⁵⁶ and House of Images⁵⁷ will need to be examined. As metadata can be created through CORC the full suite of tools used by these packages may not be needed. It is possible to speculate that as libraries are increasingly creating web content the producers of library cataloguing systems may start to include image management modules to their products. East Dunbartonshire will be reassessing its library catalogue system in the next five years and an image database may be available as an add-on to existing modules. Talis have produced the inVisage system in conjunction with Ibase and it seems likely that other manufacturers will follow suit.

Any database solution will have to fit in with platforms already in use within East Dunbartonshire. In the near future Lotus Notes software is to be used to create and maintain databases on the council website. Any solution that is found will need to be compatible with this software.

CORC Pathfinders

The Council's ICT department has given the project a costing for a database, which is out with the scope of currently available budgets and has been put on hold. In the meantime it has been agreed that a simple HTML structure could be constructed without incurring significant cost. To achieve this purpose another feature of CORC has been utilised to create a flat file database of images. CORC can create "pathfinders" which are simple HTML link pages created from records held on the CORC database. They are constructed either by selecting individual resources to

appear on a link page or by creating a dynamic search. A dynamic search involves creating a Boolean query, the results of which will appear on a links page. For each resource found on a search CORC will display the title as a hyperlink, provide the web address and the summary field from the catalogue record (an example of a pathfinder is shown in Appendix IV). Each of the items that East Dunbartonshire has digitised belongs to a collection, with two items digitised from each of twenty-five collections. For each collection a pathfinder was created. These were generated from a dynamic search that accessed the title-added entry field (740 field), containing the collection title. This meant that when a collection title was inserted into a dynamic search of CORC it would produce a links page for the items within the collection. The advantage of this is that when new images belonging to existing collections are added to the CORC database an updated pathfinder can be created by simply rerunning the query.

Each of the twenty-five pathfinders created will be sent to the ICT department for hosting on the Council website. The only additional page that needs to be created is a Collections Index that will link to each of the pathfinders.

Chapter 4:

Permanent Conservation

Digital preservation

The conservation of digital images is a problematic area, as digital surrogates are regarded as being particularly vulnerable to loss and damage. The Research Libraries Group (RLG) state that this is because:

They are stored on fragile magnetic and optical media that deteriorate rapidly and that can fail suddenly from exposure to heat, humidity, airborne contaminants, or faulty reading and writing devices. Even if the media are preserved intact, digital materials become unreadable if the playback devices necessary to retrieve information from the media become obsolete or if the software that translates digital information from machine- to human-readable form is no longer available.⁵⁸

It is the obsolescence of software and hardware that is the major risk to digitisation projects, as strategies need to be put in place to avoid materials becoming redundant. This can happen in a short space of time and so some libraries employ “refreshing” where digital items are moved from one medium to another or a new copy is produced⁵⁹. In addition, strategies are constructed to “migrate” files to new software or hardware when it becomes apparent that the way that they are stored is likely to become unusable⁶⁰. This process often involves reformatting, during which some information from the resource may be lost, and perhaps its entire integrity. This process also has cost implications if an item has to be migrated several times. In fact rescanning documents may, in the future, prove to be a cheaper option (always provided that the original documents are still available to scan).

Another option is to maintain “emulation” programs which mimic old hardware and software allowing older file formats to be accessed. Again this has large cost implications, if it has to be done several times, and it can also affect the integrity of a digital object. Emulation is something that is more likely to be carried out on a national or international scale with large banks of redundant software and hardware platforms being created and stored to allow older formats to be accessed.

In *Digital culture: maximising the nation’s investment*⁶¹ the issue of digital preservation is discussed. The growing amount of material appearing solely in a digital form is causing concern in the national repositories as this material is not being preserved. As well the problems outlined above there are legal problems over copyright, selecting what should and should not be stored and what storage format should be employed. The report notes a general lack of appreciation in all communities of the importance of digital preservation. They recommend the use of open formats for saving data and the forming of strategies with regard to preserving digital content.

National repositories of digital data are still a way off and each institution needs to be aware of the preservation needs for their own digital collections. The maintenance of collections needs to be taken into account, as it is an ongoing cost that any project will have to bear.

Digital preservation and East Dunbartonshire

In approaching digital preservation East Dunbartonshire library staff consider the perceived short lifespan of digital images as significant and therefore regard digitised images as access rather than preservation surrogates. Despite this, it is important to get as long a lifespan as is possible from the digital images that are created. To do this certain strategies need to be built in to the life cycle of the digital collection. Firstly, using an open file format for digitising and minimising the number of different formats in use will make it easier to make any transitions that need to take place. Up until now digitisation has used the Kodak Photo CD format. This may continue in the short term, but a move will be made to the more widely supported TIFF format, which is likely to have a greater lifespan. In anticipating such a move, the library plans to

create TIFF versions of the existing Kodak Photo CDs, so that if migration needs to take place all digital images will be available in the same format.

Secondly, a process of assessment will take place every two years to look into the need for possible migration. This will involve an evaluation of the status of the file formats and software/hardware in use, to assess whether there is a need to consider migrating the collection to a new platform. At this time a refreshing of the digital surrogates will take place either by moving the images onto replacement CD-ROMs or onto any new medium that is considered appropriate. As the items are moved they will be checked to make sure that the images remain valid.

35mm negatives

The library already holds a collection of around 30,000 negatives that are kept under archival conditions. The expected lifespan of a negative under these conditions is around 75 years. This has been the primary method of archiving of photographs in East Dunbartonshire and if new prints are added a copy negative is made and added to the archive. The lifespan of these negatives is superior to digital means of preservation and so it will remain as the primary archive for photographs.

The archive has been in existence for only twenty-six years, but some of the negatives that have been donated are older than the archive and are now starting to show signs of deterioration. There is a need to develop a preservation strategy for these negatives. This will be developed once an archivist is appointed to the staff, under present restructuring arrangements.

Conservation microfilm

Microfilm is the most reliable conservation format currently available, providing a predicted a lifespan of 500 years⁶². At present local newspapers are preserved on microfilm, but there remains a large collection of printed material, some of which is beginning to deteriorate. As with the photographs this will be addressed when an archivist is appointed. However, as items are digitised those that are print based will

be converted, additionally, to microfilm if they are regarded as being vulnerable to loss.

Conclusion

From the research we have concluded that it is possible for a library service the size of East Dunbartonshire's to create a digital collection on a continuing basis. It does, however, involve the start-up cost of procuring an image database, securing expandable web space and training. The staff-time element is probably the most significant factor in creating a collection. Once the initial start-up is achieved it should be possible for the library to proceed at its own pace. If this is simply work done on top of existing duties progress is likely to be very slow. It is recommended that, as far as possible, specific periods of staff time should be allocated to the digitisation programme.

An adequate source of funding is central to the aim of producing and maintaining digital resources in quantity, on an ongoing basis. The preservation of these resources requires constant attention, involving additional costs. For these resources to be sustainable their creation needs to become a part of mainstream library budgets, rather than project-based schemes. It is only then that the range of projects currently under way will achieve lasting meaning.

This project emphasises the importance of interoperability and future proofing of systems and practices. The methods for digitisation outlined in the project: scanning to benchmark standards, use of open formats, use of established metadata formats and formal subject description, are all recommendations that aim to maximise the investment made. Firstly by ensuring that work does not have to be redone anymore than is necessary and secondly by enabling resources to have maximum interoperability with other systems, so that the benefits of the Internet can be fully exploited.

The methods recommended have not yet been fully absorbed into the day-to-day practices of the library. Moving the duties of metadata creation to library staff will involve a period of adjustment to new practices. It is intended that the creation of metadata, as described in this report, will seem straightforward and easy to accomplish. The use of a formal subject classification inevitably requires extra

learning time as compared to schemes using simple database entry, but the advantage of using global standards will justify the additional training required.

Bibliography

Ariadne. Available URL: <http://www.ariadne.ac.uk/> (checked 04/01/02).

Beckett, D., Miller, E., Brickley, D. An XML encoding of simple Dublin Core. *DCMI* 11/04/2001. Available URL: <http://dublincore.org/documents/2001/04/11/dcmes-xml/> (checked 04/01/02).

Bradley, Cara. The XML files: the truth will be out there. *School of Library and Information Studies: University of Alberta*. Available URL: <http://www.slis.ualberta.ca/538-99/cbradley/xml.htm> (checked 04/01/02).

CAIRNS. SLIC & CDLR. Available URL: [Centre for Digital Library Research: CDLR Projects: CAIRNS](#) (checked 04/01/02).

Californian digital library: digital image format standards. Californian Digital Library, 2001. Available URL: <http://www.cdlib.org/about/publications/CDLImageStd-2001.pdf> (checked 04/01/02).

California heritage collection: digitising the collection: image capture. Available URL: <http://sunsite.berkeley.edu/CalHeritage/image.html> (checked 04/01/02).

Centre of Digital Library Research. Available URL: <http://cdlr.strath.ac.uk/> (checked 04/01/02).

Chapman, A., Day, M., Hiom, D. Metadata: cataloguing practice and Internet subject-based information gateways. *Ariadne* 18. Available URL: <http://www.ariadne.ac.uk/issue18/metadata/> (checked 04/01/02).

Cataloguing: Marc21 comes of age. *Library Association Record* 05/2001. Available URL: <http://www.la-hq.org.uk/directory/record/r200105/article1.html> (checked 22/08/01).

CORC. OCLC. Available URL: <http://www.oclc.org/corc/> (checked 04/01/02).

Covert, Kay. How the OCLC CORC service is helping weave libraries into the web. *Online Information Review*. 25 (1) 2001. pp. 41-46.

Day, Michael. Metadata for preservation: CEDARS project document AIW01. UKOLN. Available URL: <http://www.ukoln.ac.uk/metadata/cedars/AIW01.html> (checked 04/01/02).

Deegan, Marilyn and Tanner, Simon. *Digital Futures: strategies for the information age*. London: Library Association, 2002. p.189.

Developing a digital preservation strategy for JSTOR, an interview with Kevin Guthrie. *RLG diginews* 4 (4), 2000. Available URL: <http://www.rlg.org/preserv/diginews/diginews4-4.html> (checked 04/01/02)

D-Lib Forum. Available URL: <http://www.dlib.org/> (checked 04/01/02).

Dublin Core elements, qualifiers, and schemes for CORC resource records. OCLC. July 2000. Available URL: http://www2.oclc.org/oclc/pdf/corc_dc_format_7032000.pdf (checked 04/01/02).

Dublin Core Metadata Initiative (DCMI). 2001. Available URL: <http://dublincore.org/> (checked 04/01/02).

Dublin Core metadata element set, version 1.1: reference description. DCMI 02/07/99. Available URL: <http://dublincore.org/documents/1999/07/02/dces/> (checked 04/01/02).

Dublin Core qualifiers. DCMI 11/07/2000. Available URL: <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/> (checked 04/01/02).

E-GIF: e-government interoperability framework. *Cabinet Office: ukgovtalk*. Available URL: <http://www.govtalk.gov.uk/interoperability/egif.asp> (checked 04/01/02).

El-Sherbini, Magda. Metadata and the future of cataloguing. *Library Review*. 50 (1) 2001. pp. 16-27.

End of MARC?, The. *ALA*. Available URL: <http://www.ala.org/online/ts/ts400.html> (checked 04/01/02).

Feeney, Mary (ed.). *Digital culture: maximising the nation's investment, a synthesis of JISC/NPO studies on the preservation of electronic materials*. London: National Preservation Office, 1999. p.12.

General guidelines for descriptive metadata creation & entry. *Colorado Digitization Project*. 1999. Available URL: <http://coloradodigital.coalliance.org/glines.html> (checked 04/01/02).

Greenstein, Daniel and George, Gerald (Digital Library Federation). Digital reproduction quality: benchmark recommendations. *RLG Diginews* 5 (4), 2001. Available URL: <http://www.rlg.org/preserv/diginews/diginews5-4.html#featured> (checked 04/01/02).

Hedstrom, Margaret and Montgomery, Sheon. *Digital preservation needs and requirements in RLG member institutions*. Research Libraries Group, 1998. Available URL: <http://www.rlg.org/preserv/digpres.html> (checked 04/01/02).

Hickey, Thomas B. CORC: a system for gateway creation. *Online Information Review*. 24 (1) 2000. pp. 49-53.

Hillmann, Diane. Using Dublin Core. *DCMI* 12/04/2001. Available URL: <http://dublincore.org/documents/2001/04/12/usageguide/> (checked 04/01/02).

HILT. RSLP and JISC, 2001. Available URL: <http://hilt.cdlr.strath.ac.uk/index.html> (checked 04/01/02).

HILT. *Outcomes of HILT consultation: draft interim report*. 2001. Available URL: <http://hilt.cdlr.strath.ac.uk/Reports/Consultation.html> (checked 04/01/02).

House of Images. Available URL: <http://www.hoimages.co.uk/> (checked 04/01/02).

Hurst, Jill Ann. MARC vs. Dublin Core: summary of responses. *Imagelib*. 12/07/99. Available URL: <http://library.wustl.edu/~listmgr/imagelib/Jul1999/0015.html> (checked 04/01/02).

Ibase. Available URL: <http://www.ibase.com/heritage/home.html> (checked 04/01/02).

InVisage. Talis/Ibase. Available URL: <http://www.talis.com/products/infdelserv/invisage.htm> (checked 04/01/02).

Kenney, Anne R. and Reiger, Oya Y. *Establishing a central depository for preserving digital image collections, report of the Digital Preservation Policy Working Group*. Cornell University, 2001. Available URL: <http://www.library.cornell.edu/preservation/kodak/cover.htm> (checked 04/01/02).

Kenney, Anne R. and Reiger, Oya Y. (ed.) *Moving theory into practice: digital imaging for libraries and archives*. Mountain View, Calif. : Research Libraries Group, 2000.

Kenney, Anne R. and Reiger, Oya Y. *Using Kodak Photo CD technology for preservation and access: a guide for librarians, archivists, and curators*. Department of Preservation and Conservation, Cornell University Library, 1988. Available URL: <http://www.library.cornell.edu/preservation/pub.htm> (checked 04/01/02).

Kim, Hyun-Hee, Choi, Chang-seok. XML: how it will be applied to digital library systems. *The Electronic Library*. 18 (3) 2000. pp. 183-189.

Kodak Photo CD. Available URL:

<http://www.kodak.com/global/en/professional/products/storage/pcdMaster/aboutPCD.shtml> (checked 04/01/02).

Kodak: faq: Kodak photo cd technology/transferring images. Eastman Kodak company, 1984-2002. Available URL:

<http://www.kodak.com/global/en/service/faqs/faq1001c.shtml> (checked 04/01/02).

Lee, Stuart D. *Digital imaging a practical handbook.* London: Library Association. 2001.

Sir Thomas Lipton Digitisation Project. Mitchell Library. Available URL:

<http://www.mitchelllibrary.org/STL1.html>(checked 04/01/02).

MARC 21. Library of Congress. Available URL: <http://lcweb.loc.gov/marc/marc.html> (checked 04/01/02).

Maxted, Ian. *The Devon core: digital data collection and transfer.* Devon Library and Information Services. Available URL:

<http://www.devon.gov.uk/library/locstudy/devcore.html> (checked 04/01/02).

Medeiros, Norm. Making room for MARC in a Dublin Core world. *Online.*

November 1999. Available URL: <http://www.cbuc.es/5digital/3.pdf> (checked 04/01/02).

Medlane experiment MARC to XML. Lane Medical Library, Stanford University Medical Centre. Available URL: <http://xmlmarc.stanford.edu/> (checked 04/01/02).

Metadata for digital preservation the cedars project outline specification draft for public consultation. *The Cedars Project Team and UKOLN.* March 2000. Available URL: <http://www.leeds.ac.uk/cedars/OutlineSpec.htm> (checked 04/01/02).

Miller, Dick R. XML: libraries' strategic opportunity. *Library Journal Netconnect*. Summer 2000. Available URL: <http://www.libraryjournal.com/xml.asp> (checked 04/01/02).

Miller, Dick, R. XML and MARC: a choice or a replacement?, in, *MARBI (Machine-Readable Bibliographic Information)*, *ALCTS/LITA/RUSA joint meeting with CC:DA: ALA annual conference, July 8-10, 2000*. ALA: Chicago, IL. Available URL: <http://www.ala.org/alcts/organization/div/marbi/marbi-ccda.html> (checked 04/01/02).

Miller, Paul. Interoperability: what is it and why should I want it? *Ariadne* 24, 21/06/00. Available URL: <http://ariadne.ac.uk/issue24/interoperability/>(checked 04/01/02).

Miller, Paul. Z29.50 for all. *Ariadne*, 21. 1999. Available URL: <http://www.ariadne.ac.uk/issue21/z3950/> (checked 04/01/02)

Nicholson, Dennis. Electronic compilation. *Scottish Libraries*, 83 no.5 (14), 2000, p12.

Nicholson, Dennis., Denham, Mark., Dunsire, Gordon. and Gillis, Helena. *CAIRNS project final report: an embryonic cross-sectoral, cross-domain national networked information service for Scotland*. Glasgow: CAIRNS, 2000.

Puglia, Steven. *U.S. National Archives and Records Administration: electronic access project scanning and file format matrix*. Available URL: <http://www.nara.gov/nara/vision/eap/eapspec.html> (checked 04/01/02).

RLG diginews. Research Libraries Group. Available URL: <http://www.rlg.org/preserv/diginews/> (checked 04/01/02).

SCONE. Centre for Digital Library Research. Available URL: <http://scone.strath.ac.uk/service/index.cfm> (checked 04/01/02).

Scottish Cultural Resource Access Network (SCRAN). 2001. Available URL: <http://www.scran.ac.uk/homepage/> (checked 04/01/02).

Scottish Library and Information Council (SLIC). Slainte, 2001. Available URL: <http://www.slainte.org.uk/Slic/slichome.htm> (checked 04/01/02).

Scottish Newspapers and Microfilm Unit. Available URL: <http://www.snm.co.uk> (checked 04/01/02)

Senecal, Kirstin. The twofold promise of the CORC project. *OCLC systems and services*. 16 (2) 2000. pp. 84-90.

Tanner, Smith and Robinson, Brian. *JISC image digitisation initiative: feasibility study, final report*. Hertfordshire: University of Hertfordshire/HEDS, 1998. Available URL: http://heds.herts.ac.uk/resources/papers/jidi_fs.pdf (checked 04/01/02).

Tanner, Simon and Smith, Joanne Lomax. *Digitisation: how much does it really cost?* HEDS: London 1999. Available URL: <http://www.bodley.ox.ac.uk/scoping/digitization.html> (checked 04/01/02).

Virtual Mitchell. Mitchell Library. Available URL: <http://www.libarch.glasgow.gov.uk/vm/> (checked 04/01/02).

Visual evidence: the photographic presentation of landscape and people. RSLP. Available URL: <http://www.rslp.ac.uk/projects/> (checked 04/01/02).

Z39.50 maintenance agency. Library of Congress Network Development & MARC Standards Office. Available URL: <http://www.loc.gov/z3950/agency/> (checked 04/01/02) Appendix i: Sample MARC 21 record from CORC

Appendix I: Sample MARC 21 record from CORC

006 m c l
040 CX@EDLM †c CX@
092 0_ 331.76223340941434 †2 21
245 00 [Petticoat Dan: full length portrait of Dan holding shovel, while resting against wheelbarrow] †h [Electronic resource].
256 Electronic image data (1 file).
260 Kirkintilloch, Scotland: †b East Dunbartonshire Libraries, †c 2001.
270 William Patrick Library, Reference Department, †a 2 West High Street, †b Kirkintilloch, †c East Dunbartonshire, †d Scotland. †e G66 1AD †m ref@edlib.freeserve.co.uk
500 Greyscale.
500 Title supplied by cataloguer.
500 Part of the East Dunbartonshire Council website.
520 "Petticoat Dan" was a well-known character in Kirkintilloch. Renowned for his simple nature and his practice of wearing a petticoat (both the result of retardation caused by a childhood bout of measles), Dan would shovel coal for the residents of the town. His notoriety resulted from the innocence with which he approached the world. When shown a photo of himself (possibly this one) he is reported as saying "I'm no' shair o'myself', but its awfu' lek my shool!", i.e. I'm not sure about myself, but it's awfully like my shovel.
530 Available as a b&w photographic copy print; †b Reference Department, William Patrick Library, Kirkintilloch; †c For purchase; †d N2757.
530 Available as an electronic image in PhotoCD format; †b Reference Department, William Patrick Library, Kirkintilloch; †c For use in the library only. †d PCD/Creating a Permanent Digital Archive/cd1/43.
530 Available as a 35 mm b&w photographic negative; †b Reference Department, William Patrick Library, Kirkintilloch; †c For use by Library staff only.
534 †p Reproduction of: †a [Portraits, Cooper (Dan) "Petticoat Dan" Kirkintilloch worthy]. †e 1 photograph †l Reference Library Photographs, William Patrick Library. †n Part of: Isaac Black collection, Kirkintilloch local photographs.
538 Mode of access: World Wide Web.
540 There are copyright restrictions applying to the commercial reproduction of this image; †b East Dunbartonshire Council.
600 04 Petticoat Dan †d 1835-1913.
740 0_ Isaac Black collection, Kirkintilloch local photographs.
856 40 www.slainte.org.uk †d edpdalm †f N2757a.jpg †u <http://www.slainte.org.uk/edpdalm/N2757a.jpg> †q JPEG

Appendix II

MARC constant data from CORC

Constant Data Name: EDFull

- [006](#) m c l
- [040](#) CX@EDLM †c CX@
- [245](#) 00 *Title* †h [Electronic resource].
- [256](#) Electronic image data (1 file).
- [260](#) Kirkintilloch, Scotland: †b East Dunbartonshire Libraries, †c 2001.
- [270](#) William Patrick Library, Reference Department, †a 2 West High Street, †b Kirkintilloch, †c East Dunbartonshire, †d Scotland. †e G66 1AD †m ref@edlib.freeserve.co.uk
- [500](#) *Greyscale/Colour*.
- [500](#) Title supplied by cataloguer.
- [500](#) Part of the East Dunbartonshire Council website.
- [520](#) *Summary*.
- [530](#) Available as a b&w photographic copy print; †b Reference Department, William Patrick Library, Kirkintilloch; †c For purchase; †d *order number*.
- [530](#) Available as an electronic image in PhotoCD format; †b Reference Department, William Patrick Library, Kirkintilloch; †c For use in the library only. †d *PCD/Creating a Permanent Digital Archive/cd1/*
- [530](#) Available as a 35 mm b&w photographic negative; †b Reference Department, William Patrick Library, Kirkintilloch; †c For use by Library staff only.
- [534](#) †p Reproduction of: †t *Title of original*. †e *Physical description* †l *Location of original*. †n Part of: *Collection title*
- [538](#) Mode of access: World Wide Web.
- [540](#) There are copyright restrictions applying to the commercial reproduction of this image; †b East Dunbartonshire Council.
- [740](#) 0_ *collection title*
- [856](#) 40 www.slainte.org.uk †d edpdalm †f *filename*.jpg †u http://www.slainte.org.uk/edpdalm/*filename*.jpg †q JPEG

Appendix III:

Format for cataloguing local history images

(245) Title: put in [square brackets] no full stop	
(500) Colour:	Colour/greyscale.
(520) Summary: (short description of image)	
(530 \$d) Negative no.	
(530 \$d) CD no.	
(534) Reproduction of: \$a Title of original (if it has one, include photo subject heading)	
(\$c) publisher of original (if relevant, and different from us).	
(\$e) medium of original. (include extent of item i.e. no. of pages)	
(\$l) location of original. shelf-mark and physical location	Shelf-mark: and/or one of the following: Reference storeroom: Reference department, William Patrick Library, Kirkintilloch. Reference archive: Reference department, William Patrick Library, Kirkintilloch. Reference photographs: Reference department, William Patrick Library, Kirkintilloch. Brookwood photographs: Brookwood Library, Bearsden. Brookwood archive: Brookwood Library, Bearsden.
(740) Title added entry: Collection name	
(856) URL no.	Will be the same as the negative no.
Subject headings:	

Appendix IV:

Sample pathfinder

Kirkintilloch Burgh collections, photographs.

A collection of photographs relating to the burgh of Kirkintilloch.

[\[Thomas Johnston signing the Kirkintilloch Burgess Roll as part of a ceremony giving him the freedom of the burgh, 4th July](http://www.slainte.org.uk/edpdalm/N1245b.jpg)

[1931\]](http://www.slainte.org.uk/edpdalm/N1245b.jpg) (http://www.slainte.org.uk/edpdalm/N1245b.jpg)

Photograph of Thomas Johnston signing the Burgess Roll in front of gathered audience. Thomas Johnston was born in Kirkintilloch in 1881 and had a wide-ranging career which embraced politics, history and journalism. In 1906 he became editor of Forward, a socialist weekly, that provided him with a stage to promote his socialist ideals. He used the publication as stage for producing historical accounts that attacked the exploitation of the poor and working classes. His political career began in local government and he eventual rose to become Secretary for State for Scotland in Winston Churchill's wartime government, 1941-45. He died at his home in Milngavie in 1965.

[\[Provost Fairservice presenting keys to Kirkintilloch's first overspill tenants from Glasgow, June 1959\]](http://www.slainte.org.uk/edpdalm/N3200a.jpg) (http://www.slainte.org.uk/edpdalm/N3200a.jpg)

Provost Fairservice shaking hands and handing over the first key. During the 1950s it became apparent that Glasgow's population was living in more high density and sub-standard housing than in any other city in Britain. The Housing & Town Development Act of 1957 contained provisions "to enable Scottish local authorities to provide housing accommodation and other development in relief of the needs of districts other than their own". As a result of this act, towns such as Kirkintilloch became involved in creating "overspill" housing for Glasgow residents who came from overcrowded areas that were to be demolished and rebuilt.

Copy prints of images are available for purchase. Please contact East Dunbartonshire libraries for further information.

East Dunbartonshire Libraries

All images are copyright East Dunbartonshire Libraries, unless otherwise stated. No commercial reproduction without prior consent.

(20011224) Send comments or requests for prints to: Reference department, William Patrick Library

References

- ¹ Lee, Stuart D. *Digital imaging a practical handbook*. London: Library Association. 2001.
- ² Kenney, Anne R. and Reiger, Oya Y. (ed.) *Moving theory into practice: digital imaging for libraries and archives*. Mountain View, Calif. : Research Libraries Group, 2000.
- ³ Besser, Howard and Trant, Jenifer. *Introduction to imaging: issues in constructing an image database*. Getty Museum, 2000. Available URL:
<http://www.getty.edu/research/institute/standards/introimages/index.html> (checked 04/01/02)
- ⁴ *Ariadne*. Available URL: <http://www.ariadne.ac.uk/> (checked 04/01/02).
- ⁵ *D-Lib Forum*. Available URL: <http://www.dlib.org/> (checked 04/01/02).
- ⁶ RLG diginews. Research Libraries Group. Available URL: <http://www.rlg.org/preserv/diginews/> (checked 04/01/02).
- ⁷ *Scottish Library and Information Council (SLIC)*. Slainte. Available URL:
<http://www.slainte.org.uk/Slic/slichome.htm> (checked 04/01/02).
- ⁸ *MARC 21*. Library of Congress. Available URL: <http://lcweb.loc.gov/marc/marc.html> (checked 04/01/02).
- ⁹ *Z39.50 maintenance agency*. Library of Congress Network Development & MARC Standards Office. Available URL: <http://www.loc.gov/z3950/agency/> (checked 04/01/02).
- ¹⁰ *CAIRNS*. Centre for Digital Library Research. Available URL: <http://cairns.lib.strath.ac.uk/> (checked 04/01/02)
- ¹¹ Miller, Paul. Z29.50 for all. *Ariadne*, 21. 1999. Available URL:
<http://www.ariadne.ac.uk/issue21/z3950/> (checked 04/01/02)
- ¹² E-GIF: e-government interoperability framework. *Cabinet Office: ukgovtalk*. Available URL:
<http://www.govtalk.gov.uk/interoperability/egif.asp> (checked 04/01/02).
- ¹³ *Medlane experiment MARC to XML*. Lane Medical Library, Stanford University Medical Centre. Available URL: <http://xmlmarc.stanford.edu/> (checked 04/01/02).
- ¹⁴ Kim, Hyun-Hee, Choi, Chang-seok. XML: how it will be applied to digital library systems. *The Electronic Library*. 18 (3) 2000. p.184.
- ¹⁵ Tanner, Simon and Smith, Joanne Lomax. *Digitisation: how much does it really cost?* HEDS: London 1999. Available URL: <http://www.bodley.ox.ac.uk/scoping/digitization.html> (checked 04/01/02)
- ¹⁶ *ibid.*
- ¹⁷ *ibid.*
- ¹⁸ Lee, Stuart D. *op. cit.* p.78.
- ¹⁹ *ibid.* p.48.
- ²⁰ *Californian digital library: digital image format standards*. Californian Digital Library, 2001. Available URL: <http://www.cdlib.org/about/publications/CDLImageStd-2001.pdf> (checked 04/01/02).

-
- ²¹ Puglia, Steven. *U.S. National Archives and Records Administration: electronic access project scanning and file format matrix*. Available URL: <http://www.nara.gov/nara/vision/eap/eapspec.html> (checked 04/01/02).
- ²² General guidelines for descriptive metadata creation & entry. *Colorado Digitization Project*. 1999. Available URL: <http://coloradodigital.coalliance.org/glines.html> (checked 04/01/02).
- ²³ Kenney, Anne R. and Reiger, Oya Y. *Establishing a central depository for preserving digital image collections, report of the Digital Preservation Policy Working Group*. Cornell University, 2001. Available URL: http://www.library.cornell.edu/preservation/IMLS/image_deposit_guidelines.pdf (checked 04/01/02).
- ²⁴ Greenstein, Daniel and George, Gerald (Digital Library Federation). Digital reproduction quality: benchmark recommendations. *RLG Diginews* 5 (4), 2001. Available URL: <http://www.rlg.org/preserv/diginews/diginews5-4.html#featured> (checked 04/01/02).
- ²⁵ Tanner, Smith and Robinson, Brian. *JISC image digitisation initiative: feasibility study, final report*. Hertfordshire: University of Hertfordshire/HEDS, 1998. Available URL: http://heds.herts.ac.uk/resources/papers/jidi_fs.pdf (checked 04/01/02).
- ²⁶ *NOF-Digitise Technical Standards. op.cit.*
- ²⁷ Lomax, Joanne. *Images of the past, technology of the future: creating digital image collections*. Higher Education Digitisation Service (HEDS), 1990. p.3 Available URL: <http://heds.herts.ac.uk/resources/papers/HEDSLIS.pdf> (checked 04/01/02)
- ²⁸ Kenney, Anne. *Joint RLG and NPO preservation conference: guidelines for digital imaging*. RLG, 1999. Available URL: <http://www.rlg.org/preserv/joint/kenney.html> (checked 04/01/02).
- ²⁹ Kenney, Anne R. and Reiger, Oya Y. (ed.) *Moving theory into practice: digital imaging for libraries and archives*. Mountain View, Calif. : Research Libraries Group, 2000. p.27.
- ³⁰ Developing a digital preservation strategy for JSTOR, an interview with Kevin Gutherie. *RLG diginews* 4 (4), 2000. Available URL: <http://www.rlg.org/preserv/diginews/diginews4-4.html> (checked 04/01/02)
- ³¹ *Scottish Newspapers and Microfilm Unit*. Available URL: <http://www.snm.co.uk> (checked 04/01/02)
- ³² *Ibid.* p.39
- ³³ *Kodak: faq: Kodak photo cd technology/transferring images*. Eastman Kodak company, 1984-2002. Available URL: <http://www.kodak.com/global/en/service/faqs/faq1001c.shtml> (checked 04/01/02).
- ³⁴ Kenney, Anne R. and Reiger, Oya Y. *Using Kodak Photo CD technology for preservation and access: a guide for librarians, archivists, and curators*. Department of Preservation and Conservation, Cornell University Library, 1998. Available URL: <http://www.library.cornell.edu/preservation/kodak/cover.htm> (checked 04/01/02).
- ³⁵ *California heritage collection: digitising the collection: image capture*. Available URL: <http://sunsite.berkeley.edu/CalHeritage/image.html> (checked 04/01/02)

-
- ³⁶ Developing a digital preservation strategy for JSTOR, an interview with Kevin Gutherie. *RLG diginews* 4 (4), 2000. Available URL: <http://www.rlg.org/preserv/diginews/diginews4-4.html> (checked 04/01/02)
- ³⁷ Lee, Stuart. *op.cit.*p.45.
- ³⁸ *Centre for Digital Library Research*. Available URL: <http://cdlr.strath.ac.uk/> (checked 04/01/02)
- ³⁹ Miller, Paul. Interoperability: what is it and why should I want it? *Ariadne* 24, 21/06/00. Available URL: <http://ariadne.ac.uk/issue24/interoperability/> (checked 04/01/02).
- ⁴⁰ *CAIRNS. op. cit.*
- ⁴¹ *SCONE*. Centre for Digital Library Research. Available URL: <http://scone.strath.ac.uk/service/index.cfm> (checked 04/01/02).
- ⁴² Nicholson, Dennis. Electronic compilation. *Scottish Libraries*, 83 no.5 (14), 2000. p12.
- ⁴³ Dublin Core Metadata Initiative. Available URL: <http://dublincore.org> (checked 04/01/02)
- ⁴⁴ Maxted, Ian. *The Devon core: digital data collection and transfer*. Devon Library and Information Services. Available URL: <http://www.devon.gov.uk/library/locstudy/devcore.html> (checked 04/01/02).
- ⁴⁵ *Scottish Cultural Resource Access Network (SCRAN)*. 2001. Available URL: <http://www.scran.ac.uk/homepage/> (checked 04/01/02).
- ⁴⁶ *Virtual Mitchell. Mitchell Library*. Available URL: <http://www.libarch.glasgow.gov.uk/vm/> (checked 04/01/02).
- ⁴⁷ *Sir Thomas Lipton Digitisation Project. Mitchell Library*. Available URL: <http://www.mitchelllibrary.org/STL1.html> (checked 04/01/02).
- ⁴⁸ *Ibase*. Available URL: <http://www.ibase.com/heritage/home.html> (checked 04/01/02).
- ⁴⁹ *Visual evidence: the photographic presentation of landscape and people*. RSLP. Available URL: <http://www.rslp.ac.uk/projects/> (checked 04/01/02).
- ⁵⁰ *InVisage*. Talis/Ibase. Available URL: <http://www.talis.com/products/infdelserv/invisage.htm> (checked 04/01/02).
- ⁵¹ MARC 21. Library of Congress. Available URL: <http://lcweb.loc.gov/marc/marc.html> (checked 04/01/02).
- ⁵² Nicholson, Dennis., Denham, Mark., Dunsire, Gordon. and Gillis, Helena. *CAIRNS project final report: an embryonic cross-sectoral, cross-domain national networked information service for Scotland*. Glasgow: CAIRNS, 2000.
- ⁵³ *HILT*. RSLP and JISC, 2001. Available URL: <http://hilt.cdrl.strath.ac.uk/index.html> (checked 04/01/02)
- ⁵⁴ *HILT. Outcomes of HILT consultation: draft interim report*. 2001. Available URL: <http://hilt.cdrl.strath.ac.uk/Reports/Consultation.html> (checked 04/01/02)
- ⁵⁵ *CORC. OCLC*. Available URL: <http://www.oclc.org/corc/> (checked 04/01/02).
- ⁵⁶ *Ibase. op. cit.*
- ⁵⁷ *House of Images*. Available URL: <http://www.hoimages.co.uk/> (checked 04/01/02).

⁵⁸ Hedstrom, Margaret and Montgomery, Sheon. *Digital preservation needs and requirements in RLG member institutions*. Research Libraries Group, 1998. Available URL:

<http://www.rlg.org/preserv/digpres.html> (checked 04/01/02)

⁵⁹ Deegan, Marilyn and Tanner, Simon. *Digital Futures: strategies for the information age*. London: Library Association, 2002. p.189.

⁶⁰ Feeney, Mary (ed.). *Digital culture: maximising the nation's investment, a synthesis of JISC/NPO studies on the preservation of electronic materials*. London: National Preservation Office, 1999. p.12.

⁶¹ *ibid.*

⁶² Deegan, Marilyn and Tanner, Simon. *Digital Futures: strategies for the information age*. London: Library Association, 2002. p.189.

Acknowledgements

The project could not have been completed without the help of the Centre for Digital Library Research and in particular Gordon Dunsire who contributed a great deal of time and effort; also Dennis Nicholson.

Grateful mention should also be made of the staff at East Dunbartonshire Libraries and East Dunbartonshire ICT Department, the Scottish Newspapers Microfilming Unit, Eugene Duffy, Duncan Irving, David McMenemy, Ian Maxted and Alan Poulter.

Glossary

AAT	- Art and Architecture Thesaurus.
AACR2	-Anglo American Cataloguing Rules (second edition).
CDLR	- Centre for Digital Library Research.
CAIRNS	- Co-operative Academic Information Retrieval Network for Scotland.
CORC	- Cooperative Online Resource Cataloging.
DDC	- Dewy Decimal Classification.
dpi	- dots per inch.
DTD	- Document Type Definition.
eGif	- electronic Government Interoperability Framework.
GIF	- Graphical Interchange Format.
GRS-1	- General Record Syntax 1.
JPEG	- Joint Photographic Experts Group format.
JIDI	- JISC Image Digitisation Initiative
HEDS	- Higher Education Digitisation Service. (UK)
HILT	- High Level Thesaurus.
HTML	- HyperText Markup Language.
LCSH	- Library of Congress Subject Headings.
MARC	- MACHine Readable Cataloging.
NARA	- U.S National Archives and Records Administration.
nof	- New Opportunity Fund.
nof-digitise	- New Opportunity Fund's program of digitisation.
OPAC	- Online Public Access Catalogue.
PCD	- Photo CD.
RDF	- Resource Description Framework.
RLG	- Research Libraries Group.
SCONE	- Scottish COllections Network Extension project.
SCRAN	- Scottish Cultural Resource Access Network.
SCURL	- Scottish Confederation of University Research Libraries.
SLAINTE	- Scottish Libraries Across the Internet.
SLIC	- Scottish Library & Information Council.
SNMU	- Scottish Newspapers Microfilming Unit.
TIFF	- Tagged Image File Format.
UDC	- Universal Decimal Classification
UNESCO	- United Nations Educational, Scientific and Cultural Organisation
URL	- Uniform Resource Locator.
XML	- Extensible Markup Language
