


Digitisation: Workflows, Costing and management

Dr. Shalini R. Urs
Professor
Department of Library and Information Science
University of Mysore
Shalini@vidyanidhi.org.in

15-20 April 2002 Digitization: workflows, costing and management PI-7 1



Digital Content and collections

- Building digital libraries begins with creating digital content and collections
- Digital content may be created in the following two ways-
 - Born digital
 - Digitisation

15-20 April 2002 Digitization: workflows, costing and management PI-7 2



Digitisation

- What is digitisation?
- Creating a digital object by converting a source object in analog form
- What is a digital object?
- Digital objects are individual entities in a digital collection
- Any object which can be digitised can be a digital object



Digitisation: Processes

- Identification of the items for the collection
- Selection of the 'kind' of content
- Choice of formats
- Choice of hardware
- Choice of software
- Selection of medium
- Storage and archiving
- Management



What to digitize?

- A documented collection development policy
- Identification of the 'domain'
 - scope
- Selection criteria
 - List the criteria
- Estimate the size of content and collection
- Growth rate



Digitisation: processes

- Scanning
- OCRing
- Conversion to the chosen format



Digitisation:Workflow

- Original analog document procurement
- Identification of batches
- Scanning
- Clean up and Editing
- Creation of metadata
- organising the scanned files
- Naming the files
- Back up



costing

- Costing is a very complex issue as it depends on many variable
- The variables include-
 - **Source object**
 - **Hardware**
 - **Software**
 - **Manpower**



Costing...

- Digitization cost maybe categorized as
 - Manpower cost
 - Process cost
 - Overheads



Manpower costs

- Costing -an exercise
- Manpower costs are worked out on the basis of unit 'time' per task
- On an average one person can scan 100-150 pages per day
- One day is calculated as 8 working hours (effective 6 hours)
- Thus we say it costs per page approx 3+mins



Manpower costs...

- A manpower cost is worked out on the basis of the following assumptions-
- A monthly pay of Rs. 3500
- A working month consists of 25 days
- In other words 2500 pages per person per month
- Per page cost= $25000 / 3500$



Process cost

- Process cost is based on the hardware and software costs
- Two main hardware items are –
 - Computer
 - Scanner
- Software cost
- The cost of purchasing the software



Process costing...

- Assumptions
 - A computer costing Rs.45,000
 - A Scanner costing Rs.25,000
 - Software costing Rs.10,000
- Total investment cost of Rs.80,000
- A life of 3 years



Process costing

- Based on the assumption of 3 years of life
- The investment can result in $25 \times 12 \times 3$ working days
- This results in $100 \times 25 \times 12 \times 3$ pages of digitisation
- The process cost per page works out to be



Costing

- Total cost per page = manpower cost + process cost
- Thus per page it works out to be Rs.
- We have excluded the overheads such as
 - Electricity costs
 - Space cost
 - Other indirect costs such as supervision, general housekeeping etc



Digitisation management

- The quality of digitization depends on several factors
 - The condition of the source document
 - Such factors as coloring of the source document
 - Font type and size
 - Person doing the scanning
 - Un known factors