Good Practice Handbook

Version 1.2

edited by the Minerva Working Group 6
Identification of good practices and competence centres

November, 2003

Texts by
Karl-Magnus Drake, Borje Justrell (Riksarkivet, Sweden)
Anna Maria Tammaro (University of Parma. Italy)

WP6 Secretariat:
Marzia Piccininno

TABLE OF CONTENTS

1. Acknowledgements

This handbook had not been more then a good idea without Ciaran Clissman. Besides all his knowledge about structuring information, his way of putting good practice into proper English has been an invaluable contribution to this work. Remaining peculiarities in the text should be blamed only on the editors, because we have not always followed Ciaran Clissman´s advices.

We will also acknowledge the members of the Minerva project and the National Representative Group who have shared their experience and great knowledge with us. At last but not least, a special thanks to the Italian national working group on good practice for their great interest in the handbook and their valuable contributions to its content.

2. Introduction

2.1 Document Overview

This document is a result of the Minerva project's good practice working group (WP 6). It presents a practical handbook to the establishment, execution and management of digitisation projects, with particular focus on the cultural area (libraries, museums, archives). The target audience of this handbook is teams within and across cultural institutions who are contemplating, or are already executing, digitisation projects. The document reflects the outcome of the work carried out by WP 6 of the Minerva project, including the substantial research represented by the national questionnaires completed in connection with the National Representatives Group (NRG) meeting in Alicante, May 2002.

2.2 Document Structure

This document has the following chapters:

- Background
- Practical Guidelines
- Relevant Standards
- A Selected List of Digitisation Guidelines

Background - This reviews the relevant aspects of the Minerva project, and states the role of this document in the overall progress of the project. It also covers the work carried out to date, in order that the reader shall have a clear picture of the context in which this document should be considered.

Practical Guidelines - The most important practical lessons learnt and information collected by the Minerva project best practice team are presented. This focuses on a significant number of practical 'rules of thumb' which should be considered by organisations which are establishing, executing or managing digitisation projects in the cultural sphere. The guidelines are divided into the following areas, each of which reflects a stage in the life-cycle of a digitisation project:

- Digitisation project planning
- Selecting Source Material for Digitisation
- Preparation for Digitisation
- Handling of Originals
- The Digitisation Process
- Preservation of the Digital Master Material
- Meta-data

- Publication
- IPR and Copyright
- Managing Digital Projects

The guidelines are presented in a pragmatic manner, aimed at the hands-on project team, and are supported by relevant references to examples of best practice, competence centres and role models which are being carried out in the European cultural field, as well as by global links to appropriate and useful online resources.

It may be noted that there are several other sources of guidelines on digitisation and the creation of digital cultural content. The most important ones are noted in the selected list of digitisation guidelines which is a part of this handbook (see below). However, the target groups of this handbook and those mentioned in the selected list are different. This handbook is aimed at European cultural bodies contemplating or involved in digitisation projects, as well as stakeholders in the developing European content industry.

Standards - An overview of the relevant technical standards is provided in a separate chapter. The Minerva team recognises the wide range of standards available, and have not attempted, in this handbook, to cover any except the most important. The major focus is on technology standards which impinge on the decisions which need to be made during a digitisation project, and include standards in the following areas:

- Image
- Audio
- Digital Video
- 3D
- Meta-data
- Taxonomy and Naming

Digitisation Guidelines: a selected list - A selected list of digitisation guidelines is presented, where each guideline is described in a standardised way: Author, Contributor (if existing), Title, Description, Date, Format and URL. The list is limited to guidelines for digitisation of paper-based documentary heritage like manuscripts/records, printed books and photographs in libraries, archives and museums. The aim is to give the reader an overview of the most important guidelines.

3. Background

3.1 The Lund Principles

On 4th of April 2001, representatives and experts from the European Commission and Member States met at Lund in Sweden (under the Swedish Presidency) to discuss how to coordinate and add value to national digitisation programmes, at a European level. The meeting resulted in the publication of a set of general principles to govern public digitisation initiatives and their coordination. These principles, called the Lund principles, were transformed into the Lund Action plan, which establish a list of actions to be carried out by Member States, by the Commission, and by Member States and the Commission jointly, to improve the digitisation landscape across Europe.

3.2 The Minerva Project

This document is an output of the Minerva project, which was established in 2002 under the leadership of the Italian Ministry of Culture (IST contract 2001-35461). The project comprises representatives of the relevant government ministries or central state agencies from many EU member states, with the common objective of promoting a shared approach and methodology for the digitisation of European cultural material. The project recognises the unique value of the European cultural heritage, and the strategic role which it can play in the growing digital content industry in Europe. It also recognises the value of coordination of the efforts of national governments and cultural organisations, in order to increase the level of synthesis and synergy between and among digitisation initiatives.

The Minerva project has a number of focused working groups within the overall consortium. Each working group is made up of several project partners, working together on a particular aspect of the project objectives. The objectives of each working group are described on the project web site at http://www.minervaeurope.org. The working group structure allows the project to examine a number of the most important areas of the digitisation sphere, in parallel.

The following working groups exist within the project:

- Benchmarking framework
- Identification of good practices and competence centres
- Interoperability and service provision
- Inventories, discovery of digitised content, multilingualism issues
- Identification of user needs, content and quality framework for common access points

Each working group is responsible for a project work-package, as outlined in the project plan. The activities of the working groups include meetings, public workshops, publications (such as this handbook), international coordination and cooperation, etc.

The Minerva project is made up of representatives the EU member states, who are dedicated to the following objectives:

- co-ordination of their strategies and policies for digitisation of cultural content;
- provision of a European dimension to their policies and programmes;
- definition, exchange and dissemination of digitisation good practices across the European Union;
- support of the development of national and international inventories of cultural and scientific content.

Minerva is an acronym for Ministerial Network for Valorising Activities in Digitisation, and its members are representatives of national governments or central state authorities given the task. The Minerva project also includes major national cultural players such as national archives, national libraries, and museums. The project aims to co-ordinate national programmes, and its approach is strongly based on the principle of embedded ness in national digitisation activities.

The work plan of the Minerva project includes activities to:

- organise work groups to provide the political and technical framework for improving digitisation activities of cultural and scientific contents, and defining a common platform;
- facilitate the adoption of the Lund principles, both in EU Member States and other European countries, to amplify the impact of the eEurope initiative;
- set-up an international Forum, and electronic publication, supporting collaboration on scientific research;
- make visible, promote and exchange information about National Policy profiles concerning digitisation;
- identify users' needs, define training schemes and develop recommendations;
- make available test-beds, defining mechanisms for evaluating models, methodologies, techniques and approaches, aiming at the selection of guidelines for harmonising activities and trying to reach agreement among Member States, on a common basis;
- implement the benchmarking framework on digitisation, able to compare and improve quality of national approaches and promote best practice across Europe;

- organise a plenary meeting every six months, hosting also thematic workshops to present and discuss results achieved by the specific work groups;
- promote concertation events open to both EU and other national projects, to create clusters of projects;
- promote dissemination and training activities at national level, acquisition of new skills and access to existing resources;
- identify Road Maps suitable for activities to be launched in the near future, to support Member States in the definition of their policy, through exchange of experience, priorities agenda and work programmes.

The direct involvement of governmental organisations intends to contribute at bringing together a wide network of research centres, cultural organisations and companies interested in digitisation aspects, to co-ordinate their activities in order to advance towards common strategic goals.

3.3 The Role of This Document

This handbook document is an interim output of the best practice task force. It contributes to the achievement of objectives of the project by providing a concrete, pragmatic output from the deliberations of the project, which will allow the benefit of the knowledge and research within the project to be capitalized upon by the widest possible audience. This handbook is aimed at cultural bodies contemplating or involved in digitisation projects, as well as at the stakeholders in the developing European content industry.

This document presents a first harvest from the research carried out to date within that task force, in the form of an easy-to-use and pragmatic set of guidelines for digitisation projects. The handbook makes available the results of the work carried out so far, in a timely manner, and allows third parties to benefit as soon as possible from the work of the project. It also underlines the practical, real-world applicability of the work of the project, and its relevance to its target audience.

It may be noted that there are several other documents available, which share scope with this document. A range of Internet sites provide large amounts of information regarding best practice for digitisation projects. The most important ones are listed in chapter 6.

3.4 Work to Date

This document is one of a series of outputs from the best practice work-package of the Minerva project. The work-package (WP6) has already published a deliverable (state of the art report) describing best practice and competence

centres (D6.1), and is on course to establish appropriate web architectures for digitisation projects. The work carried out includes background research across the world on digitisation projects and on sources of knowledge and guidance which may be of relevance. Several of these are referenced in this document, as well as in D6.1. In addition, all cultural ministries in the EU have provided nominations of projects, competence centres and initiatives, in their home countries, which are examples of good practice in one or more areas. This material is presented on the Minerva web site (List of Good practice) and provides a unique insight into ongoing work within each member state.

4. Practical Guidelines

4.1 Introduction

This chapter presents the core of the handbook. It provides practical guidelines for organisations and bodies contemplating, or involved in, digitisation projects. The emphasis is on the cultural sphere; however, the material is to a large degree relevant to other spheres (e.g. tourism, general document management).

The material in this chapter is broken down in accordance with the stages in the digitisation life-cycle. This means that a reader can easily identify material which is relevant to his work, regardless of how far his own project has progressed. It is anticipated that many users of this handbook will be at the first stage of the project (planning); however, at least some of the material provided here should be of value to any digitisation project.

The digitisation life-cycle stages are used as the basis for the breaking down of the guidelines and mirrored in the table of contents.

Each guideline description is structured in the following way:

- A Guideline Title
- An Issue Definition, which sets the scene for the guideline and/or introduces the problem which the guidelines addresses
- The Guideline Text, a set of pragmatic suggestions which aim to facilitate the relevant aspect of setting up or executing a digitisation project
- Notes or Commentary, where any additional information is provided. This is sometimes empty
- References, which are broken into two parts –
    - Online References, usually links to competence centres and their publications, which address a particular issue explicitly
    - References Nominated by Member States. These references refer to a list of nominated good practice (projects, competence centres, initiative etc) published on the Minerva web site. Each reference may or may not address the particular area explicitly; the link is provided either because the reference can be expected to have experience in a particular area, or because it addresses this area in detail.

Neither the guidelines nor the references is exhaustive – however they provide the most important information needed by a project which is addressing a particular task or tasks within the life-cycle of a digitisation project.

4.2 Digitisation Project Planning

4.2.1 Introduction

Planning of the project is the first step in any digitisation project. Time spent on planning the project will pay dividends in the easier management and execution of the project. Normally, the following questions have to be answered:

- What (work needs to be done)?
- Who (should do it)?
- Where (should it take place)?
- When (will it take place)?
- How (will it be done)?

A digitisation project should have clearly specified goals and objectives – these will impact directly on the selection, copyright and publication. The project should have suitable personnel, with appropriate knowledge and skills, as well as a training plan in place to provide any additional expertise that the project may require.

A project should not begin until some research has been carried out into other projects in the same area. Such research will identify issues which need to be addressed, will stimulate new ideas and areas which might not yet have been considered, and will add value and credibility to the project output.

Research will also help to indicate the amount of work which may be planned for the execution of the project, by meeting or talking with organisations which have completed similar projects. Such interactions will help to establish whether your organisation has the personnel, the skills and the technology infrastructure to carry out the project, or whether significant training and preparation will be required.

Some time may profitably be invested in ascertaining the copyright status of the material which is to be digitised. Failure to secure permission to digitise and to publish on the web can cause the failure of a digitisation project, despite any technical expertise and experience.

A technical pilot may also be considered, at the start of the project, in order to ensure that any anomalies or problems with the technical workflow are resolved before commencing the main project.

4.2.2 Guideline Title: The Reasons for the Project

Issue Definition

Each digitisation project has its own reason for being executed. Often, the reasons involve providing access over the Internet to cultural holdings which would otherwise be underused, or protecting fragile holdings from the wear and tear of hands-on access. In other cases, the projects are exercises in inter-body cooperation, and involve the establishment of portals, networks, etc.

The reasons for the project will have a profound effect on the criteria for selecting the material to be digitised. They will also affect the project management, the meta-data, the online publication (if any) of the project output, the quality control etc. 'Why' is the most important question to raise before starting a digitisation project.

Guideline Text

- The project must have concrete, explicit aims, and these aims must be documented.

- The aims of the project should be realistic, when compared with the resources available.

- All steps of the project should be validated against these aims, in order to ensure that work carried out in the project contributes towards the achievement of the guidelines.

- The project aims should document the value which the project will bring to the institutions involved in the project. If time and effort are to be invested in the project, the justification for the project, from an institutional point of view, must be clear.

Notes/Commentary

-

References

Online

- NOF-Digitise Technical Advisory Service Manual : http://www.ukoln.ac.uk/nof/support/manual/
- Arts and Humanities Data Service : http://www.ahds.ac.uk
- American Memory : http://lcweb2.loc.gov/ammem/ftpfiles.html
- Council on Library and Information Resources (CLIR) : http://www.clir.org/pubs/reports/reports.html

- Sun Microsystems Digital Toolkit : http://www.sun.com/products-n-solutions/edu/libraries/digitaltoolkit.html
- Guides to Quality in Visual Resource Imaging: http://www.rlg.org/visguides/ (esp. Guide 1 – planning).
- US National Digital Library Programme Project Planning Checklist : http://lcweb2.loc.gov/ammem/prjplan.html
- Planning Your Digitization Project : www.infopeople.org/training/past/2001/digitization/Agenda.pdf
- An Introduction to Digital Projects for Libraries, Museums and Archives, http://images.library.uiuc.edu/resources/introduction.htm

Nominated by Member States

- **France :** National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy** *: Diplomatico :* http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **Portugal** : Endovelliccus : www.ipa.min-cultura.pt
- **Portugal :** MatrizNet : http:// www.matriznet.ipmuseus.pt
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass
- **UK :** NOF-Digitise Technical Advisory Service Manual : http://www.ukoln.ac.uk/nof/support/manual/

4.2.3 Guideline Title: Human Resources

Issue Definition

Before a project can start, it is important that the personnel required to work on the project are available. Many cultural bodies do not have large corps of staff who have a great deal of free time to carry out digitisation projects, over and above their usual duties. Also, the knowledge required for digitisation projects may be different to from the skills necessary to carry out the tasks for the daily operation. Hardware and software solutions required for a digitisation project, therefore, need to be identified.

Guideline Text

- Ensure that sufficient staff is available to carry out the project.
- Assign staff to each task or work-package of the project plan.
- Identify training requirements, including information technology training and education in the handling of delicate artefacts and documents.
- Carry out, if possible, training by using the hardware and software solution which will be used during the project, before the project commences (vendors sometime offer technical solutions free of charge for training or relevant equipment can be hired short-time).
- Aim for a small core of skilled staff dedicated to the project, rather than a large group of 'occasional' staff.

Notes/Commentary

While the material presented in this guideline is common to all project management scenarios, it is worth repeating, particularly since there is possible risk to irreplaceable artefacts and documents if the resourcing is not properly handled.

References

Online

- Canadian Heritage Information Network: Planning your digitisation project: http://www.chin.gc.ca/English/Digital_Content/Small_Museum/planning.html
- Colorado Digitisation Programme: Questions to Ask: http://www.cdpheritage.org/resource/ introduction/questions.html
- Library of Congress, National Digital Library Program NDLP Project Planning Checklist at http://lcweb2.loc.gov/ammem/prjplan.html

- NOF-Digitise Technical Advisory Service Manual: http://www.ukoln.ac.uk/nof/support/manual/ has sections on resourcing, job specification, recruitment, etc.

Nominated by Member States

- **Denmark :** "The soldier in the Backyard – an interactive children's story on the Internet" **:** http://www.soldatenibaghaven.dk (especially multi-partner projects)
- **Spain** : Virtual Sites Re-creation : www.patrimonionacional.es (especially multi-partner projects)
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **France :**National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Ireland** : ACTIVATE : http://www.activate.ie (includes methodology guides and templates)

4.2.4 Guideline Title: Research

Issue Definition

Regardless of the scope of any particular project, it can be assumed that similar projects have been carried out in the past. There is a strong likelihood that information about such projects will be available on the Internet, or else published in appropriate journals, etc.

Researching the area as part of the project planning process can help to identify candidate hardware and software solutions, to plan workflow and process, and to avoid issues and obstacles which have been experienced by other projects.

Guideline Text

- As early as possible in the planning process, carry out research into any other projects which are addressing similar issues to the project being planned. This handbook provides a starting point; however the amount of material available on the Internet is the largest and most comprehensive resource.

- Research helps avoiding mistakes. It can also put the project team in contact with others who have completed similar projects, and give the opportunity to learn from their experiences.

- Having carried out research adds credibility and value to the output of any project. Assurance that your project has not been carried out in a vacuum, by taking into account the work of others, enhances the results of your project.

Notes/Commentary

Many cultural digitisation projects are funded with public funds, and have a requirement to publish their findings and their reports. Such publication is almost always on the Internet, as well as using other appropriate media.

Project teams are usually very happy to share their experiences and their results – this adds value to their work.

References

Nominated by Member States

The following references include some of the nominated projects who may be in a position to assist in the targeting of background and pre-project research.

- **Belgium** : Culture net Flanders

- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **France :**National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **Sweden**: The Oxenstierna Project.: http://www.ra.se/ra/Oxenstierna/oxenstierna1.html

4.2.5 Guideline Title: Risks

Issue Definition

At the start of any project, plans have to be made to guarantee a successful outcome. However, the goal is not to eliminate all risks but to prepare for them by creating a project framework which responds to the unforeseen in a resourceful and effective way. The aim is to create a project with staff and procedures that can accommodate changes. Therefore, all project planning need to have a risk analysis.

Guideline Text

- Distribution of digitised images over the Internet is a form of publication, and is by this reason covered by laws of copyright and intellectual property right (IPR). Examples of questions in an risk analysis are:
    - What could be the consequences of using the material without specific permission?
    - Has attempts been made to find the rights holder?
    - If copyright infringement does occur, what would the impact be on the project?

- For public information the legal value of the information is an important issue. What steps have been made to guarantee that a digitised source material is not corrupt and has actually been produced by an authorised institution?

- The authenticity must also be guaranteed. What actions have be undertaken to maintain the image files, and what tools have been used?

- Financing the project could be a problem and, therefore, a potential risk for the possibilities to reach the goals of the project.

- A key question is the level of skill in the project. Is it possible to hire new highly skilled and experienced persons? If not, will it have an effect on the work plan of the project?

Notes/Commentary

None

References

Online

- The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural heritage Material**:**
  http://www.nyu.edu/its/humanities/ninchguide/
- Gregory W. Lawrence, William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney, Risk Management of Digital Information: A File Format Investigation (CLIR 2000).
  http://www.clir.org/pubs/abstract/pub93abst.html

4.3 Selecting Source Material for Digitisation

4.3.1 Introduction

The selection of the material to be digitised is an important decision for any digitisation project. Typically, the ideal choice is to digitise all the material in a collection or holding; however, this is rarely feasible, so choices must be made. The criteria for selection will differ, depending on the goals of the digitisation project; an online resource for schools may choose to digitise material in line with a syllabus, while a museum may digitise its best-known holdings in order to stimulate visitor numbers or it's most fragile artifacts in order to minimize demand for 'hands-on' examination. These are of course not the only issues to be addressed in the selection criteria – the reasons for choosing to digitise particular material will vary from project to proposal, as will the reasons for deciding not to digitise. Examples of other reasons include legal constraints, institutional policies, technical difficulty of digitisation, already-extant digital copy, etc.

4.3.2 Guideline Title: Establish Selection Criteria

Issue Definition

When planning a digitisation project, the choice of which material to digitise is critical. The criteria for selection will depend on the goals of the project, as well as on technical and financial constraints, copyright and IPR issues, and the activity of other projects in the area.

Guideline Text

- It is essential to establish criteria for the selection of material to be digitised. The selection criteria must reflect the goals of the overall project. At least the following criteria may be considered
    - Access to material which would otherwise be unavailable, or of limited availability
    - Wider and easier access to very popular material
    - Condition of the originals.
    - Preservation of delicate originals, by making digital versions available as an alternative
    - Project theme
    - Copyright and IPR
    - Availability of existing digital versions
    - Cost of digitisation
    - Appropriateness of the source material for online viewing

- The criteria for selection should be explicit and discussed with, and endorsed by, all relevant stakeholders, prior to selection or digitisation.

- The selection criteria should be fully documented, so that the reasons for any decisions to digitise or not to digitise are clear throughout the project.

Notes/Commentary

Most commonly, cultural bodies have a core of high-value, high-user-interest material which is, by default, included in any digitisation project which is meant to represent the institution.

A large proportion of all digitisation projects have online web publication as a goal. This means that the copyright and IPR issues which surround any material which may be digitised must be considered before selection.

References

Online

- RLG/NPO Guidelines and Selection Criteria **:** http://www.rlg.org/preserv/joint/selection.html
- Columbia University Libraries Selection Criteria For Digital Imaging : http://www.columbia.edu/cu/libraries/digital/criteria.html
- Selection Criteria for Digitiz ation Projects : www.wils.wisc.edu/events/dgtdev/present/maritime.doc
- Brown University Library Selection Criteria for Digitization : http://www.brown.edu/Facilities/University_Library/digproj/digcolls/selection.html
- Old Dominion University : Selection Criteria For Digitization **:** http://www.lib.odu.edu/services/dcenter/digselection.html

Nominated by Member States

- **Denmark** : Kongens Kunstkammer (Royal Chamber of Art) **:** http***://***www.kunstkammer.dk
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** Diplomatico ***:*** http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy :** Mediceo avanti il Principato on line***:*** http://www.archiviodistato.firenze.it/Map/
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **Sweden** : The Oxenstierna Project : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.3.3 Guideline Title: Selection Against the Criteria

Issue Definition

Having established the criteria against which material is selected to be digitised, the actual selection process can take place. This guide suggests how to manage this process.

Guideline Text

- Each candidate for digitisation must be evaluated against the selection criteria. In case that any selection criterion is not met, this should be noted. In the event that this results in the rejection of important or critical objects, it may be necessary to review the selection criteria. Should this occur, the new criteria should be noted.

- Once an object has been selected for digitisation, its details should be entered into the digitisation management knowledge base (se chapter 4.11.2).

Notes/Commentary

At this stage, the project is engaging with each of the items to be digitised, for the first time. This is the optimum opportunity for the project to create a knowledge base of all the items in the scope of the project. Having such a knowledge base will support the management of the project, and help to ensure that, for example, the appropriate expert knowledge is acquired for handling rare artifacts, as well as more mundane issues such as the location of originals.

References

Online

- RLG/NPO Guidelines and Selection Criteria **:** http://www.rlg.org/preserv/joint/selection.html
- Columbia University Libraries Selection Criteria For Digital Imaging : http://www.columbia.edu/cu/libraries/digital/criteria.html
- Selection Criteria for Digitization Projects : www.wils.wisc.edu/events/dgtdev/present/maritime.doc
- Brown University Library Selection Criteria for Digitization : http://www.brown.edu/Facilities/University_Library/digproj/digcolls/selection.html
- Old Dominion University : Selection Criteria For Digitization **:** http://www.lib.odu.edu/services/dcenter/digselection.html
- UK : Library and Information Commission 'Full Disclosure' report at http://www.ukoln.ac.uk/services/lic/fulldisclosure/

- French directory of digitised collections: http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f_02.htm

Nominated by Member States

- **Denmark** : Kongens Kunstkammer (Royal Chamber of Art) **:** http**://**www.kunstkammer.dk
- **Italy :** Diplomatico ***:*** http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy*:*** Mediceo avanti il Principato on line***:*** http://www.archiviodistato.firenze.it/Map/
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html

4.4 Preparation for Digitisation

4.4.1 Introduction

An appropriate environment and hardware/software system must be in place before digitisation can begin. The elements of such an environment include hardware for the digitisation process itself (e.g. scanners, digital cameras, copy stands, other hardware), a computing infrastructure to which the hardware is connected, software for image capture and processing software, as well as software for metadata and quality control. The working environment should be appropriate to the material being digitised, paying special attention, for example, to light, humidity, vibration, disturbance, movement of the originals, etc.

4.4.2 Guideline Title: Hardware

Issue Definition

The appropriate technical equipment must be in place for the digitisation to go ahead. Typically this will consist of digital image capture equipment (digital cameras, scanners for books, documents or microfilm, audio and video hardware, if appropriate) connected to an appropriate computing platform (computer, operative system, network, etc**).** Two different digitisation methods, using different hardware, can be distinguished: scanning and the use of digital cameras.

Guideline Text

- Appropriate hardware must be installed and its quality and functionality controlled before digitisation begins.

- Relevant test targets should be used for the evaluation of digital image capture devices.

- No source material should be present until the hardware environment has been fully established and tested with non-sensitive materials.

- Most digitisation projects will require a flatbed scanner, for material which is not harmed by being pressed flat against a hard surface (e.g. unbound printed material and manuscripts).

- The largest possible scanner should be acquired by the project. The folding or mosaiced scanning of materials should be avoided. The project should bear in mind that the transportation of large (e.g. A0) scanners is not trivial.

- Usually, a flatbed scanner should only be used where the material is already flat, and will not be damaged by being held against a hard, flat surface. A scanner with a book cradle may be appropriate for many bound articles, up to the appropriate size limits. Most digitisation projects will require a digital camera, for capture of material which cannot be flattened or held on a scanner book cradle.

- If a scanner is used, it should ideally be at least as large as the item to be scanned.

- Image capture (by scanner or digital camera) should be carried out at the highest reasonable resolution. This will often result in very large master files; smaller files can be extracted from the master, for purposes such as

web delivery. However, a higher-quality image can never be derived from a lower-quality image.

- The definition of a 'reasonable' resolution will depend on the nature of the material being scanned, and on the uses to which the scanned image will be put. For example, if the scanned images are only ever to be used as thumbnails, this can allow scanning at a low resolution. Equally, the resolution must capture the most significant details of the item – if scanning at a high resolution yields no more information than at a lower resolution, the high resolution scanning is difficult to justify.

- Image capture should create a file format which is loss-less, i.e. not compressed. Typically, the Tagged Image File Format (TIFF) is used.

- If a digital camera is used, a project shall choose the most powerful and flexible one which can be afforded. The limitations of the digitisation hardware cannot be overcome by any subsequent processing. It should be noted that 'digital zoom' does not provide a better quality picture; it merely displays less pixels per unit of view. In order to capture detail, three parameters are most important – the number of pixels in the image, the bit-depth, and the quality of the optical lens being used.

- It is important to have appropriate stands for holding material while it is being digitised.

- A digital camera with a dedicated copy stand should be used. The camera should be tripod-mounted, and have supplementary lighting, filters, etc, as appropriate. Consultation with an experienced digital photographer with a background in similar projects is advised, if at all possible, before setting up the hardware environment

- The photographic plane and the plane of the material being digitised must be exactly parallel, if the image is not to be distorted.

- Appropriate lighting must be part of the photographic set-up when using a digital camera; it is very rare for ambient light to be sufficient.

- Suitable filters should be used in order to reduce colour distortion.

- A computer with significant storage should be connected to the devices. This computer should be backed up very regularly – this requirement reflects the high costs in time, technology and possible wear on the originals, of the digitisation process.

- If an item must be scanned in multiple parts, an overlap of several centimetres should be provided, in order to ensure that there are no gaps between the parts. The same settings, light, etc should be used for all parts, in order to avoid any 'patchwork' effect.

Notes/Commentary

The hardware used is a major constraint on the quality of the end result of any digitisation project. Unless the project is digitising only flat materials which can be scanned without damage to bindings, frames or the source material itself, the use of a digital camera will be important. While an analogue camera can be used, and the slides or prints scanned, the advantages in terms of time, effort and quality of a high-specification digital camera are many.

If the project has a limited life-span, renting hardware may be appropriate. Another alternative is the use of external agencies to carry out the digitisation on behalf of cultural bodies involved in the project.

References

Online

- The comprehensive TASI site has a section on hardware and software for digitisation projects at http://www.tasi.ac.uk/advice/creating/hwandsw.html
- The University of Arizona has a substantial amount of online guidance, including hardware and software, at http://www.dlapr.lib.az.us/digital/dg_a3.html
- The Colorado Digitisation Program includes hardware in its list of guidelines at http://www.cdpheritage.org/resource/scanning/std_scanning.htm
- Harvard University publishes notes on the choice of appropriate digitisation hardware at http://preserve.harvard.edu/resources/imagingsystems.html
- The Preservation Administration Discussion Group covers a range of topics in the digitisation area. It can be found at http://palimpsest.stanford.edu/byform/mailing-lists/padg/
- Canadian Heritage provides notes on hardware at http://www.chin.gc.ca/English/Digital_Content/Capture_Collections/capturing_images.html
- Base Enluminures (Manuscript illuminations) : http://www.enluminures.culture.fr

Nominated by Member States

- **Austria**: www.bildarchiv.at. (special digital photography setup)
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy** *:* Diplomatico *:* http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy :** Mediceo avanti il Principato on line: http://www.archiviodistato.firenze.it/Map/
- **Potugal** : Endovelliccus : www.ipa.min-cultura.pt
- **Portugal :** MatrizNet : http:// www.matriznet.ipmuseus.pt
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.4.3 Guideline Title: Software

Issue Definition

Having created a digital version of the object, the resulting file is likely to require processing before it can be used. Colour may need correction; extraneous detail may need to be cropped (removed) from the edges of the image, etc. Also, the master files are typically very large, so a smaller file in a compressed format will often be needed (e.g. as a thumbnail image, or for web delivery).

Guideline Text

- When the scanner or the digital camera is turned on a calibration routine should start automatically.

- Suitable image processing software will be needed to utilise the master files for whatever the purpose of the digitisation project may be. While digitisation hardware will typically be provided with some software included, this is usually not of sufficient power and flexibility for many projects.

- The requirements on the software depend on the aims of the project. It is worthwhile to note that, once the master files are not modified in any way; various different types of software can be used to process them. However, the cost in time and effort may be significant, and will usually overshadow the cost of a more powerful software package.

- The project should acquire the most appropriate and powerful software package which it can afford.

- As an absolute minimum, the software must be capable of:
  - opening very large image files
  - modifying the resolution and the colour depth
  - saving multiple different versions, in different file sizes.
  - selecting and copying a part of the image, and saving this as another file.
  - exporting images in different file formats, including the web standards
    JPEG and GIF.

  Several free software packages provide this level of functionality; however investing in a commercial product is likely to pay dividends in time, effort, documentation and technical support.

- In the event that the digitisation project has a OCR component, the choice of software is also critical. Any OCR exercise has a certain amount of

manual editing and correction; the manner in which this is supported by the software product in use can have a significant effect on the time and effort required by the project. Better OCR packages may enable review and editing on a single screen, suggest possible corrections for mis-read words, support the use of multiple text columns (e.g. newspaper layout), etc.

The evaluation of multiple OCR packages is likely to be worthwhile, if the project exceeds, for example, one person-year in size

Notes/Commentary

The right software will save a digitisation project a large amount of time and effort. If the project is of significant duration (e.g. more than two persons for more than six months), evaluation of several software packages may be worthwhile, in order to establish the best match for the requirements of the project.

References

Online

- The comprehensive TASI site has a section on hardware and software for digitisation projects at http://www.tasi.ac.uk/advice/creating/hwandsw.html
- The University of Arizona has a substantial amount of online guidance, including hardware and software, at http://www.dlapr.lib.az.us/digital/dg_a3.html
- The Colorado Digitisation Program includes hardware in its list of guidelines at http://www.cdpheritage.org/resource/scanning/std_scanning.htm
- Gallica Digital Library: http://gallica.bnf.fr

Nominated by Member States

- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy *:*** Diplomatico ***:*** http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Portugal:** MatrizNet : http:// www.matriznet.ipmuseus.pt (Matriz is a museum management software solution).
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass (project includes significant software development)

4.4.4 Guideline Title: Environment

Issue Definition

Many rare or delicate materials require a particular environment. It is critical to any digitisation project that the digitisation process has the least negative effect on the source materials. An appropriate digitisation environment is important to many digitisation projects.

Guideline Text

- The environment in which digitisation takes place is of considerable importance.

- Expert opinions should be sought in order to ensure that all aspects of handing of original material are addressed as well as possible. These include the environment for digitisation.

- The area used for digitisation should be dedicated to the digitisation project for the duration of the project. Excessive movement, rearrangement etc of the workspace can lead to damage, loss or other negative effects on the source materials, as well as to loss of time by the project.

- If the source materials have particular requirements in terms of light, humidity, etc, then these should be replicated as closely as possible in the digitisation environment. For certain materials, such as leather documents, a short-term increase in humidity may assist in relaxing the materials prior to flattening for photography or scanning.

- In almost all cases, direct exposure to bright light (e.g. sunlight) for extended periods is not recommended. Smoking, eating and drinking in the vicinity of the items should of course not be permitted – keep coffee away from the work area!

Notes/Commentary

Depending on the size and budget of the project, a dedicated digitisation environment may not be feasible. However, the aims outlined here, to minimize movement, disruption and handling of the materials, should be kept in mind.

As with the handling of heritage material, no references should be taken as a substitute for discussion with those whose responsibility includes the care of the material.

References

Online

- The Australian Consortium for Heritage Collections and their Environment publishes guidelines at amol.org.au/craft/publications/hcc/
- environment_guide/environ_1.pdf (hosted by Australian Museum Online – AMOL)
- AMOL also publishes a FAQ for conservation of artworks; although focused on Australian concerts, it includes much of value, at http://www.amonline.net.au/materials_conservation/faq/
- The University of Melbourne publish a useful guide to conservation, including the handling of fragile materials, at http://home.vicnet.net.au/~conserv/prepast1.htm

Nominated by Member States

- **Germany**: Workflow and tools for providing access to larger quantities of archival material: http://www.lad-bw.de
- **Italy:** DADDI**:** http://www.uffizi.firenze.it/Dta/daddi-eng.html

4.5 Handling of Originals

4.5.1 Introduction

This section considers how a digitisation project should treat the material which is being digitised. In many cases, the source material is rare or valuable; the negative effects of digitisation on the source material must be minimized.

In every case, it must be emphasized that the specialist knowledge of the individuals who are responsible for the source material on a day to day basis will be valuable to the project team.

4.5.2 Guideline Title: Moving and Manipulating Original Material

Issue Definition

In many cases, the material to be digitised is of particular sensitivity or fragility. Replacing hands-on access with online access is often an important reason for digitisation projects in the first place. It is critical that any digitisation project takes steps to ensure that no damage is done to the original material during the digitisation process. These steps may range from the use of the correct hardware to the establishment of a suitable micro-climate or the movement of the digitisation centre of operations to the location of the material, rather than vice versa.

Guideline Text

- Consult the person usually responsible for the source material, before moving or handling it. Include any information on appropriate handling, in the digitisation project knowledge base (see chapter 4.11.2)

- Be prepared to be flexible – an inconvenience to the digitisation project can be overcome, while damage to a unique artifact may be irretrievable.

- If necessary, bring the digitisation equipment (e.g. digital camera) to the source item, rather than transporting the item itself.

- Avoid unbinding of bounded books and records. Use instead of a flatbed scanner a scanner with a book cradle or a digital camera.

- Always remove staples, paper clips, and other fasteners; they can damage both the digitisation devices and the source material.

  - Expert advice (e.g. from the curator of the item to be digitised) should be sought before any handling of the original.

  - This advice should be sought prior to digitisation, ideally at the time that the article is selected for digitisation. The advice should be recorded in the digitisation management knowledge base, and consulted before movement or digitisation of the article. If necessary, the expert should be briefed on the capabilities of each possible hardware solution.

Notes/Commentary

While much of this material is quite obvious, it is important to establish and maintain a discipline while handling the source material.

References

Online

- The Australian Consortium for Heritage Collections and their Environment publishes guidelines at amol.org.au/craft/publications/hcc/environment_guide/environ_1.pdf (hosted by Australian Museum Online – AMOL)
- AMOL also publishes a FAQ for conservation of artworks; although focused on Australian concerts, it includes much of value, at http://www.amonline.net.au/materials_conservation/faq/
- The University of Melbourne publish a useful guide to conservation, including the handling of fragile materials, at http://home.vicnet.net.au/~conserv/prepast1.htm
- The Preservation Administration Discussion Group covers a range of topics in the digitisation area. It can be found at http://palimpsest.stanford.edu/byform/mailing-lists/padg/
- Base Enluminures (Manuscript illuminations): http://www.enluminures.culture.fr/

Nominated by Member States

- **France:** National digitisation programme – annual project calls: http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy :** Mediceo avanti il Principato on line**:** http://www.archiviodistato.firenze.it/Map/
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.6 The Digitisation Process

4.6.1 Introduction

This chapter provides some practical guidelines for the actual digitisation process. The technical solutions for digital capture can differ. Scanners, digital cameras or software applications for optical character recognition (OCR) are areas covered in some detail, as being most relevant to the largest number of projects. The digitisation of transparent originals like microfilm is not considered.

4.6.2 Guideline Title: Using Scanners

Issue Definition

Flatbed scanners are a very common digitisation tool. The most common A4 and A3 models are relatively cheap, require limited skills to use, and can manage a fast throughput of material, once a workflow has been established. Larger models (up to A0) of flatbed scanners and scanners equipped with book cradles are very expensive and thus require long-term projects/programs, high-volume digitisation., or oversized source material.

Guideline Text

- Only scan material on a flatbed scanner which will not be damaged by being pressed flat onto a hard surface. Consult the experts, if in doubt.

- Ensure that the glass scanning plate is completely clean at all times. This both leads to better image quality and also protects the source material from soiling.

- If possible, scan only items which fit on the flatbed scanner or the scanner equipped with a book cradle in one piece.

- If it is necessary to scan an item in multiple parts, ensure that there is sufficient overlap to allow the image to be reassembled, during post-processing (by using mosaicing software).

- Test the scanner, and its output, on non-sensitive material before beginning to scan original source material. Train users with the same non-sensitive material.

- Establish a file-naming convention for the files produced by the scanner, for example by using the existing cataloguing system or giving them meaningful names. The file name should allow mapping between the file and the source item.

- In order to maximise the portability of files across computer platforms, a file name with a maximum of eight characters, followed by an extension of at most three characters, should be adhered to.

- Before establishing workflow or work-batching process, carry out some end-to-end scanning and image processing, in order to ensure that the end result of the workflow will be what is anticipated.

- Scan at the highest resolution and bit-depth that is feasible given the reasons for the project, the limitations of the scanner, the conditions for data storage, and the attributes of the source material (see chapter 4.4.2).

- Scan with the maximum appropriate colour depth, given the same limitations.

o Back up the hard disk where the data is stored, on a daily basis.

Quality control of the digital image and of metadata is important – at scanning time is the most convenient time to address any quality issues. The following points may be borne in mind:

- Establish minimum resolution and colour parameters (mainly the spatial resolution and the bit-depth) for groups of items to be scanned.

- Examine the scanned output on screen, on paper and in any other format that you expect it to be used for (e.g. on a mobile device).

- Ensure that the screens (monitors) being used are reliably calibrated. Avoid having other material on and around the screen, which may affect the perception of the item

- Master images must be created with visible scaled rulers, and colour or greyscale images must include also a standardised colour/greyscale reference target.

Notes/Commentary

Scanning is in itself a relatively simple operation. However, in order to increase efficiency and minimize errors, having a workflow system in place will be worthwhile.

Scanning of oversize items, or very high quality scanning, takes a significant investment of time and effort per item. This can be reduced by using hardware appropriate to the item (e.g. a larger scanner, a book cradle); in the event that large hardware resources are not available, allow plenty of time. Training on oversize or irregular materials should not be neglected.

References

Online

- A good guide to workflow and process management is on the TASI site at http://www.tasi.ac.uk/advice/managing/jidi_workflow.html

- A user-friendly site on the scanning process is provided at www.scantips.com
- A short overview on how to use a scanner is provided at http://www.aarp.org/computers-howto/Articles/a2002-07-16-scan

- There are countless scanning pages on the Internet – use Google or a similar search engine to browse them.

Nominated by Member States

Most or all of the projects nominated by Member States as examples of good practice have used scanners at some stage. Some examples are given here

- **Germany** – Digital Conversion Forms : http://www.lad-bw.de
- **Germany** - Workflow and tools for providing access to larger quantities of archival material http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi The Nordic library: http://tiden.kb.se
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy** *:* Diplomatico*:* http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy :** Edit16 **:** http://edit16.iccu.sbn.it
- **Italy :** www.pinacotecabologna.it
- **Italy** *: Mediceo avanti il Principato on line:* http://www.archiviodistato.firenze.it/Map/
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **Potugal** : Endovelliccus : www.ipa.min-cultura.pt
- **Portugal :** MatrizNet : http:// www.matriznet.ipmuseus.pt
- **Portugal:** BND: Biblioteca Nacional Digital (the National Digital Library Initiative): http://bnd.bn.pt
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.6.3 Guideline Title: Using Digital Cameras

Issue Definition

The use of digital cameras is becoming increasingly common in digitisation projects. This reflects their flexibility in terms of being able to digitise non-flat objects, such as bound books, folded or wrinkled manuscripts, and 3D objects. However, a scanner equipped with a cradle is normally to prefer when digitising bounded books and over-sized material like maps and drawings.

Guideline Text

- Consider renting a high-quality camera, if the scope of the project is limited.

- Put up the digital camera on a motorized carriage on a column and place the items to be digitised on a steady copy stand board with specially tailored lights.

- Organise training from a specialist digital photographer –the difference in quality between pictures taken by an amateur and the same photos taken by a specialist can be striking.

- Ensure that backgrounds will show the item clearly.

- Avoid changing the light conditions between shots, and between images of different parts or sides of an item – this can lead to erroneous impressions of colour variation.

- Use apochromatic lenses and appropriate lens filters to combat colour misregistration and image distortion.

Notes/Commentary

The increasing use of digital cameras in digitisation projects reflects their availability as a mainstream consumer product, and the resulting decrease in price. However, there remains a significant difference, in both price and quality, between specialist digital cameras and mass-produced low-end consumer products.

References

Online

- A good guide to Workflow and process management is on the tasi site at http://www.tasi.ac.uk/advice/managing/jidi_workflow.html

- A guide on the basics of using a digital camera is provided at http://www.pcphotoreview.com/basic3040crx.aspx
- The TASI page on hardware and software may be useful – see http://www.tasi.ac.uk/advice/creating/hwandsw.html
- NCSU provide a guide to the practical use of digital camera at http://www.ncsu.edu/sciencejunction/route/usetech/digitalcamera/
- Base Mémoire: http://www.inventaire.culture.gouv.fr/documentation/memoire/pres.htm

Nominated by Member States

Most of the projects nominated by Member States as examples of good practice (see Minerva website www.minervaeurope.org) will have used a digital camera extensively. Of particular interest is the Italian Daddi project.

- **Germany** – Digital Conversion Forms : http://www.lad-bw.de
- **Germany** - Workflow and tools for providing access to larger quantities of archival material : http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy** *:* Diplomatico*:* http://www.archiviodistato.firenze.it/progetti/attivite.htm
- **Italy :** Edit16 **:** http://edit16.iccu.sbn.it
- **Italy :** www.pinacotecabologna.it
- **Italy** *:* Mediceo avanti il Principato on line*:* http://www.archiviodistato.firenze.it/Map/
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- Italy : Virtual Archaeological Tours around the Lost Cities : http://www.archeologia.beniculturali.it (especially Virtual Reality)
- **Potugal** : Endovelliccus : www.ipa.min-cultura.pt
- **Portugal :** MatrizNet : http:// www.matriznet.ipmuseus.pt

Portugal: BND: Biblioteca Nacional Digital (the National Digital Library Initiative): http://bnd.bn.pt

- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.6.4 Guideline Title: Software Applications for Optical Character Recognition (OCR)

Issue Definition

Many digitisation projects involve the digitisation of printed documents, such as a books and newspapers. This occurs most often (though not exclusively) in tandem with the use of scanners. The use of OCR software is a popular way to extract the information from such scanned information, and to open opportunities for processing the information. OCR software recognises the letters and numbers which make up the scanned image (bit mapped image file), and exports them as ASCII text files, rather than as image files. This enables searching, indexing, format conversion, and other data processing operations to be carried out.

Guideline Text

- Evaluate multiple OCR software offerings before selecting a particular product. While OCR software is often included with the sale of a scanner, more powerful software is typically sold separately.

- A major element of any OCR project is the identification and manual editing of mistakes, ambiguities and locations where the text could not be processed. An OCR package which provides a friendly user interface for carrying out this task can save considerable time and effort.

- OCR works best with documents which are in good condition – folding, wrinkling and discoloration of the source material will increase the number of errors and faults in the OCR process. Pre-treatment, where possible, of the source material should be carried out to avoid this.

- The use of image processing software, to remove discoloration and improve contrast, before the use of OCR software, should be considered for material which is not in perfect condition.

- The availability (or not) of dictionaries in the language of the source material, as part of the OCR package, should be verified.

Notes/Commentary

Relevant products in this market include
- OmniPage
- TextBridge and
- Adobe Capture.
- Abby FineReader

The last of these has excellent editing and fault resolution functionality.

References

Online

- The University of Maryland hosts a major OCR resource at http://documents.cfar.umd.edu**/**
- A brief OCR overview is provided by computer world magazine at http://www.computerworld.com/softwaretopics/software/apps/story/0,10801,73023,00.html
- A worthwhile technical report on OCR is provided by the University of New York, Buffalo, at http://www.cedar.buffalo.edu/Publications/TechReps/OCR/ocr.html
- A report on OCR, newspapers and microfilm is provided by IFLA at http://www.ifla.org/VII/s39/broch/microfilming.htm

Nominated by Member States

- **Austria** : Digital Image Archive www.bildarchiv.at (automated.indexing)
- **Germany** – Digital Conversion Forms : http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **Portugal**: BND: Biblioteca Nacional Digital (the National Digital Library Initiative): http://bnd.bn.pt

4.7 Preservation of Digital Master Material

4.7.1 Introduction

In the longer term, it is an important goal of any digitisation project to protect and keep accessible the data which it has created. This involves dealing with the inevitable obsolescence of digital file formats and various types of computer storage media.

Preserving the digital master material and corresponding metadata helps to avoid having to re-digitise any items, thus protecting the fragile source material and avoiding repetition of the labour-intensive digitisation process including generating metadata.

4.7.2 Guideline Title: File formats

Issue Definition

The digital output of the digitisation process is usually a master file in uncompressed TIFF format with some metadata embedded (see chapter 4.8 2). The file format as well as the compression used will have a major impact on the usability of the digitisation output. At this time, issues such as file format, standards file size, network transmission time, and different kind of outputs (monitor or printer) need to be taken into account.

Guideline Text

- Before deciding on a file format, take into account the relevant standards, the established global user base and the degree to which file formats are supported by software in use by your organisation and your target audience. The size of the global user base is a good indicator of the future, ongoing, support for a particular file format. It also indicates the likelihood of sustainable migration paths, when file formats change.

- The default digitisation output file for digital images is Tagged Image File Format (TIFF). Unless your project has a clear, justified reason for using some other file format, digitisation output, and so master files, should use this format

- The output file will typically be quite large. It is common to have a large master file, which is stored locally but not transmitted over the Internet. From this master file, smaller versions can be created using image processing software, either in TIFF, or more commonly in a delivery format such as JPEG 2000, PNG or GIF (see the section on image standards, later in this document).

- More information on file formats is provided in the survey of standards provided later in this document (see chapter 5).

- Regardless of how attractive a proprietary or national format may appear to be from a technical standpoint, it is important to bear in mind that failure to use standard formats and media will act as a major obstacle to international exchange of raster image files and corresponding embedded metadata as well as the creation of networked resources.

Notes/Commentary

File format choice must be governed by the imperative to create the highest quality digitisation output, and by the availability of migration paths for future preservation of the digital master. The role of standards in this area is very great.

References

Online

- The AHDS provides a directory of material on the preservation of digital content at http://www.pads.ahds.ac.uk:81/padsProjectLinksDirectory/PreservationDigitalMaterial
- The Australian PADI initiative hosts a huge range of information on digital preservation, at http://www.nla.gov.au/padi/, particularly at http://www.nla.gov.au/padi/topics/44.html
- Reference Model for an Open Archival Information System. http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html
- Gregory W. Lawrence, William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney, Risk Management of Digital Information: A File Format Investigation (CLIR 2000). http://www.clir.org/pubs/abstract/pub93abst.html
- The Masters of European Comics (French National Library Exhibition): http://www.expositions.bnf.fr/bd

Nominated by Member States

- **Germany**– Digital Conversion Forms : http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK** : NOF-digi technical standards http://www.peoplesnetwork.gov.uk/nof/technicalstandards.html
- **Portugal:** BND: Biblioteca Nacional Digital (the National Digital Library Initiative): http://bnd.bn.pt

4.7.3 Guideline Title: Media Choices

Issue Definition

The issue of media choice is an important one for projects which wish to maintain their digital collections over a several-year period. Important projects such as the UK Domes day book initiative have been lost due to media obsolescence.

Guideline Text

- The output of the digitisation project will be held on server machines, including those which serve digital content to Internet users. However, these machines need to be backed up. Also, if a server is not dedicated to a digitisation project, the digital content should be stored on removable media, separate to other data on the server.

- All master files (including metadata) should normally be backuped on two kinds of media separately stored.

- Currently (early 2003), the use of CD-Rs as a common backup medium is in the process of being replaced by the use of DVDs. DVDs offer significantly larger storage, and the hardware needed to read them is becoming ubiquitous on new PCs and laptops. DVD writers remain more expensive, but are already well within the means of all but the smallest projects.

- However, DVDs are not expected to replace magnetic tape media like Digital Linear Tape (DLT) as the storage medium of choice for backup of computer storage, in the near future. Both of these technologies should be seriously considered as candidates for preservation of digital content.

- Regardless of the choice of medium, it must be borne in mind that the medium will become obsolete in near to mid-term future. Within five years, migration to new storage media is likely to be a necessity.

Notes/Commentary

The rapid change of media layouts, driven primarily by the consumer electronics industry, has had major effects on digitisation projects in the past.

However, the increasing trend to store data 'on the Internet' on large server machines, and as data on mobile hard drive units, facilitates the migration of data from place to place and from medium to medium. Once servers are backed up and migrated to new servers over time, the dependence on removable media as the only record of a digitisation process can be expected to decrease.

In the meantime, the issue of media selection is still an important one. There is no indication that the limits of compressed, small-footprint digital storage are being reached.

References

Online

- The AHDS provides a directory of material on the preservation of digital content at http://www.pads.ahds.ac.uk:81/padsProjectLinksDirectory/PreservationDigitalMaterial
- The Australian PADI initiative hosts a huge range of information on digital preservation, at http://www.nla.gov.au/padi/, particularly at http://www.nla.gov.au/padi/topics/44.html
- Reference Model for an Open Archival Information System. http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html
- Gregory W. Lawrence, William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney, Risk Management of Digital Information: A File Format Investigation (CLIR 2000). http://www.clir.org/pubs/abstract/pub93abst.html

Nominated by Member States

- **Germany** – Digital Conversion Forms : http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **Portugal:** BND: Biblioteca Nacional Digital (the National Digital Library Initiative): http://bnd.bn.pt

4.7.4 Guideline Title: Migration Strategies

Issue Definition

As noted above, the choice of file format and storage medium must take into account the feasibility of moving data to a new file format and/or a different storage medium, in the foreseeable future.

Guideline Text

- Examine the relevant standards for file formats and storage medium, as noted in the previous two guidelines. Compliance with standards is a reasonable indicator that a particular format or medium will have some support into the future.

- Proprietary file formats and non-standard media formatting should be adopted only with great care.

- Migration from one format to another should avoid migrating from a lossless file format (e.g. TIFF in the image domain) to a lossy one (e.g. JPEG), for master digital material. Once information is lost, it cannot be replaced.

- Bear in mind that any choice of file format and/or storage medium will become obsolete in the foreseeable future (possibly less than five years, probably less than ten years).

- The size of the market for storage media provides an indication of how likely it is that migration from one medium to a new one will be feasible, as the medium becomes obsolete.

- Having created the digitised material, storage media (e.g. CD-R, DVD) should be refreshed periodically (once every two to three years), to combat data loss. This involves copying all data to new media.

- The status of digitised material, including when it was last refreshed, should be recorded in an appropriate log.

- Copies of digitised material should be stored in multiple locations whenever feasible, to reduce the risk of catastrophic data loss in the event of fire, etc.

Notes/Commentary

None

References

Online

- The AHDS provides a directory of material on the preservation of digital content at http://www.pads.ahds.ac.uk:81/padsProjectLinksDirectory/PreservationDigitalMaterial
- The Australian PADI initiative hosts a huge range of information on digital preservation, at http://www.nla.gov.au/padi/, particularly at http://www.nla.gov.au/padi/topics/44.html

Nominated by Member States

- **Germany** – Digital Conversion Forms : http://www.lad-bw.de
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **Italy :** "I dipinti della Galleria Spada": no web site
- **UK :** Digital Preservation Workbook : http://www.jisc.ac.uk/dner/preservation/workbook/

4.8 Meta-data

4.8.1 Introduction

The area of meta-data is one of the most actively researched and dynamic in the whole digitisation area, as well as in areas such as information retrieval, web searching, data exchange, enterprise application integration, etc.

The selected meta-data model is of particular importance as it influence the choice of attributes to describe an object. Related to this is the choice of a standard model, as will be described in the chapter on standards (see chapter 5).

4.8.2 Guideline Title: The Scope of Meta-data Used For Object

Description

Issue Definition

Before selecting a metadata model for a digitisation project, the material to be described with the metadata should be reviewed. This will help to identify existing meta-data models, as well as to pinpoint any omissions or gaps between what is covered by an existing meta-data model and the important meta-data for your project.

Guideline Text

- The use of appropriate metadata is very important for enabling search and retrieval of material from digital collections. This is even more the case when searching across multiple collections, stored in different locations, is the overall objective (logical union catalogues, virtual combined museums, etc.).

- There exist already many meta-data models. Therefore, each project has to choose as meta-data model based on its own goals. It is advisable to avoid creating a new one, unless the requirements of your project are badly underserved by all existing standards.

- Time spent modelling the important characteristics of the material being digitised, and identifying its key attributes and descriptors will be well invested. Such a model can then be compared with the scope and features of existing meta-data models.

- Possible controlled vocabularies (e.g. to describe a location, or an artist) should be identified. Several such vocabularies already exist and can greatly increase the success of searches, etc. See the section on meta-data standards and controlled vocabularies, below, for details.

Notes/Commentary

Comments: The Making of America II project (Library of Congress) used three categories of meta data

- Descriptive – for description and identification of information
- Structural – for navigation and presentation
- Administrative – for management and processing
-

Of particular importance are the metadata models to be selected for a digitisation project -the choice of which set of attributes that will be used to characterize the

works and items to be digitised, the resulting images, the description of the undertaken processes, techniques and technology, the rights management, etc. The National Library of Australia has a powerful model for this.

The plethora of existing models and competing standards for meta-data has led to projects which focus purely on translating from one standard to another.

References

Online

- The TASI page on meta-data is at www.tasi.ac.uk/advice/delivering/metadata.html
- The Colorado guidelines for meta-data creation and entry are at http://coloradodigital.coalliance.org/glines.html
- PADI's meta-data page is at http://www.nla.gov.au/padi/topics/30.html
- An unusual approach to user-generated meta-data is used at www.gimp-savvy.com
- Cornell University tutorial is found at http:// www.library.cornell.edu/preservation/tutorial/metadata/metadata-o1.html
- Bibliographic structure
    - MARC21 http://www.loc.gov.marc/
    - MARC21: UNINMARC: http://wwwunimarc.net
- Structural metadata
    - METS: http://www.loc.gov/standards/mets/

Note: The National Library of Portugal is the actual host institution for the international UNIMAC Program. The service http://www.unimarc.net is not currently available, but will be released in October 2003.

-
- The Dublin Core is covered at www.dublincore.org
- The Encoded Archival Description (EAD) home page is at www.loc.gov/ead/

Nominated by Member States

- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **France :** National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Germany** – BAM Portal http://www.bam-portal.de/
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- ICONCLASS in Italian : www.iccd.beniculturali.it
- **Italy** : Information Network dei Beni Culturali : www.iccd.beniculturali.it
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it

- **Italy:** SBNonline **:** http://sbnonline.sbn.it
- **Portugal:** BND: Biblioteca National Digital (the National Digital Library Initiative): http://bnd.bn.pt
- **Portugal:** PORBASE: Base Nacional de Dados Bibliográficos (National Union Catalogue): http://www.potbase.or
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass

4.8.3 Guideline Title: Appropriate Meta-data Standards

Issue Definition

Certain important standards already exist for meta-data. In the bibliographic domain (and increasingly in non-library cultural domains), the Dublin Core standard is of great importance.

Guideline Text

- Review existing meta-data models and standards before creating your own.
- Creating a totally new meta-data model for cultural collections should be avoided.
- The meta-data work carried out by similar projects in the past is likely to be relevant to your project – meta-data models travel well between projects in the cultural area.
- Unless your project has good reason not to do so, the Dublin Core fields should be included in the meta-data model. While museums may find the CIMI model better fits their holdings, a common core set of attributes should be aimed for, which will enable cross-collection searching.
- If a proprietary meta-data model is to be used, a mapping from this model to the Dublin Core should also be developed.
- While a naming scheme or national naming convention may be very useful, a full meta-data model is better, both in terms of the amount of data that can be stored about an item, and also to enable more powerful searching and interoperation with other projects and other countries.

Notes/Commentary

There are an impressive number of existing standards, covering various aspects of meta-data. However, there is also significant overlap across standards, and a very large population of institution-specific models, where sectoral or cross-domain models have been neglected.

References

Online

- **The TASI page on meta-data is at**
  www.tasi.ac.uk/advice/delivering/metadata.html
- The Colorado guidelines for meta-data creation and entry are at
  http://coloradodigital.coalliance.org/glines.html

- PADI's meta-data page is at http://www.nla.gov.au/padi/topics/30.html
- An unusual approach to user-generated meta-data is used at www.gimp-savvy.com
- The Dublin Core is covered at www.dublincore.org
- The Encoded Archival Description (EAD) home page is at www.loc.gov/ead/
- IFLA index on "Metadata Resources" at http://www.ifla.org/II/metadata.htm

Nominated by Minerva Partners

- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **France :** National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- ICONCLASS in Italian : www.iccd.beniculturali.it
- **Italy** : Information Network dei Beni Culturali : www.iccd.beniculturali.it
- **Italy:** SBNonline **:** http://sbnonline.sbn.it
- **Portugal:** BND: Biblioteca NAcional Digital (the National Digital Library Initiative): http://bnd.bn.pt
- **Portugal:** PORBASE: Base Nacional de DAdos Bibliográficos (the National Union Catalogue): http://www.porbase.org
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html

4.9 Publication

4.9.1 Introduction

At this stage of the project, the digital master material has been created and stored/backed up. A suitable meta-data model has been identified, and the meta-data associated with each article has been created.

Preparation for publication involves processing the newly-created material prior to publication. Typically, publication means display on the Internet, and processing means reduction in image/audio/video file size, quality, and downloads, to fit the operational characteristics of the Internet.

4.9.2 Guideline Title: Image Processing

Issue Definition

The TIFF files created during the digitisation process are typically very large (a few to many megabytes). Such files are not appropriate for Internet publication, due to the great length of time that they would require to download to the end user.

Guideline Text

- Create delivery versions of master material. As a minimum, there must be one delivery version. A second version, a 'thumbnail', may also be useful, depending on the layout of the web site on which the material is to be published.

- Delivery versions are created by opening the master TIFF file in an image processing package, and exporting it in JPEG, PNG file format(see 'Image Standards', below).

- Typically, colour resolution can be reduced, to 256 colours. If this shows an appreciable loss of quality, a higher colour resolution can be used. Choosing the right colour resolution usually requires some subjective decision to be made.

- An image created at 72 DPI will show at approximately its original size on many computer monitors. This makes 72DPI a reasonable choice for many images which are to be viewed on-screen. For lower resolutions, a subjective decision of 'acceptable quality' will be required.

- Choosing file format, colour resolution and pixel resolution involved deciding on what is 'acceptable' quality. A balance must be found between quality and file size.

- In general, the total image files on a web page should not greatly exceed 100 kilobytes. Larger images can certainly be published; however, these should be accessed via a link fom the web page, with suitable warning text that the download may be prolonged.

- Unless material is being streamed, video and audio material will typically involve large file sizes, with the file downloaded before viewing offline. However, the download time can be adjusted by changing the frames per second of the video, the sampling rate of the audio, etc.

Notes/Commentary

Decisions regarding image processing depend to a large degree on personal judgement. The guidelines provided here may be considered too strict or too lax, depending on the project and the end user audience.

Image processing software such as Paint and Paintshop is freely available online. More powerful image processing software may save sufficient time and effort to justify the expense of the software.

Audio and video editing software is also available freely online. Equally, audio and video hardware is usually supplied with the software required to edit and process the data created.

References

Online

- The open source GNU Image Manipulation Program is available at www.gimp.org
- Image optimization is addressed at http://www.yourhtmlsource.com/optimisation/imageoptimisation.html
- The University of Oregon provides a very brief look at image optimization at http://libweb.uoregon.edu/it/webpub/images.html as well as a more detailed section at http://www.uoregon.edu/~jqj/inter-pub/images/
- The University of Minnesota provides practical material on image manipulation at http://www.geom.umn.edu/events/courses/1996/cmwh/Stills/manipulating.html
- Montana State University provides guidelines for images in web pages at www.msubillings.edu/tool/

Nominated by Minerva Partners

Spain: Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
Greece: ODYSSEUS: http://www.culture.gr
Italy: DADDI**:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
Italy: Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
Sweden: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
UK: Compass**:** http://www.thebritishmuseum.ac.uk/compass

4.9.3 Guideline Title: 3D and Virtual Reality Issues

Issue Definition

The guidelines provided above for image publication are not immediately applicable to digital renderings of 3D and virtual reality material. However the balance between quality and file size is a common one on the Internet.

Guideline Text

- Viewers for 3D and VR material are not yet widely distributed with operating system software. This contrasts with image, audio and video, which are commonly provided with Windows software.

- Ensure that viewers for any 3D or VR material are readily available. Make the viewer software available from the same site as the material. This helps to overcome any issues with other software download sources becoming unavailable.

- Evaluate multiple viewers before endorsing one or another. Compatibility across file formats and viewers is not as standardized as in the still image domain.

- Modern PCs, with a focus on games, will often have hardware accelerators and increased graphics memory. This can have a profound effect on the VR viewing experience.

Notes/Commentary

A VRML viewer which has been successfully used in one of the reference projects (the Irish ACTIVATE project) is the Blaxxun Contact viewer).

References

Online

- The VRML standard is covered in some detail at www.web3d.org.
- Shockwave 3D is covered at www.macromedia.com and at http://www.3dlinks.com/community_shockwave3D.cfm
- Washington University has a very large, but slightly out-of-date knowledge base on virtual reality at http://kb.hitl.washington.edu/onthenet.html
- The US NIST also hosts a page on virtual reality resources at http://www.itl.nist.gov/iaui/ovrt/hotvr.html
- The AHDS has a guide to VR for cultural bodies at http://vads.ahds.ac.uk/guides/vr_guide/index.html
- Great archaeological sites: http://www.culture.gouv.fr/culture/exp/exp-htm

Nominated by Member States

- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy : DADDI :** http://www.uffizi.firenze.it/Dta/daddi-eng.html

4.9.4 Guideline Title: Online Publication

Issue Definition

Many digitisation projects in the cultural area lead to the creation of online cultural resources, usually a web site with images, meta-data, 3D artifacts, etc. They range from the simplest content sites to complex, software-driven portals and viewing engines. A large body of knowledge covers the creation of web sites; only a few guidelines are provided here, as well as links to examples of web sites nominated as best practice examples by Minerva partners.

Guideline Text

- Web sites should be easy to navigate – links to the front page or to a table of content should be available throughout.
- Due attention should be paid to universal access and to the utilisation of web sites by the partially sighted and other disabled persons.
- Web pages should be short enough to minimize the amount of scrolling necessary by the user.
- Images should be small enough not to disrupt the browsing experience. Larger images should be linked to from the web pages, with a note to the effect that the image is large and download may be slow.
- The use of animations, pop-ups, pop-unders, Flash and similar technologies should be treated with care. It should be possible to bypass lengthy introductory animation sequences.
- Web sites should ideally be multilingual, with at least the host country language and one or two other languages (commonly including English, as the de facto online language standard) supported.
- Links to external resources should be verified on a periodic basis, in order to minimize dead links and the annoyance associated with these.

Notes/Commentary

The actual process of making material available on the web is one which is widely understood and documented. This handbook does not provide guidance on how the create websites, program in HTML, build web-enabled databases and carry out the other tasks which are needed to create and maintain a web presence. It is anticipated that many of the cultural institutions which utilise these

guidelines will already have some web server functionality availability, which they will exploit for their digitisation project.

There are many more recommendations for the creation of web sites – the above being simply rules of thump. In the references below, some examples of different types of website are noted:

Simple information website : ACTIVATE (www.activate.ie), **: "**le piazze storiche**":** http://cantieri.theranet.it/piazze

Large multi-element websites: Biblioteca Virtual Miguel de Cervantes: http://cervantesvirtual.com/

High-tech websites with significant proprietary software : DADDI : http://www.uffizi.firenze.it/Dta/daddi-eng.html

Interactive websites with tours, etc : Compass : http://www.thebritishmuseum.ac.uk/compass

References

Online

The creation of web sites is one of the most documented topics on the web. Examples include the following, but a search with any search engine will show literally thousands more

- Web page design : http://www.essdack.org/webdesign/
- Web page authoring : http://www.htmlgoodies.com
- IASL web page awards – a source of ideas : http://www.iasl-slo.org/web_award.html
- The Louvre web page : http://www.paris.org/Musees/Louvre/
- Sun Microsystems list of library web pages, Europe section : http://sunsite.berkeley.edu/Libweb/Europemain.html

Nominated by Minerva Partners

Almost every project listed in Appendix A has a website. Some examples of websites which are interesting due to their size, or their simplicity, include the following

- **Spain**: Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **France**: INA digitisation programme of National Audio-Visual Archives.: http://www
- **Greece:** ODYSSEUS: http://www.culture.gr
- **Ireland:** ACTIVATE: http://www.activate.ie

- **Italy:** DADDI: http://www.uffizi.firenze.it/Dta/daddi-eng.html

- **Italy:** Edit16**:** http://edit16.iccu.sbn.it
- **Italy :** www.pinacotecabologna.it
- **Italy : "**le piazze storiche**":** http://cantieri.theranet.it/piazze
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it
- **Portugal:** MatrizNet: http:// www.matriznet.ipmuseus.pt (high quality web site).
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK:** Compass**:** http://www.thebritishmuseum.ac.uk/compass

4.10 IPR and Copyright

4.10.1 Introduction

The publication of any material online must be accompanied by some consideration of the intellectual property rights (IPR) associated with the material. For material which is in the public domain (such as particularly old books or newspapers, or material placed explicitly in the public domain), there is relatively little difficulty. However, many cultural institutions derive revenue from the use of images of artifacts or images in their collections, and so are defensive of copyright. Material, the copyright of which is held by third parties, can only be published with the consent of such third parties.

Fortunately, a range of technical options are available to protect the copyright of material placed on the Internet. These are surveyed here.

4.10.2 Guideline Title: Establishing Copyright

Issue Definition

The initial step when exploring the copyright situation for a cultural item is to establish the ownership of that copyright.

Guideline Text

- Establish the legal situation with regard to copyright and publication in the country where the project is being carried out. Each country has its own copyright laws, usually dating back to at least the 19th century. Such laws usually apply to all forms of publication, including online publication. They may, or may not, cover the act of digitisation, which may be construed to be an archiving process, or may be considered copying.

- On no account should online publication go ahead without copyright being sought.

- Certain items, e.g. old newspapers, have clear copyright rules governing them. Typically these allow free copying once the papers are of a certain age. Items which fit into this category can be freely digitised and published.

- For items whose copyright is vested in the institution carrying out the project, internal permission will be required for digitisation and online publication.

- For items whose copyright is held by a third party, such as the lender or donor of a collection of historical items, that party's permission must be sought, in writing. Only when such permission has been received, should publication go ahead.

- Securing permission to digitise and publish may involve payment. The amount of payment must be balanced against the value of including the relevant item(s) in the online resource.

Notes/Commentary

The copyright situation varies from country to country.

References

Online

- TASI copyright page : http://www.tasi.ac.uk/advice/managing/copyright.html
- PADI copyright page : http://www.nla.gov.au/padi/topics/28.html
- IFLA copyright page : http://www.ifla.org/II/cpyright.htm
- University of New York, Buffalo, includes many links to copyright pages at http://ublib.buffalo.edu/libraries/units/cts/preservation/digires.html
- UK : Cedar's guide to IPR : http://www.leeds.ac.uk/cedars/guideto/ipr/
- UK : MCG Copyright in Museums and Galleries : http://www.mda.org.uk/mcopyg/index.htm
- UK : Library Association Copyright Paper : http://www.la-hq.org.uk/directory/prof_issues/pospaper.html
    - IRCAM digital library on music: http:// mediatheque.ircam.fr

Nominated by Member States

- **Italy***:* Mediceo avanti il Principato on line*:* http://www.archiviodistato.firenze.it/Map/
- **Italy:** SBNonline **:** http://sbnonline.sbn.it
- **Italy** : TRADEX : http://www.tradex-ist.com

4.10.3 Guideline Title: Safeguarding Copyright

Issue Definition

The publication of items online on the web is an open invitation to make copies of the items. It is infeasible to prevent some level of copying of material displayed on the web. However, there are a number of possible procedures which can be considered, each of which has some effect in the safeguarding of copyright.

Guideline Text

- Establish whether or not copyright must be safeguarded.
- Agree the procedures to be used to safeguard copyright, with the copyright holders.
- The following procedures are among those which could be considered
    - Addition of a visible watermark or copyright stamp on each image.
    - Addition of an invisible digital watermark on each image. Such marks can be used to prove the ownership of a 'stolen' image, as well as to track the use of the image across the Internet.
    - Encryption of images, with the issuing of the appropriate key only to registered users. This, of course, reduced the value of the online image to the rest of the public.
    - Restricting publication to low-resolution images, such as 75 dpi for screen viewing. This restricts the degree to which images can be used in other domains, such as printing, clothing, etc.
    - Restrict publication to only small parts of an image. The Italian DADDI project (see references) is an excellent example.
- Display images only to registered, authorized members of a particular community.
- Test the results of the copyright protection process using the first few items, in order to ensure that the process does not have any unexpected or unwanted effects.

Notes/Commentary

The approach which is most appropriate for any one project will depend to a large degree on the goals of the project and the cultural institution, as well as on

the nature of the material. In general, the publication of a small selection of images, at low resolution, is a common approach for online galleries and museums. The relative uniqueness of many cultural holdings provides proof of ownership of copyright in many situations.

References

Online

- TASI copyright page : http://www.tasi.ac.uk/advice/managing/copyright.html
- PADI copyright page : http://www.nla.gov.au/padi/topics/28.html
- IFLA copyright page : http://www.ifla.org/II/cpyright.htm
- University of New York, Buffalo, includes many links to copyright pages at http://ublib.buffalo.edu/libraries/units/cts/preservation/digires.html
- Digimarc digital watermarks – www.digimarc.com
- Signumtech digital watermarks – www.signumtech.com
- Audio digital watermarks – www.musicode.com
- Watermarking overview - http://www.webreference.com/content/watermarks/
- General UK copyright information - http://www.copyrightservice.co.uk/copyright/protecting(02).htm
- AHDS has a copyright FAQ at http://ahds.ac.uk/copyrightfaq.htm

Nominated by Member States

The following are nominated projects with a particular interest in, or focus on, copyright.

- **Italy :** Mediceo avanti il Principato on line**:** http://www.archiviodistato.firenze.it/Map/
- **Italy** : TRADEX : http://www.tradex-ist.com
- **Italy:** DADDI: http://www.uffizi.firenze.it/Dta/daddi-eng.html (hi-tech, interesting approach).

4.11 Managing Digitisation Projects

4.11.1 Introduction

The success of any project, including digitisation projects, is influenced to a large degree by the management of the project. This section provides a small number of guidelines specific to the management of digitisation projects in particular.

4.11.2 Guideline Title: Digitisation Process Management

Issue Definition

A typical digitisation project will involve dozens, hundreds or even thousands of items to be digitised. In order to achieve an efficient project, it is important to establish a work-flow that maximises the through-put of the digitisation team. In addition, information resources such as a digitisation project knowledge base will be of significant importance.

Guideline Text

- Establish and document each of the steps that an item must go through during the digitisation process. These will include, for example,
  - retrieval from storage / usual location
  - cleaning or preparation
  - scanning or photography
  - return to usual location
  - file naming
  - file storage
  - creation of online delivery versions of large master files
  - backup of servers / storage media

- Develop a digitisation project knowledge base that can be used to track the object through the digitisation process, and enables the status of the project to be reviewed at any time. This knowledge base may take the form of a database (e.g. in MS Access, Oracle, MySQL, etc), or may use a simple spreadsheet or even a collection of documents. The important issue is not the format of the knowledge base, but the process which ensures the recording of actions which are carried out.

The name, identifier and other relevant information for each item to be digitised should be entered in the digitisation project knowledge base, as soon as the item has been selected. The status of the item (i.e. which step it is has last completed) must also be recorded, on an ongoing basis.

- Procedural choices must be made – for example, should items be collected at the digitisation workstation at the start of each day, each week, or on a per-item basis.

- Articles which require similar activities or hardware setups should be digitised together. This reduces time spent setting up digital cameras, configuring scanners, etc. The parameters for hardware setup should be documented, in order to allow any digitisation to be replicated in the event of file loss, etc.

- The location, phone numbers and backup staff of key service delivery personnel (e.g. IT support) should be noted at the start of the project, and remain available throughout.

Notes/Commentary

The larger the project, the more worthwhile it is to establish a process and workflow. The efficiencies which this introduces will greatly repay the time spent setting them up. The references below include some projects which concentrate purely on this aspect of digitisation.

References

Online

- A guide to digitisation project management and workflow is provided at http://www.tasi.ac.uk/advice/managing/jidi_workflow.html
- A comprehensive manual for many aspects of the digitisation project process is provided by the **NOF-Digitise** Technical Advisory Service Manual : http://www.ukoln.ac.uk/nof/support/manual/
- The Colorado Digitisation Program has a section on project management at http://www.cdpheritage.org/resource/project%20management/rsrc_project_management.html
- So does Canadian Heritage at http://www.chin.gc.ca/English/Digital_Content/index.html
- AHDS has a section on managing digitisation projects at http://www.ahds.ac.uk
- Chapman, Stephen and William Comstock. "Digital Imaging Production Services at the Harvard College Library." (http://www.rlg.org/preserv/diginews/diginews4-6.html#feature1). RLG DigiNews (Dec. 5, 2000). A look inside the planning and workflow design of a project at the Harvard College Library in 1999.
- Fleischhauer, Carl. *Steps in the Digitization Process.* National Digital Library Program, Library of Congress (1996). (http://lcweb2.loc.gov/ammem/award/docs/stepsdig.html).
- Hughes, Carol Ann. "Lessons Learned: Digitization of the Special Collections at the University of Iowa Libraries." *D-Lib Magazine* (June 2000). (http://www.dlib.org/dlib/june00/hughes/06hughes.html).
- The UK HEDS Matrix provides some input on budgeting for digitisation projects at http://heds.herts.ac.uk/resources/matrix2.html

Nominated by Member States

The following are examples of nominated projects which may be in a position to provide guidance on the practical management of digitisation projects.

- **Austria** : Meta-e engine for workflow management http://meta-e.uibk.ac.at/
- **Germany** : Workflow and tools for providing access to larger quantities of archival material http://www.lad-bw.de
- **Denmark :** "The soldier in the Backyard – an interactive children's story on the Internet" **:** http://www.soldatenibaghaven.dk
- **Spain** : Biblioteca Virtual Miguel de Cervantes (Miguel de Cervantes Digital Library) : http://cervantesvirtual.com/
- **Finland:** Digital historical newspaper Library 1771-1860 (ready), continuing to 1890: http:// digi.lib.helsinki.fi. The Nordic library: http://tiden.kb.se
- **France** : INA digitisation programme of National Audio-Visual Archives.: http://www.ina.fr/index.en.html
- **France :** National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Greece :** ODYSSEUS : http://www.culture.gr
- **Sweden**: The Oxenstierna Project.: http://www.ra.se/ra/Oxenstierna/oxenstierna1.html
- **UK :** Compass : http://www.thebritishmuseum.ac.uk/compass
- **UK :** NOF-Digitise Technical Advisory Service Manual : http://www.ukoln.ac.uk/nof/support/manual/

4.11.3 Guideline Title: Team Development

Issue Definition

Digitisation projects often expose the staff of cultural institutions to new technologies for the first time. Such technologies include digitisation hardware, web publication, image processing, meta-data tagging, database development and population, etc.

Guideline Text

- If possible, include at least one person with appropriate information technology skills in the project team.

- Assess the state of knowledge of the personnel to work on the project, and the IT skills that they will need, well in advance of the project. Identify training needs and fill these before the project starts.

- IT skills are not the only ones which may be needed. Specialist skills may be needed, as noted above, in the handling of delicate documents, artifacts, etc. Appropriate training maybe available from the individuals whose responsibility includes the source material.

Notes/Commentary

It is better to have a small core of skilled personnel working on a project than a larger population of occasional participants. However, while developing and using a particular skill is efficient for the project, staff may prefer to be exposed to the full digitisation life-cycle. Digitisation and meta-data tagging is not in itself particularly rewarding work – exposure to other elements of the project will increase staff satisfaction.

References

Online

-

Nominated by Member States

Denmark: "The soldier in the Backyard – an interactive children's story on the Internet" **:** http://www.soldatenibaghaven.dk (especially handling large collaborative projects)
France: National digitisation programme - annual project calls: http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
Ireland: ACTIVATE: http://www.activate.ie

4.11.4 Guideline Title: Staff Training

Issue Definition

Unless the staff working on the project has significant experience from prior projects, there will be a requirement for staff training. This will include two quite different areas – the technology to be used, and the handling of the source material.

Guideline Text

- Do not assume that no staff training is required, nor that archives, library or museum staff automatically has all the relevant expertise.

- Ensure that the training requirements of the staff on the project are identified at the start of the project, i.e. already in the planning phase. These training requirements should be included in the digitisation project knowledge base, and acted upon before the training is needed in the project.

- Certain training, such as the use of the digitisation technology, may be to learn 'on the job'; other training, such as handling of source materials, requires training in advance.

- A smaller core of personnel, who are trained and develop experience during the whole project, is to be preferred to a larger, more casual group which changes its membership more frequently.

- Technology training may be well delivered from another project in the same institution; alternatively an outside digitisation agency may be able to provide training..

- Curator training may best be provided by the individuals who are responsible for the care of the original material.

Notes/Commentary

A lack of staff training can lead to unfortunate and irreversible accidents or incidents early in the project; the same may result at any time if staff is removed from the project and new personnel start to work. A small, well-trained core is a desirable aspect of such projects.

Time invested in training at the start of the project should be repaid in extra productivity and less problems during the life of the project.

References

Online

- HATII digtisation summer school: http://www.hatii.arts.gla.ac.uk/SumProg/Digiss03/
- TASI training: http://www.tasi.ac.uk/training/training.html

4.11.5 Guideline Title: Working with Third Parties for Technical Assistance

Issue Definition

It is often appropriate for a digitisation project to engage the services of one or more third parties during the project. The services which are most commonly provided include the actual digitisation itself, the management of the project, integration with third party systems, software development, etc. This allows the cultural body to concentrate on its own areas of expertise, without need to train and retain staff with advanced IT or other skills.

Guideline Text

- As with any other project, the relationship between technical partners and other project members should be governed by clear, strict contracts. A documented and signed specification of the products or services to be provided should be agreed before any work is carried out.

- The work being carried out should be reviewed on a regular basis, to ensure that what is being delivered is in fact what the project wants or needs.

- While the use of third parties can be convenient, it should be borne in mind that any expertise or experience to be gained during the execution of the outsourced work will be lost to the cultural institution at the end of the project. This also applies to temporary staff who is employed for the duration of a project. It may be better to dedicate a long term member of staff to a project, while replacing him in the short term with a contractor.

Notes/Commentary

Certain large projects, such as the French national digitisation programme, have identified a preferred supplier, the relationship with whom may stretch for several projects and several years. Having established a working relationship with a supplier, the value of changing supplier between projects may need to be questioned.

References

Nominated by Member States

France: National digitisation programme - annual project calls: http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
Ireland: ACTIVATE: http://www.activate.ie
Italy: DADDI**:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
Italy: TRADEX: http://www.tradex-ist.com
UK: Compass**:** http://www.thebritishmuseum.ac.uk/compass

4.11.6 Guideline Title: Working with Third Parties in Cooperative Projects and Content Sharing

Issue Definition

Many digitisation projects are either cooperative efforts which involve two or more cultural bodies, or else EU-funded Framework projects, which almost always have multiple partners in multiple countries. The guidelines for establishing and managing multi-partner projects are many, and go beyond the scope of this document. However, a few pointers are included

Guideline Text

- Ensure that all partners are aware of, and have endorsed, their roles and responsibilities within the project. Refresh this knowledge on a regular basis.

- Establish a common mode of communication across partners, and ensure that all partners receive the information which is aimed at them. Electronic mail is ideal for this purpose, so long as partners read and reply to such mail.

- Subcontractors should be governed by strict commercial contracts, with their deliverables clearly and unambiguously defined.

- The IPR of all partners should be clearly documented and formally signed by all partners. A partnership agreement which clearly states the IP Rights covering material which is being brought to the project, and material which is created by the project, should be agreed in advance of the project commencing.

- Each partner should have a clear role in the project – if a partner's role is not clear, review whether or not the partner is necessary to the project.

Notes/Commentary

The notes above are only a small part of the possible material that could be provided on the establishment and management of multi-partner projects. Partners and suppliers are a major source of delay and confusion within a project – clear agreement and common endorsement of the roles and responsibilities of all partners at all times can help to avoid this.

References

Online

- TASI has a section on the use of sub-contractors, at http://www.tasi.ac.uk/advice/managing/manage.html
- Musenor projects for museums of Northern France: http://www.musenor.org/

Nominated by Member States

Many of the nominated projects worked with third parties. Some examples are:

- **Denmark :** "The soldier in the Backyard – an interactive children's story on the Internet" **:** http://www.soldatenibaghaven.dk
- **France :**National digitisation programme - annual project calls : http://www.culture.gouv.fr/culture/mrt/numerisation/index.htm
- **Ireland :** ACTIVATE : http://www.activate.ie
- **Italy :** DADDI **:** http://www.uffizi.firenze.it/Dta/daddi-eng.html
- **Italy :** Rinascimento Virtuale-Digitalepalimpsest Forschung (RV) : www.iccu.sbn.it , www.bml.firenze.sbn.it (large network of 42 partners)
- **Italy:** SBNonline **:** http://sbnonline.sbn.it
- **Italy** : S.I.T.I.A : www.archeologia.beniculturali.it
- **Italy** : TRADEX : http://www.tradex-ist.com
- **Potugal** : Endovelliccus : www.ipa.min-cultura.pt
- **Sweden**: The Oxenstierna Project. : http://www.ra.se/ra/Oxenstierna/oxenstierna1.html

4.11.7 Guideline Title: Costs

Issue Definition

Building a digital collection is normally very expensive. Projects, therefore, have to take into account all start-up and infrastructural costs as well as costs for running the project. That means costs for initial planning, data specifications, tracking and documentation systems, staff training, and so forth, as well as the incremental cost for digitising the actual source material.

Guideline Text

- Digitisation projects should consider the following costs involved in the design, implementation and maintenance of a digital collection:
  - Staff development
  - Facilities needed
  - Operational costs
  - Costs for storage and for delivery systems

- Staff development includes salaries for project management, web programmer, educational officer, technical support, etc. but also travel costs and training

- Costs for facilities are often connected to questions concerning required functionalities and the need for tradeoffs. Projects have for instance to decide whether keeping costs on a low level is more important to the overall success of the project than achieving the highest possible standard for image capture.

- Operational costs to consider are:
  - Time for handling source material (from shelf to image capture device and back again) as a percentage of total salary cost per day.
  - Preparation of source material (conservation, cleaning etc).
  - Capture time (from set-up to naming and saving) provided as a percentage of the operators total salary costs per day.
  - Cataloguing and handling of metadata as a percentage of total salary costs.
  - Hardware and software cost per digitised item (preferably based on depreciations or replacements costs rather than acquisition costs.
  - Quality assurance time as a percentage of salary costs.
  - Hardware and software maintenance.
  - Technical support time related to capture.
  - Project management time related to capture.
  - Training related to capture.

- Be aware of that image capture often is the least costly part of a digitisation project. On average, one third of total costs are connected to digital convention, slightly less than one third to creation of metadata, a bit more then one third to administrative and quality assurance tasks. The rest is long term maintenance costs.
- Storage costs to consider should normally be total costs for maintenance per gigabyte.

Notes/Commentary

Sustainability in the long term is often pushed down the list of priorities by more immediate and pressing concerns. Regardless of the quality and robustness of the digital resources created by a digitisation project, they will not last long if the project in mind cannot find funds for their maintenance.

References

Online

- Research Libraries Group: http://www.rlg.org
- Library of Congress: http://www.loc.gov
- Online Computer Library Center: http://www.oclc.org/home/
- The report Building and Sustaining Digital Collections: Models for Libraries and Museums: http://www.clir.org/pubs/reports/pub100/pub100.pdf

5 Standards

5.1 Introduction

This chapter surveys some of the many technical standards which exist in the digitisation and online publication areas. Some of the most important of these (e.g. the Dublin Core meta-data standard) were created for other domains, but have found application in the digitisation area. Others are 'pure' technology, such as the TIFF, JPEG and GIF image format standards. Others again are 'de facto' industry standards which, while widely supported and used today, may become obsolete in a relatively short period of time.

This section surveys standards which apply to the various stages of the digitisation life-cycle. These include

- technology standards
- meta-data standards
- taxonomy and naming Standards

It should be noted, however, that the list of standards presented here is selective; the guidelines, procedures, models, ontologies, thesauri etc. which exist in this area are very numerous. For example, Minerva Deliverable D6.1 provides links to standards bodies in ISO, CEN etc whose work may be relevant to digitisation projects.

It may also be noted that it is worthwhile for any digitisation project to survey the state of the digitisation art before beginning – this will provide an updated version of the standards which are most widely supported at the time of the project. The standards discussed in this section have already demonstrated longevity, and so can be expected to persist, or else have such a dominant industry position that the size of the user base is expected to dictate support and migration paths, going forward.

A huge range of technology standards is applicable to, or can be applied to, the digitisation area. This reflects the long history of digitisation and the computer graphics industry, as well as the ability of the IT world to create new standards on an ongoing basis. Practically any area in the IT domain has a wide range and choice of standards covering it. The most relevant from a digitisation project point of view are those which cover

- Images
- Audio material
- Video material
- 3D material

Standards which are widely used, preferable ISO standards, should be used.

It must be emphasized that the number of standards and the material which has been written about them are both very large. The amount of this material which is available online is impressive. A targeted online search using a search engine such as Google is likely to fulfil almost any information need in this area. Alternatively, exploration of the references provided in this document will also be fruitful.

5.2 Image Standards

The use of relevant image standards is critical to any digitisation project that wishes to share or publish the image files which it creates. Fortunately, this area has a small number of very dominant standards, and these standards enjoy widespread support.

TIFF (Tagged Image File Format)

TIFF (latest version 6.0) is a proprietary format. The ISO standardised TIFF/EP is valid to TIFF (6.0). This standard is relevant to the creation of high-quality digital images. There is no compression involved, and so TIFF images are typically very large, high-quality files. TIFF output can be anticipated from any scanner or digital camera, either as its native format or (more commonly) as an export option from the proprietary software provided with the hardware.

Master images should be stored in TIFF format unless there is a good reason for using some other format.

A TIFF specification can be found at
http://www.dcs.ed.ac.uk/home/mxr/gfx/2d/TIFF-6.ps.gz

JPEG (Joint Photographic Experts Group)

This ISO-standard is widely used to deliver images across networks with limited bandwidth, such as the Internet and most intranets. The standard utilizes mostly "lossy" file compression to reduce the size of the file being transmitted across the network. The display of JPEG files is supported by all web browsers and by a large number of desktop applications.

JPEG images can be created from TIFF images, using image processing software.

For more information on JPEG, see www.jpeg.org, or the user-friendly JPEG FAQ at http://www.faqs.org/faqs/jpeg-faq/

A jpeg specification can be found at
http://www.dcs.ed.ac.uk/home/mxr/gfx/2d/JPEG.txt

JPEG2000

The ISO-standardised JPEG2000 format is the successor of the above traditional JPEG format. Contrary to the technology of JPEG, JPEG2000 technology provides very high "lossy" or "lossless" compression without compromising image quality. The layered (multi-resolution) structure of JPEG2000 can eliminate the need to maintain several different resolution version of the same images in an images database (i.e. just ***one*** "global" image could be used as a master file as well as a surrogate image for web access).

GIF (Graphics Interchange Format)

In common with JPEG, this format is widely used to deliver images across networks with limited bandwidth, such as the Internet and most intranets. The format utilizes lossless file compression to reduce the size of the file being transmitted across the network. Depending on the nature of the image, GIF or JPEG may be more appropriate. GIF is well suited to cartoons, icons and simpler graphics, JPEG suits scanned photographs and complex images better. However, both are orders of magnitude smaller in file size than TIFF. The display of GIF files is supported by all web browsers and by many desktop applications.

It may be noted that GIF is in fact a proprietary file format, covered by patent.

GIF images can be created from TIFF images using image processing software.

A GIF specification can be viewed at
http://www.dcs.ed.ac.uk/home/mxr/gfx/2d/GIF87a.txt

PNG (Portable Network Graphics)

PNG images are supported by the most recent versions of the mainstream browsers. They offer a higher quality image than GIF or JPG for many pictures, but at the cost of a somewhat larger file size.

Support for PNG beyond the web technologies area is still somewhat sparse.

PNG images should be created using image processing software, which imports a TIFF image and exports PNG images.

The PNG specification is at http://www.w3.org/TR/REC-png-multi.html

5.3 Audio Standards

The standards surveyed briefly here are those most relevant to the web publication of audio material. Their support in the mainstream desktop environment is of great importance, since this decides to a large degree the size of the audience which they will address.

Audio standards for commercial and professional sound engineering are not covered here.

For a general coverage of audio file formats, see the Audio File Format FAQ at http://home.sprynet.com/~cbagwell/audio.html or the Duke University Audio site at http://cit.duke.edu/resource-guides/tutorial-web-multimedia/06-audio-formats.html

WAV

This is the standard Windows audio file format, and is supported by modern versions of Windows using the inbuilt Windows Media Player. As a result it has a very large market penetration.

However WAV is not particularly well suited to the online publication of digitised sound, due to the large file sizes it creates. For instance 1 minute of CD quality audio recorded at 16-bit rate and sampled at 44 kHz gives a file size of about 10mb in WAV format.

MP3

This digital audio standard has a large user base, particularly on the Internet, due to its small file size and high quality. It is part of the MPEG family of multimedia standards. It is also supported by the widespread Windows Media Player. Information on the MP3 standard is available at www.mp3-tech.org

Real Audio

This is a proprietary digital audio format created and supported by Progressive Networks (www.real.com). It has a significant user base due to the free availability of the player software and its early market penetration. File sizes are smaller again than MP3, though the quality of the sound is also slightly less.

5.4 Digital Video Standards

Again, this section focuses on the standards for online publication of digitised content. Video is a powerful tool for the presentation of a continuous view of all sides of an object, or for the presentation of three-dimensional spaces, without the need to create full virtual reality content. The availability of economical digital video camera equipment also makes this technology accessible for small or pilot digitisation projects.

The material covered here can be researched in much greater detail at Duke University's comprehensive site (cit.duke.edu).

MPEG (Motion Pictures Expert Group)

This format is popular on web sites, due to the relatively short download time and the widespread availability of player software (including the Windows Media Player). Sound and video are often combined in a single file. MPEG gives high quality and a relatively small file size.

The MPEG standards can be investigated further at www.mpeg.org

Real Video

This is a proprietary format created and supported by Progressive Networks. Its popularity is based on a good quality picture and the free availability of player software. The quality of the image can be adjusted in order to take into account the desired file size. However, the MPEG standard is becoming dominant in this area, and the proportion of online Real Video material is decreasing.
Real Video is accessed at www.real.com

QuickTime

QuickTime is the dominant video format specifically for the Apple platform. The popularity of the Mac in the multimedia domain means that a great deal of material is created and published in this format. Very high quality can be achieved; however, the large size of the files makes it less appropriate for mainstream Internet use.

The QuickTime file format can be accessed at
http://developer.apple.com/techpubs/quicktime/ qtdevdocs/QTFF/qtff.html

5.5 3D Standards

The creation and publication of three-dimensional material is a powerful tool for cultural content. This is particularly the case for museums, whose holdings are primarily three-dimensional (3D) objects, and for historic buildings and heritage landscapes.

As noted above, digital video is a low-cost alternative to the creation of true 3D models; however, such an approach does not support the attractive interactive manipulation of objects and exploration of landscapes that a true 3D model enables.

Online 3D technologies are well covered in the site of the Web3D consortium, which includes a range of industry players. See http://www.web3d.org A more casual coverage can be found at www.3dsite.com and at http://www.tnt.uni-hannover.de/subj/vrml/

VRML (Virtual Reality Markup Language)

The VRML standard is the dominant 'official' standard for the modeling of virtual reality and 3D material. Despite having been available for several years, however, its take-up has been sporadic. While several players exist for the browsing of VRML content, it has not yet entered the mainstream desktop in the manner of audio or video. Virtual tours of museums and galleries are relatively common, however, with some excellent examples available online.

In common with video, VRML content cannot usually be 'streamed' to the end user, due the significant size of the files involved. Instead, VR material is downloaded as a compressed (zip) file, and then viewed locally.

The VRML standard is covered in some detail at www.web3d.org.

Shockwave 3D

Shockwave 3D is a new technology allowing 3D models to be imported into 'Macromedia director ' (The industry standard for publishing interactive online/ CD based content). 3D interactive content can then be published as a 'Shockwave' file, viewable by anybody with the latest version of the free, cross platform 'Shockwave' viewer plug-in, which has the best market penetration of any plug-in technology (estimated at 69.9% of the online market in March 2002)(source: macromedia).

The main disadvantage of Shockwave 3D is that it is not as mature as VRML for creating these kinds of online experiences. S3D does not allow a simple navigational 3D experience to be constructed as easily as VRML. And S3D does not have VRMLs extensible design. Really all that Shockwave 3D does, at present, is it allows a 3D animation to be played back within director and has a few predefined 'behaviours' for camera moves etc. Anything else needs to be scripted from the ground up. Shockwave 3D has the scope to offer all that VRML does, and more, but for the present VRML is a better, faster development environment for small scale projects.

A great deal of information about this popular format is available online. This includes the manufacturer's site at www.macromedia.com as well as third party content such as that at http://www.3dlinks.com/community_shockwave3D.cfm

5.6 Meta-data Standards

Dublin Core

The use of meta-data to describe the content of digital files is central to the discovery of particular or relevant items in large collections. Metadata helps to remove the ambiguity of free-text searching, and to add some semantic aspects which narrow and focus an information retrieval search activity.

In order to be of value, metadata must follow conventions and standards, so that those searching an information resource can use the same meta-data tags and values as those who create and maintain the resource.

Fortunately, in the information retrieval domain, one standard is very dominant. This, the Dublin Core standard (named after Dublin, Ohio), provides a short list of the most commonly used meta-data terms, as well as an extension mechanism. While Dublin Core was originally intended for libraries, it has been widely adopted on the Internet and across into other domains. It is an official ANSI Standard, Z39.85

A detailed description of the Dublin Core standard, and an exploration of the fields which it includes, can be had from http://au.dublincore.org/documents/dces/ or from www.dublincore.org

Other Meta-data Standards

There are a very large number of meta-data standards and models available. A partial directory of some of the most important is provided at http://www.ulb.ac.be/ceese/meta/meta.html
In addition, there are major meta-data sites at the WorldWide Web Consortium (http://www.w3.org/Metadata/) and at IFLA (http://www.ifla.org/II/metadata.htm).

Of particular interest is the W3C work on self-describing data, represented by the Resource Description Framework (RDF) standard. See www.w3c.org/rdf. RDF can be used as enabling technology for Dublin Core, for example. See, among others, www.ukoln.ac.uk/metadata/resources/ dc/datamodel/WD-dc-rdf/

Some standards which impinge on the libraries and cultural area include

- Government Information Locator Service (GILS) at http://www.dtic.mil/gils-input/htgi/htgiinp.html
- Computer Interchange of Museum Information (CIMI) model for museums
- Encoded Archive Description (EAD) at http://www.lcweb.loc.gov/ead/
- Text Encoding and Interchange (TEI) at http://www.hti.umich.edu/docs/TEI/

- NCITS L8 proposed draft ANSI standard for meta-data at http://pueblo.lbl.gov/~olken/X3L8/
- Machine-Readable Cataloguing (MARC) at http://www.loc.gov/marc/ and elsewhere.

The range and scope of the meta-data standards varies significantly. A meta-data standard that covers almost any aspect of feasible digitisation projects will already have been created – creating a new one is not recommended.

Despite the range of meta-data standards available, the Dublin Core work is the most widely used and referenced; unless there is a good reason not to, DC fields should be included in whatever meta-data standard a new project utilizes.

5.7 Taxonomy and Naming Standards

Significant effort has been invested in the creation of standard taxonomies and naming schemes for the cultural domain. These attempts to enforce some consistency on the semantics of commonly used terms, as well as to identify synonyms and alternative names for the same concept or person.

The Dublin Core meta-data standard, surveyed briefly above, recommends that many meta-data fields be populated from restricted, recognised populations of terms. This greatly facilitates searching for particular information.

The number of taxonomies and naming standards which have been created is quite large – some samples are provided here, but a great deal more information on this topic is available online, at resources such as TASI (Technical Advisory Service for Images) at http://www.tasi.ac.uk and VADS (Visual Arts Data Service) at http://www.vads.ahds.ac.uk

Controlled vocabularies, thesauri and classification systems available on the WWW http://www.lub.lu.se/metadata/subject-help.html.

The High Level Thesaurus Project (HILT) is a clearinghouse of information about controlled vocabularies, including related resources, projects, and an alphabetical list of thesauri. http://hilt.cdlr.strath.ac .uk/Sources/index.html

The Getty Vocabulary Program builds, maintains, and disseminates several thesauri for the visual arts and architecture:

- Art & Architecture Thesaurus® (AAT) http://www.getty.edu/research/tools/vocabulary/aat/
- Union List of Artist Names® (ULAN) http://www.getty.edu/research/tools/vocabulary/ulan/

- Getty Thesaurus of Geographic Names™ (TGN) http://www.getty.edu/research/tools/vocabulary/tgn/

Some other controlled vocabularies:

- Library of Congress Subject Heading List-Available through OCLC, RLG and other cataloging services and on CD ROM from the Library of Congress.
- Thesauri of Graphic Materials I: http://lcweb.loc.gov/rr/print/tgm1/
- Thesauri of Graphic Materials II: http://lcweb.loc.gov/rr/print/tgm2/
- Thesaurus of Graphic Names: http://www.gii.getty.edu/vocabulary.tgn.html

6 Digitisation Guidelines: A Selected List

Selection criteria

The list is not exhaustive but wants to be selective. The list is limited to guidelines for digitization of paper based on documentary heritage, that is manuscripts, printed books and photographs of libraries, archives and museums, not for digitization of multimedia materials. Toolbox and tutorial have been included too, considering these learning resources as valuable as guidelines
The selected Guidelines have been produced by public and private institutions: some are for guiding the digitization projects, others are related to digitization programs where the Guidelines want to reach the strategy and mission of single institutions: the criteria followed for inclusion was that of general interest for professionals worldwide.

The list of digitization guidelines is a work in progress, to be updated constantly.
For updating, see: http://www.minervaeurope.org/guidelines.htm.

The data chosen for description are: Author, Contributor (if existing), Title, Description, Date, Format and URL. The presentation is in alphabetical order by author. We welcome your comments and suggestions.

AHDS (Arts and Humanities Data Service)

- *Guide to Good practice in the Creation and Use of Digital Resources*
  Available formats: HTML
  http://www.ahds.ac.uk/guides
  Guidelines for: Archaeology, History, Performing Arts, Textual Studies, Visual Arts. Each of these Guides includes tips for discovering and re-using digital data, information about creating and managing new digital data, and guidance to ensure proper preparation and documentation of this data for long term archiving

- *Managing Digital Collections*
  Available formats: HTML
  http://ahds.ac.uk/managing.htm
  This guide gives a framework of strategies and standards for developing, managing, and distributing high-quality digital collections.

British Library

- *Objectives of Digitization*
  Available formats: HTML
  http://www.bl.uk/about/policies/digital.html#one
  The policy covers all materials originally produced in non-digital form (e.g. printed matter of all kinds, manuscripts, photographs, drawings, paintings, sound recordings, microforms), the digitization of which would fulfil one or more of the desired objectives. It includes objectives, scope, context and BL examples.

- *Preservation and digitization: principles, practices and policies*
  Available formats: HTML; PDF; print publication
  http://www.bl.uk/services/preservation/freeandpaid.html
  Realised by NPO (National Preservation Office), this is a series of guidelines whose aim is to provide an independent focus for ensuring the preservation and continued accessibility of library and archive material. Free and paid material is offered.

CHIN (Canadian Heritage Information Network)

- *Creating and managing digital content* (April 2002)

Available formats: HTML
http://www.chin.gc.ca/English/Digital_Content/
Capture_Collections/index.html
Series of Guidelines for creating and maintaining a digitization project. The titles include:

- Capture your collections,
- Web site development,
- Web site development resources,
- Intellectual Property,
- Collection Management,
- Standards.

- *Producing Online Heritage Projects* (August 2002)
  Available formats: HTML
  http://www.chin.gc.ca/English/Digital_Content/
  Producing_Heritage/index.html
  This handbook is for heritage professionals who are developing online content, and helps them to achieve the benefits available from Web-based education and promotion. It focuses on skills needed for the creation, management and presentation of digital content. The index includes:
  - Project planning,
  - Project development,
  - Getting ready to launch,
  - Product maintenance

  Annexes: Glossary, Bibliography, Project manager's tools and templates.

- *Program Guidelines* (April 2002)
  Available formats: HTML; PDF
  http://www.chin.gc.ca/English/Members/
  Vmc_Investment_Program/guidelines.html
  Virtual Museums of Canada Investment Program. It includes:
  - Operating principles;
  - Performance indicators;
  - Governance structures;
  - Content policy;
  - Skills development.

  Annexes: Guidelines for calculating cost/values.

- *Capture your collections. Planning and implementing digitization projects* (April 2002)
  Available formats: HTML; PDF
  http://www.chin.gc.ca/English/Digital_Content/
  Managers_Guide/index.html
  Modules and sections of a on line course on digitization. It includes:
  - Project planning;
  - Legal Issues related to Digitization;
  - Determining the costs of a Digitization Project;
  - Standards and Guidelines to Consider;
  - Implementation;
  - Maintenance/Management;

CLIR (Council on Libraries and Information Resources)

- Abby Smith. *Building and sustaining digital collections: models for libraries and archives* (August 2001)
  Available formats: HTML; print publication
  http://www.clir.org/pubs/abstract/pub100abst.html
  This guide brings together libraries, museums and academic communities. The focus

is on scholarly publishing, with presentations of business models. This is an agenda for:

- develop sound selection criteria;
- identify online audience;
- manage intellectual property rights;
- develop and share best practices for technological issues;
- implement cost recovery strategy;
- manage the institutional transformation.

Colorado Digitization Project

- *Digital Toolbox* (2002-2003)
  Available formats: HTML
  http://www.cdpheritage.org/resource/toolbox/index.html
  The purpose of this toolbox is to introduce cultural heritage institutions to the range of issues associated with digitization of primary source materials. Provides links to general resources, bibliographies, initiatives, and clearinghouses on selection, scanning, quality control, metadata creation, and other project management issues. Also offers a glossary of digital imaging terms.

Cornell University Library

- *Moving theory into practice: Digital imaging tutorial* (2002-2003)
  Available formats: HTML; PDF
  http://www.library.cornell.edu/preservation/tutorial/contents.html
  This tutorial, produced also in Spanish and French, includes:
  - Basic terminology,
  - Selection,
  - Conversion,
  - Quality control,
  - Metadata,
  - Technical Infrastructure,
    - Digitization chain
    - Image creation
    - File Management
    - Delivery
  - Presentation,
  - Digital Preservation,
  - Management,
  - Continuing Education.

CUL (Columbia University Libraries)

- Annr R. Kenney - Stephen Chapman. *Digital Imaging for Libraries and Archives*
  Available formats: HTML; print publication
  http://www.library.cornell.edu/preservation/dila.html
  The volume begins with a theoretical overview of the key concepts, vocabulary, and challenges associated with digital conversion of paper-and film-based materials. This is followed by an overview of the hardware/software, communications, and managerial considerations associated with implementing a technical infrastructure to support a full imaging program. Additional chapters present information on the creation of databases and indexes, the implications of outsourcing imaging services, converting photographs and film intermediates, issues associated with providing long-term access to digital information, and suggestions for continuing education.

- *Selection Criteria for Digital Imaging Projects* (January 2001)
  Available formats: HTML
  http://www.columbia.edu/cu/libraries/digital/criteria.html

The criteria listed are important to assure that issues of technical feasibility, intellectual property rights, and institutional support are considered along with the value of the materials and the interest of their content.

- *Technical Recommendations for Digital Imaging Projects* (1997)
  Available formats: HTML
  http://www.columbia.edu/acis/dl/imagespec.html
  Prepared by the Image Quality Working Group of ArchivesCom, a joint Libraries/AcIS committee. This document provides recommendations for image quality, file formats, and other capture and storage issues when converting paper, photographic and other physical materials into digital form.

- *Guidelines for Providing Access to Digital Images* (2001)
  Available formats: HTML
  http://www.columbia.edu/cu/lweb/projects/digital/policy.html
  Access to digital images should be provided in the most open level, consistent with the protection of intellectual property rights, and compliant with the local policies on the exercise of such rights

DLF (Digital Library Federation)

- *Digital library standards and practices* (October 2002. Last revision)
  Available formats: HTML
  http://www.diglib.org/standards.htm
  The DLF documents and promotes adoption of standards and best practices that support the effective acquisition, interchange, persistence, and assessment of digital library collections and services.

- *Guides to Quality in Visual Resource Imaging* (July 2000)
  Available formats: HTML
  http://www.rlg.org/visguides/
  This guide includes:
  - Introduction
  - Planning an Imaging Project, by Linda Serenson Colet
  - Selecting a Scanner, by Don Williams
  - Imaging Systems: the Range of Factors Affecting Image Quality, by Donald D'Amato
  - Measuring Quality of Digital Masters, by Franziska Frey
  - File Formats for Digital Masters, by Franziska Frey

DLM Forum

- *Guidelines on Best Practices for Using Electronic Information: How to Deal with Machine Readable Data and Electronic Documents* (1996 first edition; 1997, updated and enlarged edition)
  Available formats: HTML
  http://europa.eu.int/ISPO/dlm/documents/guidelines.html
  The DLM Forum, organised jointly by the Member States of the European Union and the European Commission in Brussels in December 1996, brought together experts from industry, research, administration and archives to discuss a topic of ever increasing importance: the memory of the information society. The Guidelines include:
  - from data to structured electronic information;
  - information life cycle and allocation of responsibilities;
  - design, creation and maintenance of electronic information;
  - short and long term preservation of electronic information;
  - accessing and disseminating information.

  Annexes: Terminology, Checklist for electronic information strategy, How to select metadata, Standards.

eLib

- *Preservation Studies (Supporting Studies)* (1998-2000)
  http://www.ukoln.ac.uk/services/elib/papers/supporting/
  Managed by the British Library Research and Innovation Centre, the series Preservation Studies offer several reports on creating and preserving digital image collections. One of the goals is to compare various digital preservation strategies for different data types and formats. Studies included are:
  - John Bennett. *A framework of Data Types and Formats, and Issues affecting the long term preservation of digital material*
    Available formats: HTML; PDF
  - Monika Blake - David Haynes - Tanya Jowett - David Streatfield. *Responsibility for Digital Archiving and Long Term Access to Digital Data*
    Available formats: HTML; PDF
  - Seamus Ross - Ann Gow. *Digital Archaeology: Rescuing Neglected and Damaged Data Resources*
    Available formats: Executive summary: PDF; Full Study: PDF (Mounted 15 November 1999)
  - Alan Poulter. *Preservation of digital materials; policy and strategy issues for the UK*
    Available formats: HTML
  - Denise Lievesley - Simon Jones. *An Investigation into the Digital Preservation needs of Universities and Research Funders*
    Available formats: HTML (mounted 11 November 1998)
  - Neil Beagrie - Dan Greenstein. *A Strategic Policy Framework for Creating and Preserving Digital Collections*
    Available formats: HTML; PDF; RTF
  - Tony Hendley. *Comparison of methods of digital preservation*
    Available formats: PDF; HTML; RTF

The Getty Trust

- *Introduction to Vocabularies* (2000)
  Available formats: HTML
  http://www.getty.edu/research/institute/vocabulary/introvocabs/
  The tutorial is an introduction to the topic of vocabularies and related issues - documentation, standards, and access

- Murtha Baca. *Introduction to Metadata: pathways to digital information* (May 2000)
  Available formats: HTML; PDF; print publication
  http://www.getty.edu/research/institute/standards/intrometadata/
  Version 2 of the guide, which rather than including a single crosswalk as in the previous version, is now offering a "suite" of metadata crosswalks that map different sets of metadata. The author will continue to add to and revise this section as developments arise in the development of metadata schemas that are still evolving (e.g. Dublin Core Qualified, VRA Core 3.0).

HATII (Humanities Advanced Technology and Information Institute) and NINCH (National Initiative for a Networked Cultural Heritage)

- *The NINCH Guide to Good Practice in the Digital Representation & Management of Cultural Heritage Materials* (October 2002 - Version 1.0 First edition)
  Available formats: HTML
  http://www.nyu.edu/its/humanities/ninchguide/
  The Guide describes the process of creating and distributing digital collections and looks at mechanisms by which the institution that created or holds digital collections can manage them to maximum advantage. It includes:
  - Project planning
  - Selecting materials

- Rights management
- Digitization and encoding of text
- Capture and management of images
- Audio/Video Capture and Management
- Quality Control and Assurance
- Working with others
- Distribution
- Assessment of Projects by User evaluation
- Digital Asset Management
- Preservation

In Appendixes: Equipment, Metadata, Digital Data Capture: Sampling

Harvard University Library

- *Selection for digitization: a decision making matrix* (December 1997)
  Available formats: HTML; PDF
  http://www.clir.org/pubs/reports/hazen/matrix.html
  A decision making matrix, produced as imagine, for guiding professionals in the selection. It is included in the Harvard program: Library preservation resources principles and guides.

IMLS (Institute of Museum and Library Services)

- *A Framework of Guidance for Building Good Digital Collections* (November 2001)
  Available formats: HTML
  http://www.imls.gov/pubs/forumframework.htm
  Indicators are listed for Digital objects, Metadata, Collections and Projects, within the context of networked services. Report of the IMLS Digital Library Forum on the National Science Digital Library Program
  Reference in: Priscilla Caplan et al. (2001)

Library of Congress

- *Digital strategy for the Library of Congress* (2000)
  Available formats: HTML; print publication; e-book
  http://www.nap.edu/catalog/9940.html
  LC21: A Digital Strategy for the Library of Congress discusses challenges and provides recommendations for moving forward at the Library of Congress. Topics covered include:
  - Digital collections,
  - Digital preservation,
  - Digital cataloging (metadata),
  - Strategic planning,
  - Human resources,
  - General management,
  - Budgetary issues

- *Challenges to Building an Effective Digital Library*
  Available formats: HTML
  http://memory.loc.gov/ammem/dli2/html/cbedl.html
  The staff of the NDLP (National Digital Library Program) at the Library of Congress have identified ten challenges that must be met if large and effective digital libraries are to be created during the 21st century. The challenges are grouped under the following broad categories:
  - Building the resource,
  - Interoperability,
  - Intellectual property,
  - Providing effective access,
  - Sustaining the resource.

- *Technical Notes by Type of Material*
  Available formats: HTML
  http://memory.loc.gov/ammem/dli2/html/document.html
  The notes provide general comments on digital reproductions of textual materials for American Memory, including:
  - Searchable text
  - Textual material available for use in DLI-Phase II
  - Challenges faced by NDLP (National Digital Library Program)

- *Background Papers and Technical Information*
  Available formats: HTML
  http://memory.loc.gov/ammem/ftpfile.html
  These versions represent the final document of NDL Requests for Proposals for scanning and text conversion services . Contracts have been awarded for the work described in the Requests for Proposals.

- *Manuscript Digitization Demonstration Project. Final Report* (October 1998)
  Available formats: HTML
  http://memory.loc.gov/ammem/pictel/
  The Manuscript Digitization Demonstration Project was sponsored by the Library of Congress Preservation Directorate and was carried out in cooperation with the NDLP from 1994 to 1997. The questions framed are:
  - What type of image is best suited for the digitization of large manuscript collections, especially collections consisting mostly of twentieth century typescripts?
  - What level of quality strikes the best balance between production economics and the requirements set by future uses of the images?
  - Will the same type of image that offers high quality reformatting also provide efficient online access for researchers?

- *Lessons Learned: National Digital Library Competition* (January 2001)
  Available formats: HTML
  http://lcweb2.loc.gov/ammem/award/lessons/lessons.html
  LC/Ameritech award winners are learning many lessons about digitization projects in the implementation of their award. To help award-winners, digital project managers, and others interested in this emerging field, the competition staff has summarized, extracted, and paraphrased points from some of the interim reports submitted by awardees. These include:
  - Formats and specifications for digital reproductions,
  - Production workflow and project Management,
  - Intellectual access.

- *Conservation Implications of Digitization Projects*
  Available formats: HTML
  http://memory.loc.gov/ammem/techdocs/conservation.html
  This paper was written by a group of Library of Congress conservators who have worked closely with NDLP digitization projects and NDLP project leaders since the beginning of the program in 1995. The multi-faceted and precedent setting role which conservation plays in digital image conversion projects in the NDLP in the areas of consultation, training, and treatment for scanning is discussed.

NARA (National Archives and Records Administration)

- Steven Puglia. *Guidelines for Digitizing Archival Materials for Electronic Access* (January 1998)
  Available formats: PDF
  http://www.archives.gov/research_room/arc/arc_info/

guidelines_for_digitizing_archival_materials.pdf
These guidelines have been realised to provide a method for evaluating quality of images produced, to estimate the data storage for access files (on line) and master files (off line), and to assist in determining upgrades of NARA infrastructure. Differences in document type dictate differences in approach to scanning; specifications are given for: textual documents, photographs, maps, plans and oversized records, graphic records.

National Library of Australia

- *Digitization of traditional format library materials. Standards and Guidelines*
  Available formats: HTML
  http://www.nla.gov.au/digital/standards.html
  These guidelines, created for National Library staff , provide advice on digitization projects.
  They focus on creating digital images and displaying them on the Web, including metadata and preservation issues

- *Preserving Access to Digital Information (PADI)*
  Available formats: HTML
  http://www.nla.gov.au/padi/
  The PADI site, offers a subject gateway to digital preservation resources. Includes current information on digital preservation-related events, organizations, policies, strategies, and guidelines. Also includes glossaries of terms that are relevant to digital information.

NEDCC (Northeast Document Conservation Center)

- Maxine Sitts. *Handbook for Digital Projects: A Management Tool for Preservation and Access* (December 2000)
  Available formats: PDF; print publication
  http://www.nedcc.org/digital/dman2.pdf
  Web resource providing information on the issues surrounding the digital conversion of collection materials. With contributions from many of the School for Scanning series presenters, it provides information on project selection and management, technical and copyright considerations, digital longevity and includes commentary on the transformation in scholarly access and preservation tenets required to fully utilize and maintain digital images. Given at NEDCC's school for scanning conferences, Andover, MA. It includes:
  - Rationale for digitization and preservation,
  - Considerations for project management,
  - Selection of materials for scanning,
  - Overview of copyright issues,
  - Technical primer,
  - Developing best practices: guidelines from case studies,
  - Vendor relations,
  - Digital longevity,
  - Scholar commentary.

NINCH (National Initiative for a Networked Cultural Heritage)
see **HATII**

Nordinfo. NDLC

- *Guidelines on the Establishment of Digitization Services* (July 1997 /updated November 2000)
  Available formats: HTML
  http://www.nordinfo.helsinki.fi/publications/nordnytt/nnytt3-4_97/solbakk.htm

It includes:

- Digitising documents where the original is on paper or film base
- Digitising audio
- Digitising video

NSDL/SMETE (Science Mathematics Engineering and Technology Education)

- *NSDL Metadata primer* (Last revision January 2003)
  Available formats: HTML
  http://metamanagement.comm.nsdlib.org/outline.html
  The National SMET (Science, Mathematics, Engineering and Technology Education) Digital Library (NSDL) is being constructed to support excellence in SMET for all Americans. NSDL is a comprehensive information system built as a distributed network and will develop and make accessible high quality collections. Reference: C. Manduca, F. McMartin, D. Mogk, Pathways to progress: vision and plans for developing the NSDL (2001):
  http://doclib.comm.nsdlib.org/PathwaysToProgress.pdf this primer is intended to serve NSDL partners and collaborators as they work with NSDL staff to make their metadata available through the NSDL Metadata Repository. Its primary clientele are those NSDL-funded projects which are at the beginning stages of awareness and use of metadata, but there are also sections that will be useful to others.

- *NSDL Building collections* (October 2002)
  Available formats: HTML
  http://collections.comm.nsdlib.org/cgi-in/wiki.pl?BuildingCollections
  Checklist, tools and examples are provided for those wanting to contribute to build NSDL collection, but it is useful also to others.

RLG (Research Libraries Group)

- *RLG Guidelines for Microfilming to Support Digitization* (February 2003)
  Available formats: HTML
  http://www.rlg.org/preserv/
  Offers supporting materials to institutions in their efforts to preserve and improve access to endangered research materials.

- *RLG Tools for Digital Imaging* (May 2002)
  Available formats: HTML
  http://www.rlg.org/preserv/RLGtools.html
  The tools include worksheets and guidelines for creating digital imaging services. The following documents are available:
  - The RLG Worksheet for Estimating Digital Reformatting Costs
  - The RLG Guidelines for Creating a Request for Proposal for Digital Imaging Services
  - The RLG Model Request for Information (RFI)
  - The RLG Model Request for Proposals (RFP)

  Reference: Papers given at the RLG and NPO Preservation Conference *Guidelines for Digital Imaging* (1998): http://www.rlg.org/preserv/joint/

- *RLG Preserving digital information* (August 2002)
  Available formats: HTML; PDF
  http://www.rlg.org/ArchTF/
  The Commission on Preservation and Access (CPA) and RLG formed the Task Force on Archiving of Digital Information, charged with investigating and recommending means to ensure "continued access indefinitely into the future of records stored in digital electronic form". The report is an outcomes of the Task Force

- Anne R. Kenny - Oya Y. Rieger. *RLG Moving theory into practice* (May 2001)

Available formats: HTML; print publication
http://www.rlg.org/preserv/mtip2000.html
The book advocates an integrated approach to digital imaging programs, from selection to access to preservation, with a heavy emphasis on the intersection of institutional, cultural objectives and practical digital applications.

TASI (Technical Advisory Service for Images)

- *Managing Digitization Projects* (2002)
  Available formats: HTML; printed pack
  http://www.tasi.ac.uk/advice/managing/jidi_workflow.html
  Funded by the Joint Information Systems Committee (UK), provides information on creating, storing, and delivering digital image collections. The course includes:
  - Deciding to digitise,
  - Managing the workflow,
  - Managing the project,
  - Looking after copyright, IPR, ethics and data protection,
  - Project Management,
  - Workflow guidelines,
  - Why "Archive Standard"?,
  - Copyright,
  - Coping with copyright,
  - Quick reference copyright guide,
  - Example Licence agreement,
  - JIDI digitization model,
  - Lessons learned from the JIDI project,
  - Risk Assessment,
  - Staff Training.

  Also lists events and information resources of interest to those involved in digital imaging initiatives.

TEI (Text Encoding Initiative)

- C.M. Sperberg-Mc Queen - Lou Bernard. *Guidelines for Electronic Text Encoding and Interchange* (March 2002 - P4 Edition)
  Available formats: XML
  http://www.tei-c.org/P4X/
  A new and corrected version of the TEI Guidelines, XML-compatible, edited by theTEI Consortium (The Association for Computers and the Humanities (ACH); The Association for Computational Linguistics (ACL); The Association for Literary and Linguistic Computing (ALLC). The Guidelines provide means of representing those features of a text which need to be identified explicitly, in order to facilitate processing of the text by computer programs. In particular, they specify a set of markers (or tags) which may be inserted in the electronic representation of the text, in order to mark the text structure and other textual features of interest.

UNESCO/ICA/IFLA

- *Guidelines for digitization projects for collection and holdings in the public domain, particularly those held by libraries and archives* (March 2002)
  Available formats: PDF
  http://www.ifla.org/VII/s19/pubs/digit-guide.pdf
  Guidelines for digitalisation projects including planning and setting up projects, selection, management and production processes. They deal with paper material, manuscripts, printed books and photograps. They are not concerned with digitization programs as an integral part of an institution strategy. They include checklists for each chapter.

University of California Los Angeles UCLA

- Kim Thompson. *Digital projects Guidelines and Standard* (1998)
  Available formats: HTML
  http://www.library.ucsb.edu/ucpag/digselec.html
  The list of criteria is recommended to guide collection development librarians and preservation librarians in selecting collections of analog materials (including paper, film, audio, and video) for conversion to digital format. Some of the criteria are based on conventional selection and preservation considerations common to all formats; others arise from the opportunities and constraints unique to digital technologies.

University of Virginia Library. Electronic Text Center

- *Archival Digital Image Creation* (1996-1997)
  Available formats: HTML
  http://etext.lib.virginia.edu/helpsheets/scanimage.html
  Basic Helpsheets for helping to making decisions. They include:
  - Text Scanning: A Basic Helpsheet,
  - Image Scanning: A Basic Helpsheet ,
  - The Special Collections Department.