

**Paper for the South Asia Library Group - 60th Conference, Public Records Office, Kew, London. 27th November 1998.**

**By Simon Tanner**

It is with great pleasure that I speak with you today. As a Chartered Librarian myself, I appreciate the challenge and expertise required to handle special collections, Government and official publications and newspapers and grey literature, especially in such a complex region as South Asia. I would like today to offer an insight into the work of the Higher Education Digitisation Service (HEDS) and the benefits of digitisation.

My paper is called preserving the past, access for the future and this identifies HEDS main role - which is to assist and provide services to convert any physical media, whether paper, microfilm, photographs or bound volumes for instance, into any of the many electronic formats available. Digitisation is tool in the information revolution that we currently find ourselves immersed in. This information society is about empowerment, i.e. making it possible for new, fresh things to happen and for fresh approaches to information resources. By digitising valued resources we are: opening up new avenues of research; enabling wider potential access and also giving a renewed means of perceiving our information heritage. It is this new approach that assists in preserving the past. I make no claims that the act of digitisation itself is a preservation mechanism, but I do claim that the electronic resources created ease the pressure on the physical artefacts long into the future with the added value of wider access to the resources.

**Introduction to HEDS**

HEDS is a JISC funded service, established by the University of Hertfordshire. HEDS provides a single point of contact to support the conversion into electronic formats through a total management package. Our services are low cost to Higher Education and are also available to non-HEI organisations. We provide advice and initial consultations at no cost, will carry out detailed feasibility studies, write technical specifications and will also manage contracts and digitisation production through to project completion.

There are services, like HEDS, out here to assist with every aspect of digitisation, and you may be considering digitising a collection or some other information resources. So I'd like to address the big issues to consider before embarking on such a project.

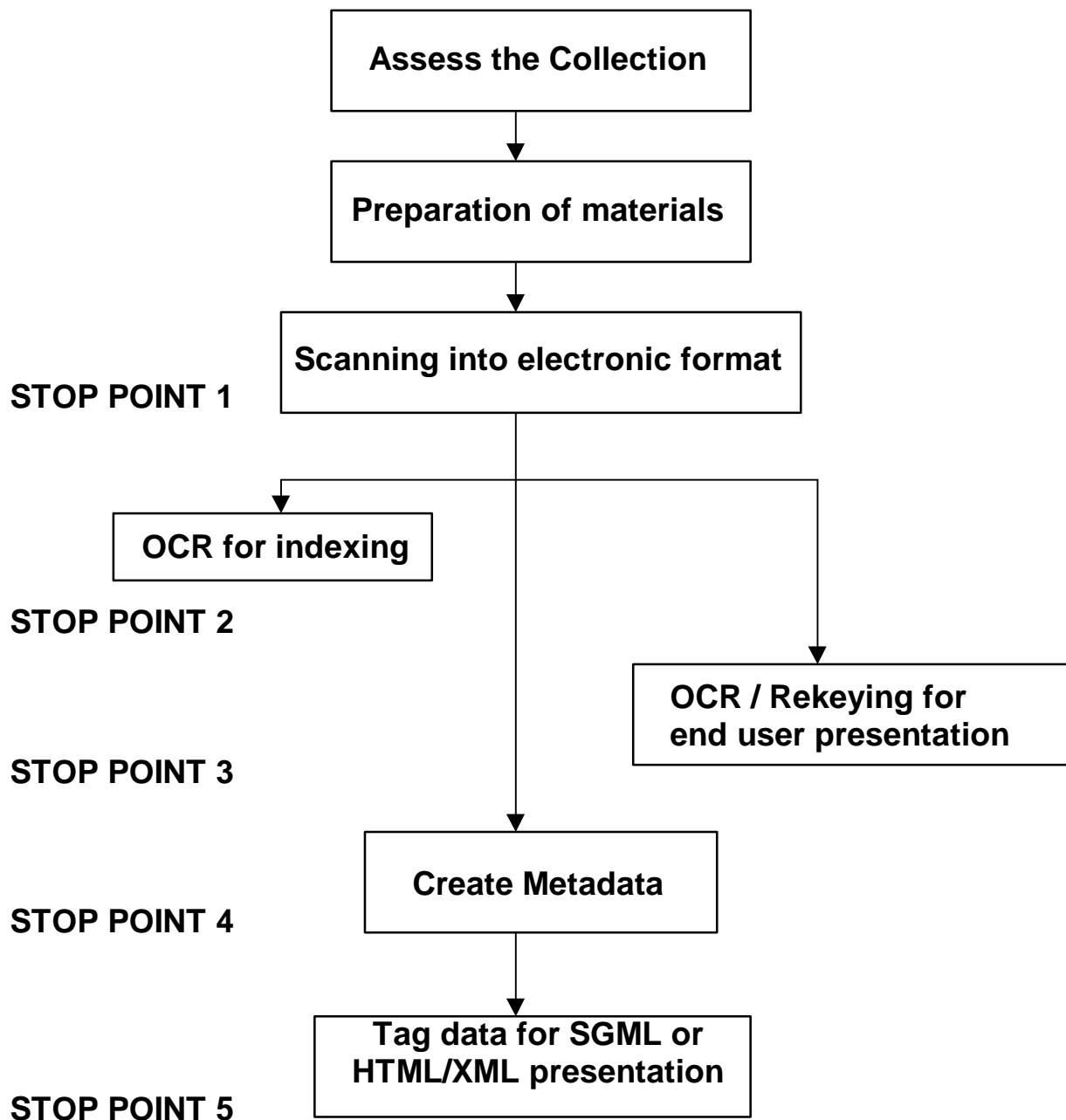
**Digitisation Issues**

When considering digitisation - and especially in the context of official publications, newspapers and grey literature – you may be surprised to note that the biggest challenges presented to HEDS when we advise clients will be very familiar to information workers and archivists. These are preparation of originals for digitisation and organisation of data resources.

Allow me to illustrate this point by briefly describing the digitisation chain of events and the various Stop Points along the way for a typical document collection. Each Stop Point represents a place in the chain where you could stop and say this is all I want to do to achieve my digitisation goals. Also please note this just one chain, there are a number of possible combinations.

**27th November 1998**

**1**



There are 3 things common to all of the processes and every single Stop Point:

1. It is essential to assess the collection to identify its unique characteristics. These might be physical characteristics or is content driven, but these unique features will drive the digitisation mechanism and help define the required provision and access routes to the electronic version. There are no standard digitisation projects and the defining the nature of the original materials to be digitised is the essential first step of any conversion project.
2. Every project requires original materials to be prepared for digitisation and this is often a significant hidden cost to the organisation. Are the materials catalogued, can an inventory of items be provided easily, possibly to the individual page? Are the materials in boxfiles, or are they bound, stapled or loose? They might need separating into individual sheets or removing from the

filing system. At the very least the materials will require movement from the place of storage to the place of digitisation, whether doing work in house or using an external service provider.

3. Organising the electronic version. Whether the end product is one data file or thousands they will all have be given file names and put into some sort of logical directory structure. This is to enable the data to be easily accessed. Often as important as structuring data is to name the data files in such a way as to ensure there is some logical linkage back to the original physical item and that the filenames are preferably unique.

There are many ways to achieve this and HEDS is involved with several journal projects. One such project has every original page as a separate file, whilst in another, every article has a separate file with multiple pages.

So whatever is done in a digitisation project 3 things must happen:

1. Assess the nature of the originals with digitisation in mind.
2. Prepare our collection for the digitisation process.
3. Decide how to organise the end information resource.

This strikes me as what librarians and information workers do every day in different contexts and I am reminded of a quote about the internet from John Rennie, Editor in Chief of Scientific American:

“At some point the Internet has to stop looking like the world’s largest rummage sale. For taming this particular frontier the right people are librarians, not cowboys. The Internet is made of information and nobody knows more about how to order information than librarians”

Rennie, John. Civilising the Internet. Scientific American, March 1997, p6.

These issues and the complexity of the materials involved are most evident in the areas of grey literature, newspapers and official publications.

### **Digitisation Costs**

So why digitise? All of these stages and processes must be adding to the cost of digitisation and the cost may appear to outweigh the benefits. But in digitisation the technology itself is not the main cost element, it’s the amount of human intervention that is required, which push the costs up. When the original is more complex in terms of presentation, preservation handling, preparation and data organisation the costs of digitisation will generally be higher. Sometimes HEDS is asked why digitisation is “so expensive” and I am reminded of this quote from Shaw’s Pygmalion:

“Pickering: Have you no morals, man?

Doolittle: Can’t afford them, Governor.”

George Bernard Shaw - Act II, Pygmalion

We tend to view things as expensive if we do not value them. Digitisation has for too long been associated with either high costs or low grade end products. HEDS is trying to rectify this by providing high grade end products at low realistic costs. When one considers the actual value added to the information resource in terms of wider access, varied services and preservation, then that has to be balanced against the costs for the large amount of work done to convert the original materials. Costs are also dropping over time as the technology advances and opens the door to projects that previously would not have been considered, whilst the standard of the end product is rising at the same time.

Digitisation is a decision that should be made with an eagle eye on the value and benefits gained versus the money to be expended to achieve your project goals.

## The Future of Digitisation

I would like to finish off this paper by expressing my personal views on the future of digitisation. I feel that there are so many wonderful resources that make up the treasure house that is the UK's library and archive system. I hope that many of them will be digitised and made available to a wider audience. To get there the following issues will have to be addressed:

- **Continued investment in the conversion infrastructure.**  
Resources are scarce and funds will be best spent on developing centralised resources or using vendors, rather than every project purchasing equipment with limited lifetimes and funds.
- **Technology Catch Up.**  
The equipment and software needed to achieve the high standards of conversion required to serve the Higher Education communities needs will become available and ever more efficient with less and less human intervention required. This will drive down the cost of conversion whilst increasing the standards of images. This will encourage further projects to digitise materials previously inaccessible to conversion.
- **Maximise value of materials.**  
There is a trend towards high profile, large scale conversion projects that will benefit all of Higher Education in the UK. These are extremely important, but I hope that the funding bodies do not remove funding from the smaller projects with a very focused user group in mind where the costs are low and the value added very high.
- **Planned growth of digital libraries.**  
The Electronic Libraries Programme under JISC in the UK has developed many resources to create digital libraries. However, much of this has come about through the strong interests and work of a few people without institutional review. There is a need to continue to develop the top level planning to resolve digitisation priorities based on experts within disciplines determining the core resources that would be most valuable for teaching.
- **Access Issues and Metadata.**  
The future costs for conversion will continue to drop, but the costs of creating intellectual access and metadata will remain high. It is important to develop appropriate levels of content description to enable access mechanisms if the value of the digital library resources are to be fully realised.