Planning for a Successful Project

Paper for the IT Forum, 16th Feb 2000
Joanne Lomax Smith
Digitisation Consultant
Higher Education Digitisation Service

Introduction

Who HEDS are and what we do
General information about project planning
The digitisation project:
- why digitise?
- how?
- what then?

The Higher Education Digitisation Service

HEDS was established as an eLib project in 1996 and became a full JISC Service in August 1998. The Service was established as a response to the increasing tendency for digitisation projects to buy in equipment and train up staff for ultimately short term projects. Money was being spent over and over on equipment that would be outmoded very quickly and on training staff who would leave at the end of the project and take their experience with them. HEDS therefore offered the opportunity for money to be spent centrally on a Service offering expert advice and digitisation services.

HEDS are not restricted to the HE sector only, any government or non-profit making organisation can use our services.

The service that we offer is twofold - we offer both advice and production. As a starting point a project can take advantage of the free 'First Response' service which means that a HEDS Consultant will give free, no commitment advice by phone or email. If you require it, the First Response service also includes a half-day visit by one of the consultants to view your collection.

No part of First Response commits the project to using HEDS for production services, indeed we often advise projects to use in-house resources if this is the most appropriate manner for digitising the materials in question. We can also help plan your own in-house digitisation unit and advise on equipment and staff training if required as part of the full consultancy service that we also offer.

The production services include the conversion of any kind of originals into any appropriate digital form, according to a detailed specification developed with the project team. This bespoke service means that you will never see a HEDS price list. As explained later in this paper there are many factors involved in setting a price.

A full list of current and recent projects is available on the HEDS web pages at
http://heds.herts.ac.uk/HEDCinfo/Projects.html

Project Planning

Most of what I will say is common sense but we do find that with digitisation projects the 'digitisation' part can become such a focus of attention that the 'project' part is forgotten. IT should really be a tool rather than a purpose, like any other project it is the overall outcome of a digitisation project that should be the important thing.

Project planning has its foundations in three concepts:

- Vision
- Risk management
- Resource management

Ignoring these factors can lead to failure. A study by KPMG into why technology projects fail gave the following reasons:

- 32% - inadequate project management and control
- 20% - lack of communication
- 17% - failure to define objectives
- 17% - lack of familiarity with project scope and complexity
- 14% - incorrect technology, project size & other

All bar 14% of these failures were due to management and project planning issues rather than any failure of the underlying technology.

Vision is about whether you can see the whole picture. Are you seeing the project from a single viewpoint or can you envisage the parts making up the whole? A successful project depends on knowing exactly how the elements of the project fit together although it is not always necessary for every project team member to know the detail of each part of the project.

The project should have clearly defined and solid objectives, goals and deliverables. It is not enough to know that you have, for example, a collection of letters written by a famous author that it would be useful to digitise. Why would they be useful to have, who would use them, would that person have the computer to use them? Knowing the ultimate aim for the project is as important as having a good idea.

This approach could be termed as the Hollywood Pitch - the project should be describable in just a few sentences.

Risk and resource management is about using what's available to you in the most appropriate manner. The first element of this is to clearly define a project model and plan so that everyone knows what they should be doing and when, and what they are working towards. Objectives should be closely defined, for example a calculated target for the conversion process - i.e. 2000 slides scanned by 2 months from the start date calculated at so many per day with however much lead in time rather than 'the scanning will be complete by Christmas'.

A good project manager is essential. Particularly a good project manager who has time to manage the project and is not managing 8 other projects for one day a week each. Good project management is about knowing about the project, not being a figure head who signs the invoices. Plus, who's going to be doing the rest of the project manager's day job?

People should be empowered to do the parts of the project that they are good at. If you have someone who is not so good with technology but an efficient cataloguer, put them on the part of the project that will suit them best, don't force them into a role they are not happy with.

Good communications are vital - regular debriefing sessions keep everyone informed and make sure that any problems can be aired rather than allowed to fester.

A practical point is that equipment should be ordered as early as possible in the project so that the effects of any delays are minimal. This may mean investing time before the project starts in evaluating equipment.

Invest in training - self-tuition in IT is no bad thing, but it takes time. Even if sending someone on a specialist course is expensive, it is likely to be cheaper in the long run than all the time that person would otherwise spend working things out.

Make use of advice; don't proudly struggle on making all the mistakes that other projects have already made. Involve agencies such as HEDS - although preferably from the beginning rather than for last ditch problem solving!

Be honest about potential problems and failures; you cannot learn from something if you refuse to accept it has happened!

The Digitisation Project

Why digitise?

The answer to this question is not 'because everyone else is doing it and we don't want to be left behind'!

The main reasons for HEDS clients to digitise materials are the provision of access and the aiding of preservation of the original.

Access. Creating a digital version of an item and then distributing it will almost undoubtedly improve access. At the most basic level it means that the item can be viewed in multiple locations by multiple users - if a core past exam paper has been hidden or stolen from the library the electronic copy will remain accessible.

Ease of use: Digitisation can also make an item much easier to view - for example, a collection of photographic negatives or microfilm may be underused even though they contain important content because of the difficulties of viewing them - through digitisation these materials can be made more easily accessible.

New life to old materials. Digitisation can also bring new life to out-of-print items - a recent client commissioned HEDS to create a digital version of a key student text which would have been far more expensive to re-publish; the digital version also had the advantage of being searchable and comparatively cheap to reproduce for sale.

Expanding usage without bringing in more users. (debatable) Providing remote access to digital surrogates can also ease pressure on libraries and archives by satisfying enquiries from a distance. However, it may just as easily raise your profile and lead to more visitors

Expanding use to those who are not normally allowed access. Digital versions can also allow an individual access to materials that would not normally be available to the public, for example special collection materials that are normally only accessible to certain researchers could be made more generally available.

Aid preservation through reducing use of the original. Digitisation may be used to aid preservation of the item, notably if the digital version can be made the only or the main method of access to the resource.

Make money? Selling CDs or access to web material. Potentially…

Are you ready for a digitisation project?

A digitisation project on its own is a pretty lonely thing. It should be an integrated part of the service, adding value to those things already in place. A web page with a few pretty pictures from the collection is no substitute to a good finding aid for the whole collection. However, this is often dictated by the money available for projects.

A digitisation project is for life, or should be. Spending large amounts of money on a one-off project is a complete waste if there is no money available to maintain the data at the end of the funded project period.

The Digitisation Project life cycle

As I said before, a digitisation project should be tackled in exactly the same way as any other project. Particularly when it is noted that IT is not the core factor in a digitisation project. Broadly speaking there are the following basic tasks in a digitisation project:

1. Assessment of the need for digitisation.
2. Selection of materials.
3. Deciding what you want to achieve from the information content of the originals. *(what do you want to do with it - what is the final product? Web, CD….)*
4. Deciding how you are going to reach the end product.
5. Finding the funds for the project.
6. Planning the project and assigning resources.
7. Preparing the originals for digitisation.
8. Conversion.
9. Quality assurance checks to ensure the output conforms to specification.
10. Return originals to their place in the collection.
11. Mount data.
12. Make provision for archiving and preserving the data.

One of the single largest factors is **preparation** and this is often overlooked when planning a project. Preparation includes the removal of the materials from their place in the collection and return when the conversion is complete. It includes the movement of materials to the processing area, even if this is internal. Unique identifiers should be given to each item. Items may need cleaning before scanning or disbinding. Copyright clearance is also part of the preparation. This cost may end up being as high as 30% of the project cost.

Factors affecting cost

IT DEPENDS….

As the HEDS cost matrix[1] suggests, there are basic inherent factors associated with each media type which will effect the basic price of converting each item.

Other factors include:

Preparation required - will the items need conservation, cleaning or numbering before scanning
The more preparation is done before the scanning takes place the more efficient it will be.
Condition - can any part of the process be automated
'Sameness' - a large amount of any single media type will be relatively less expensive than the same volume of lots of different media types

Is there a conservation effort planned? Scanning manuscript pages while they are disbound for rebinding is preferable to scanning them as bound volumes.

Can money be found to clean and conserve photographic materials to make digitisation simpler

Know your originals

Knowing your originals well will not only keep surprises to a minimum, it will also make the process more efficient. The types of things that you need to know are:

Photographic media (transparencies, prints, negatives)
- What size are the originals, are they all the same size?
- What proportion of the items have colour content? Is it important to capture the colour?
- What condition are they in - for example, are they dirty from heavy use?
- Are the slides in strips, frames or sleeves?
- Are the photographs flat or have they bowed?

Paper media
- What size are the pages, are all items the same?
- What general condition is the material in?
- Can books that are bound be stripped to loose pages for scanning?
- Is there any artwork - is it black and white, colour or line art?
- Is the text particularly small or large?

Each of these factors may have impact on the scanning treatment and ultimately on the cost.

Think about a Feasibility Study and a Pilot

- Risk Management tools
- Low expenditure to gain measurable results
- Forces decision making
- Provides evidence of full project costs

Provides funding body with evidence you know what you're doing. Contains the mistakes.

[1] The cost matrix is available in the HEDS paper *Digitisation, How Much Does it Really Cost?* at http://heds.herts.ac.uk/HEDCinfo/Papers.html

Scanning in-house vs. using external vendors

The conversion of the materials can, broadly speaking, be done either in-house on specially purchased or existing equipment or sent to any external agency such as HEDS.

The major reasons for sending materials to a bureau for digitisation are that the **originals are not capable of being scanned successfully in-house** (for example bound volumes or microfilm) or that **the intended product is particularly difficult** - for example requiring advanced colour management skills. The type of equipment used for the scanning of items such as bound books or microfilms tends to be so expensive that it would be difficult for a project to justify the expenditure on such equipment, particularly given the short life-span and high maintenance costs of scanning equipment.

Alternatively, the project manager may decide to use in-house resources for several reasons including that

- the collection cannot be moved out of the institution;
- that the collection is badly organised;
- that the digitisation needs to be phased in small amounts over a long period;
- or that the digitisation task is very simple.

It may also be that the project can call on existing staff knowledge and equipment which would mean the project could be done in-house with limited further capital expenditure.

There are some base line infrastructure requirements to in-house digitisation:

- A robust production level scanner.
- A powerful PC with lots of memory (at least 256Mb RAM).
- Plenty of system resources such as backup and write to media (e.g. CDROM) capacity.
- Software to assist the digitisation.
- Experienced staff to run the equipment and staff to oversee the process and QA
- Basic equipment set-up costs: £3-4000 for paper, £6-8,000 for photographs

This is assuming that the in-house operation wants to approach anywhere near the unit price of production available from outside agencies such as HEDS.

A further reason that many projects are undertaken in-house is that the staff time, overheads and some consumables such as file storage can often be swallowed up by the institution and do not become apparent as a costed factor of the project, thus making this appear to be a cheaper option than out-sourcing.

There is no easy answer to the question of whether to scan in-house or to out-source because it depends so closely on the project team, the institution and the materials.

Archive quality vs quick and dirty….

The digitisation mantra should be **Scan Once for All Purposes.** This would mean that all that costly human effort in preparation and project management would not be repeated in years to come when the images were no longer good enough to be used, for example, on new generation web browsers or were tied into old formats that could not be upgraded.

For this reason HEDS recommends scanning at the highest resolution and bit depth that is appropriate and affordable. This will create a master image file that can be used to create any number of surrogates for the various requirements. Kept properly this image file should be accessible for years to come.

However, we recognise that there is not always the money available to create and store the most high quality images. If the only use you can ever foresee for your documents is as illustrations on your web pages then a format like PhotoCD might not be such a bad option. However, should you ever want to come back to these images and use them for another purpose such as high quality colour printing they may not be suitable.

There is also a growing move that the digitisation of high volume, low cost materials should not necessarily receive the same 'scan once' treatment as more 'high value' materials. It may be more cost effective to scan to the level currently required and then scan again in a few years time if necessary.

Similarly, where the digitisation is not from the original but is from a surrogate such as microfilm it may be that it would be cheaper to rescan the microfilm if need be rather than preserve and regenerate the digital files.

Digital Preservation

This leads into digital preservation: making sure that anything you digitise now will stay accessible for as long into the future as possible. The main things we can do to mitigate against digital files becoming obsolete are:

- use a non-proprietary, open file format such as TIFF
- Save files uncompressed where possible, or use lossless compression if not
- keep geographically separate copies of the data
- do not rely on a single media type
- **budget for a regular review of the data, upgrading and moving data where necessary. If any of the above is disregarded then the review period must be made more frequent**

Some references:
Project costs:

HEDS Paper *Digitisation: How Much Does it Really Cost?*
http://heds.herts.ac.uk/HEDCinfo/Papers.html

HEDS recommends Dr Stuart Lee's scoping study for the Oxford Digital Library, which includes real-life project costs (including some HEDS projects) and comparisons on in-house versus outsource costs: http://www.bodley.ox.ac.uk/scoping/

Contact Details

Joanne Lomax Smith
Email: heds@herts.ac.uk
Web: http://heds.herts.ac.uk

Quote

"I don't see much sense in that," said Rabbit
"No" said Pooh, humbly, "there isn't. But there was *going* to be when I began it. It's just that something happened to it along the way".