BEING IN CONTROL:
HOW TO CREATE A NEW DIGITISATION PROCESS AND MONITOR THE WORKFLOWS?

The National Library of Finland
Tiina Hölttä
CCS Content Conversion Specialists GmbH
Claus Gravenhorst
6.10.2011

LIBER-EBLIDA
Workshop on Digitisation of Library Material in Europe
5. – 7.10.2011 the Hague
Table of contents

Background (TH)
Methods
Working together – developing a new process and monitoring tools
Goals
Ephemera feature specifications for three digitisation systems (RIT, ScanClient and DW)
Workflow and Item Tracking
Workflow Overview
Ephemera Workflow and Processing (collection based process) (CG)
Remote Item Tracking (1)
Remote Item Tracking (2)
Remote Item Tracking (3)
Remote Item Tracking (4)
Item Tracking at a glance
Practical innovations and tools made during developing the process (TH)
Conclusions behind the success of establishing the digitisation process for Ephemera
National Library of Finland (NLF) has in recent years built up its mass digitisation processes for different materials using external fundings.

NLF realized that it must at any time be able to know what material is being digitised, where it’s located and what step is in progress.

The Ephemera project (years 2008 – 2012) funded by the European Union and the European Social Fund was established to develop a digitisation process and functional workflows with metadata expertise for extremely heterogenous and mostly uncatalogued small printings (e.g. brochures, price lists, leaflets).
Methods

- First we made a survey of harvesting several Libraries seeking out ways and models other Libraries have done.

- Second we studied old documentation made in the past and gathered information created previously (historical analysis) by interviewing experts => creating a comprehensive written documentation of the expert’s tacit knowledge.

- Third we established various working groups across organizational boundaries in the NLF to address the special issues.

- Fourth we gathered systematic all information regarding the need for changes in the digitisation systems.

- Fifth we had weekly conference calls with the supplier of the production data systems (NLF and CCS).
Working together – developing a new process and monitoring tools
Goals

=> Integrating cataloguing to the digitisation process and recording technical metadata during it

- reason: overall economy and efficiency

Used developmental work research methods such as adapting change laboratory and knot working techniques

- Using mirror material
  - Videos, recodings, old documents as a reflector material by applying the process of planning

- Joint process dates
  - Experts discussing, solving problems and designing the new workflow and process

=> Expanding learning
Ephemera feature specifications for three digitisation systems (RIT, ScanClient and DW)

- Information system requirement specification need emerged directly from the workflow and process documentation done (action)

- A few main features needed by the NLF:
  - Possibility to mark and separate documents during scanning of the box (ScanClient – scanning utility)
  - New printout (Inventory list, Remote Item Tracking)
  - Genre list (docWORKS post processing system)
  - Customized MODS-form (docWORKS) for cataloguing
Workflow and Item Tracking

- **Situation:** Decentralized conversion workflow
  - Material for digitization assembled at the library
  - Scanning of (printed) material at a remote location
  - Conversion workflow at another location
  - Material to be returned to library

- **Goal:** Seamless overview/control of all media passing the digitization process
  - Starts with creating a new batch of material at the library
  - Ends with returning the batch to the library after scanning/conversion
  - Gapless Status & Location Tracking at a glance
  - Individuals at different locations with access to the same web interface (project managers, scan operators, QA operators ...)

THE NATIONAL LIBRARY OF FINLAND
Workflow Overview

• Helsinki delivers Ephemera documents and provides batch and item information
• Inspection and item check-in at Mikkeli
• Scanning of items using various devices
• Conversion of scanned images
• Interactive verification (QA)
• Automated validation and final QA on digital output
• Delivery/Transportation to Helsinki
• Need to track items at every stage
Ephemera Workflow and Processing
(collection based process)

Figure: Jukka Kervinen, NLF
Remote Item Tracking (1)

RIT system is tracking a:

- **Project**
  - has a source location (where media is stored)
  - has a project base type (monograph, newspaper, etc.)
  - can be tracking only (no digitization or conversion involved)

- **Batch**
  - collection of items (books, newspapers, manuscripts, etc.)
  - assembled for transportation to digitization center

- **Physical Unit**
  - Single item (e.g. book, microfilm reel, box of parchments)
  - Unique ID / Barcode

- **Logical Unit**
  - Smallest unit in item tracking
  - can be book chapter, newspaper issue, part of a series, …

Quick launch menu for all functions

One tab for each project

Main Frame

Current project

Current function

Current project

One tab for each project

Main Frame
Remote Item Tracking (2)

Source provider (library) workflow:

- Create Batch
- Edit Batch
  - add physical units (boxes)
  - finalize batch after all items have been added
- Prepare Batch for shipment
- Ship batch to digitization center
- Receive batch from digitization center
- Verify content of received batch
Remote Item Tracking (3)

Digitization Center workflow

• Receive batch from Library
Remote Item Tracking (4)

Monitoring

- View project status
- View batch status
- View physical unit status
- View docWorks document status per physical unit

Statistical Evaluation

- View current workflow snapshot for a project
- View recent activities for a specified time period
- View scanning statistics per scanner/operator
Item Tracking at a glance

• Let’s you manage, monitor and analyze mid- and large-scale digitization projects
• Supports decentralized workflows spread over various locations
• Tracks projects, batches, physical and logical units
• Offers extensive monitoring and statistical evaluation
• Easy to integrate with the docWorks digitization/conversion workflow
• Enables efficient project execution and crisis management
Practical innovations and tools made during developing the process

A tool to get single Ephemera publications catalogued: MODS-cataloguing form customized for Ephemera material

- a key feature, when it comes to creating search options on a user interface

New viable ideas were generated along the way, such as the inventory list

- practical tool for the collection management
Conclusions behind the success of establishing the digitisation process for Ephemera

- The **systematic formal documentation** from the very beginning helps in managing the entire development process.
- **Learning by doing** is the best teacher.
- Numerous experts building the workflow over functional units and organisations and the **entity lead by the project organisation**
- Genuine **enthusiasm**, a touch of **insanity** and **commitment** to resolve the difficult issues are key factors for the success.
- Precise written **requirement specifications with process descriptions** done for the data system supplier.

⇒ Well-planned processes and data systems with practical tools are essential for mass digitisation and logistic management.
Thank You!

Tiina Hölttä
tiina.holtta@helsinki.fi

Claus Gravenhorst
c.gravenhorst@content-conversion.com