

## Description of the Solution.

What does it take to start a digitizing job?

Before attempting to start a digitizing job the following are the questions that need to be answered and the appropriate requirements laid out

1. Analysis of the physical media and to be digitized
2. Appropriate hardware, i.e. scanners, workstations, servers etc
3. Standards for digitization
  - a. Digital file formats and standards
  - b. Meta data requirements and standards
    - i. Document meta data
    - ii. Structural meta data
    - iii. Administrative meta data
  - c. Naming conventions for files and folders
  - d. Search criteria and keyword analysis
4. Post Scanning Requirements
  - a. Software for image enhancements
  - b. OCR requirements
  - c. Indexing requirements for cross-linking and direct linking of document
5. Security mechanisms
  - a. For physical media
  - b. Digital media
6. Archiving systems
7. Standards for master images
8. QA criteria
9. Rendering mechanisms

While all the above requirements are not exhaustive and the complexity can increase or decrease depending on the kind of media and the end use of digitization the key factor that needs to be addressed is stringing these requirements together to create a seamless solution.

For such a job, it is first required that once the standards have been formulated and the requirements frozen, like any other production process a work mechanism is designed, rolled out, adhered and monitored. It is extremely critical in large jobs that this is standardized and adhered as redoing a job is many fold more expensive than getting it right the first time. For anyone from the stand point of a person executing a medium to large digitization whether in-house or through a service bureau following are the some of the pain areas envisaged.

1. Am I doing my job as per schedule?
2. Can I give my client or management a clear picture of the progress of the project?
3. Are the standards being met

4. Is everyone in the system doing what he has been contracted to do?
5. Am I using the right tools to achieve the objective?
6. Are there a proper check and a QA process in place?
7. How do I optimize resources and reduce the fatigue of system?
8. Can I effectively manage the physical documents/digital documents?
9. Am I in a position to have a validated billing system?
10. In case of a rework will I be able to effectively manage the situation

The Digiflo the workflow management software from Thrinaina Informatics is configurable and integrable solution that addresses the requirements and the concerns of Projects, Project managers and clients which have been listed above and many more. The tasks involved in a digitization workflow may seem simple and mundane, but considering that we need to string the data done across many processes and people at high volumes it presents a great challenge. Repetitive tasks and processes can be configured and system ensures accuracy in operations by reduction in manual intervention.

The workflow also ensures remote monitoring of the project and a facility to move some of the processes to other physical locations from where jobs like post scan processing and indexing may be done seamlessly. This results in helping one reduce the stresses due to paucity of space, people and other resources and facilitates outsourcing but still ensuring that the defined procedures and processes are adhered to.

This software has been designed, tested, and implemented across prestigious projects and has demonstrated tremendous increase in productivity, reliability, efficiency and efficacy of operations and personnel.

As an offering the workflow software comes as software that is configurable according to the requirements of the project. Further a tool kit is integrated which helps in executing basic jobs which can become complicated in their absence. Software for images processing, OCR, image acquisition can be taken by choice and these can be integrated with the software.

The software comes with a large suite of MIS reports to track the media, digital images, quality and personnel and the efficiency of their work. These reports can also be viewed by the client or the manager/management across the web giving them a perspective and progress of the project. This avoids communication gaps in the status as one has a direct online track of the progress. A Document Format Auditor is integrated for the QA and its output a page-wise report forms a part of the end deliverable. Mechanisms and means to create indexes, metadata, structural metadata and administrative metadata are incorporated and can be customized depending on the requirements of the project.

A brief description below illustrates the various features and the capabilities of the workflow software.

## DIGIFLO AND DIGITOOL

**DigiFlo and DigiTool** is a workflow and a toolbox, that aid in digitization production facilities. They ensure quality at all levels of the work flow enhance ensuring manageability, scalability and efficiency of operations. It gives a facility to integrate OEM software for scanning, processing and OCR and gives a seamless usage in the workflow. Remote access of the site details can be made possible with the client update with the happenings of his project.

Work Flow Management Software Digiflo™ ensures better management of work processes addressing the following aspects leading to the qualitatively superior output of targeted quantity, consistently.

### Features

- Reduction in Turnaround time
- Data capturing at various stages
- MIS and administration including reporting
- Quality checks at appropriate stages
- Identifying the bottlenecks to get them addressed & solved
- Managing the books and data
- Managing both Static Media and Digital Media

### Automated work-flow management from Thrinaina

Work Flow Management Software **Digiflo™** was designed to increase the efficiency of work, Manpower, Quality, Quantity and reducing the cost of ownership of Projection, Reducing manual errors, rework and reducing people dependency.

### 1.Scope

The scope of workflow management includes.

- Creation of a predefined workflow
- Creation of predefined file/folder naming conventions for storage

- Creation of Jobs profiles for different projects
- Tracking of physical documents and media
- Creation of metadata and tagging to the digital documents
- Integrated Document format Auditor to check quality of scanned documents
- Storing of the locations of the digital documents at various stages of workflow
- Creation of document index and structural metadata of documents
- Comprehensive MIS for planning production and measuring the efficiency of the persons associated with project.
- Tracking of documents and processes at various stages of the workflow.
- Managing the attendances and the productivity of the personnel

### Configurations of the workflow

- Scanning jobs can be configured based on the parameters of scanning
- File naming and folder naming conventions can be defined for a job. (Can also be taken from the metadata entered.
- Storage locations of the files and folders at the various stages of the work flow can be defined
- Retrieval of files and folders for subsequent processes possible

## 2. Overview

### Functionalities of Work Flow Management Software Digiflo™

- Tracking of document information in all stages
- Tracking of Man Power
- Tracking the quality in each stage
- Tracking of time taken for each stage
- Generating respective reports like Daily Report, Shift-Wise Reports, Operator-wise Reports etc
- Tools to help the management in training people.

### 3. Modules of the Product

The modules involved in WorkFlow Management are:

**Pre-Scanning:** This involves the management and assignment of roles and permissions for the various users of the workflow, and metadata creation.

**Scanning:** Issue of Documents for scanning and pushing in the scanner settings defined for the job the metadata can also invoke the file and folder naming mechanisms. The module captures the details of the start and completion of scanning, time stamps of files no of pages scanned and a facility for the operator to enter any discrepancies in documents into the metadata. The module has a comprehensive MIS to track the documents scanned and to measure the efficiency of the scanning setup.

**Cropping:** In a production environment to subject a group of files for a automated batch process it is essential that the peripheral noise and artifacts are removed from the images and then consigned for the subsequent stage. A cropping tool (DigiTool) is incorporated which is fast and convenient to us.

**Processing:** This is normally the next step in a scan work flow which would involve the usage of various software for image enhancement, like deskewing, despecking, noise reduction, resizing padding etc. Digiflo has a facility to integrate image-processing tools into the workflow and the metadata of the stage is stored.

**OCR:** Some jobs would require an OCR to be done and a number of OCR engines can be integrated with the workflow ensuring a seamless environment.

**Quality Control:** A multileveled quality control is configurable so that the images with errors are trapped in the early stages. This also utilizes a

Document format Auditor (DigiTool) that generates a log file on the status of the book highlighting the parameters of the scanned documents, which are not as per the job settings. The file/folder naming conventions and a comparison of pages scanned vs. the metadata pages can also be done. A viewer is also integrated with a facility to put in remarks for rejecting the files/ folders at the various QC stages.

**Post Quality Control:** It is ensured that in this stage all the deliverables are according to the client's requirement. The file formats, parameters, metadata are ensured for consistency and QC reports are appended into the folders. At this point the digital data is ready for delivery.

**MIS:** Through this module the books, digital data, QC processes and the people can be tracked and appropriate controls, measures and decisions can be made based on the numerous reports that the workflow software provides.

**Note:** Digiflo and DigiTool currently do not have any native image enhancing and OCR tools but can be integrated with any such software.

## DIGITOO

**Features:** A toolbox of utilities to enhance the productivity and the quality of work in a digitization environment. Following are utilities that are currently available. The tools have been created keeping in mind the specific and typical requirements of a production job. This reducing the cost of ownership of the solution and ensures that the specific tools are available at a click of a button.

1. Document format Auditor: This ensures parametric and objective evaluation of documents and sends a page wise output to an excel or a text file. Can be used in a batch mode for a number of books or for single books.
2. Crop Tool: A quick and simple tool for cropping and removing artifacts from images manually
3. Image viewer: An image viewer with a facility to create structural metadata and also report errors in case they are encountered.
4. File size Verifier: It is normally a problem to handle files that are too large and it could be lank in case too small. It helps identify the files of abnormal sizes for separate intervention.
5. Page renumbering tool: Tools for renumbering pages after insertion or deletion of pages.
6. Image split and Merging tool: Pages are sometimes scanned in a spread mode and may need to be split up for ensuring the sizes are consistent with a subsequent facility to stitch up the pages with a definable gutter size. This tool helps in doing this.
7. Importing and exporting of metadata from XML to Databases and vice versa.
8. File counter: Gives a count of files when delivered in multiple formats to ensure that there is parity and consistency of files.