

# **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

# Background

- 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)



Google maps

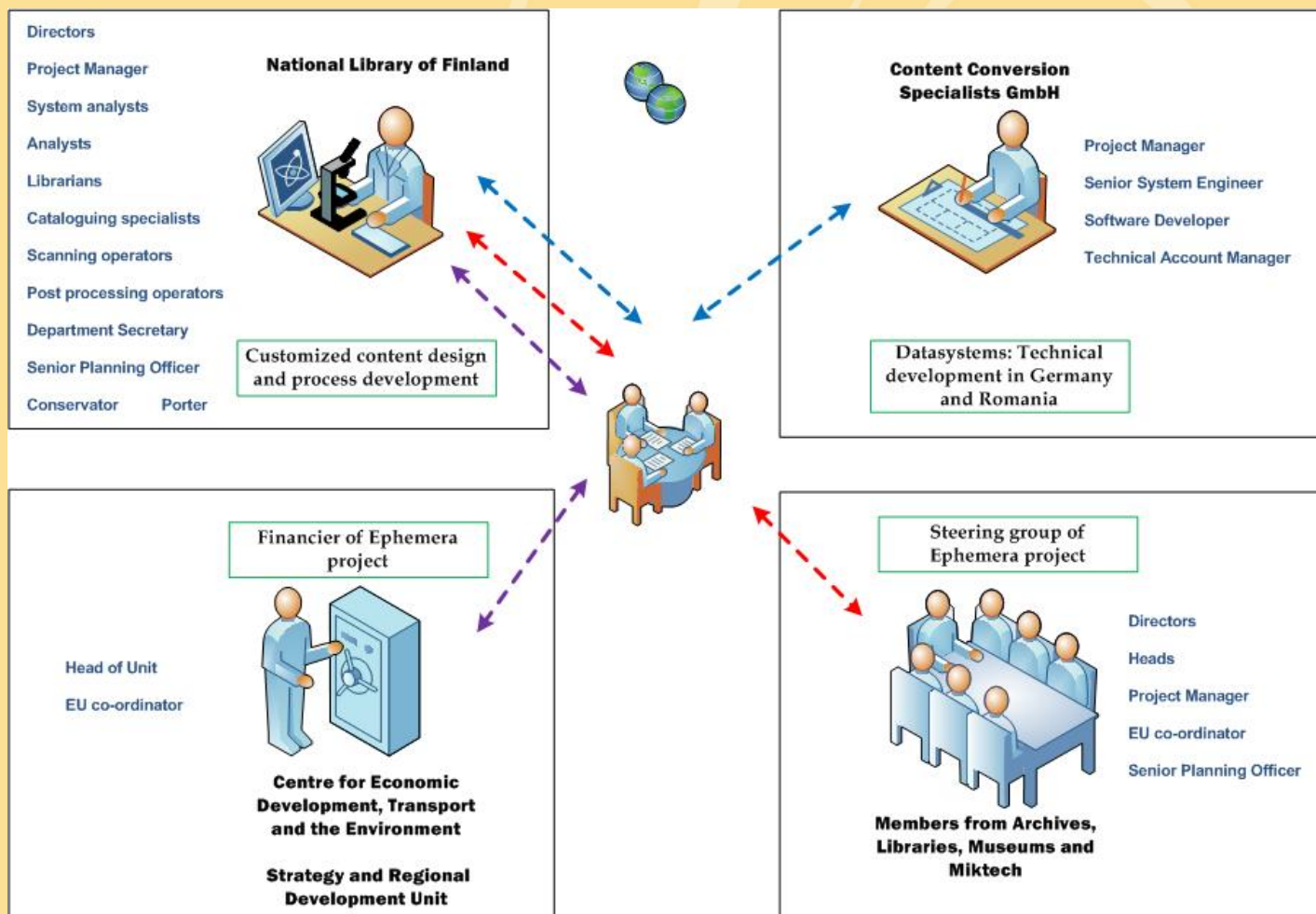


## Methods

- 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

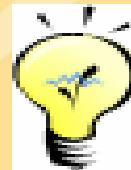
### **=> Integrating cataloguing to the digitisation process and recording technical metadata during it**

- reason: overall economy and efficiency

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

- Using mirror material
  - Videos, recordings, 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

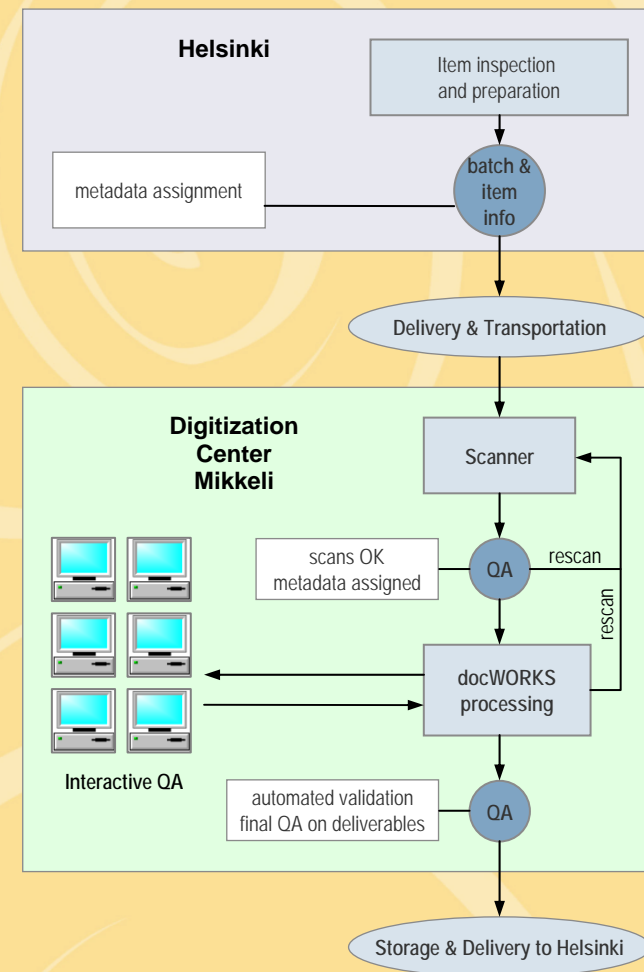
# Workflow and Item Tracking

- **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 ...)



# 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)

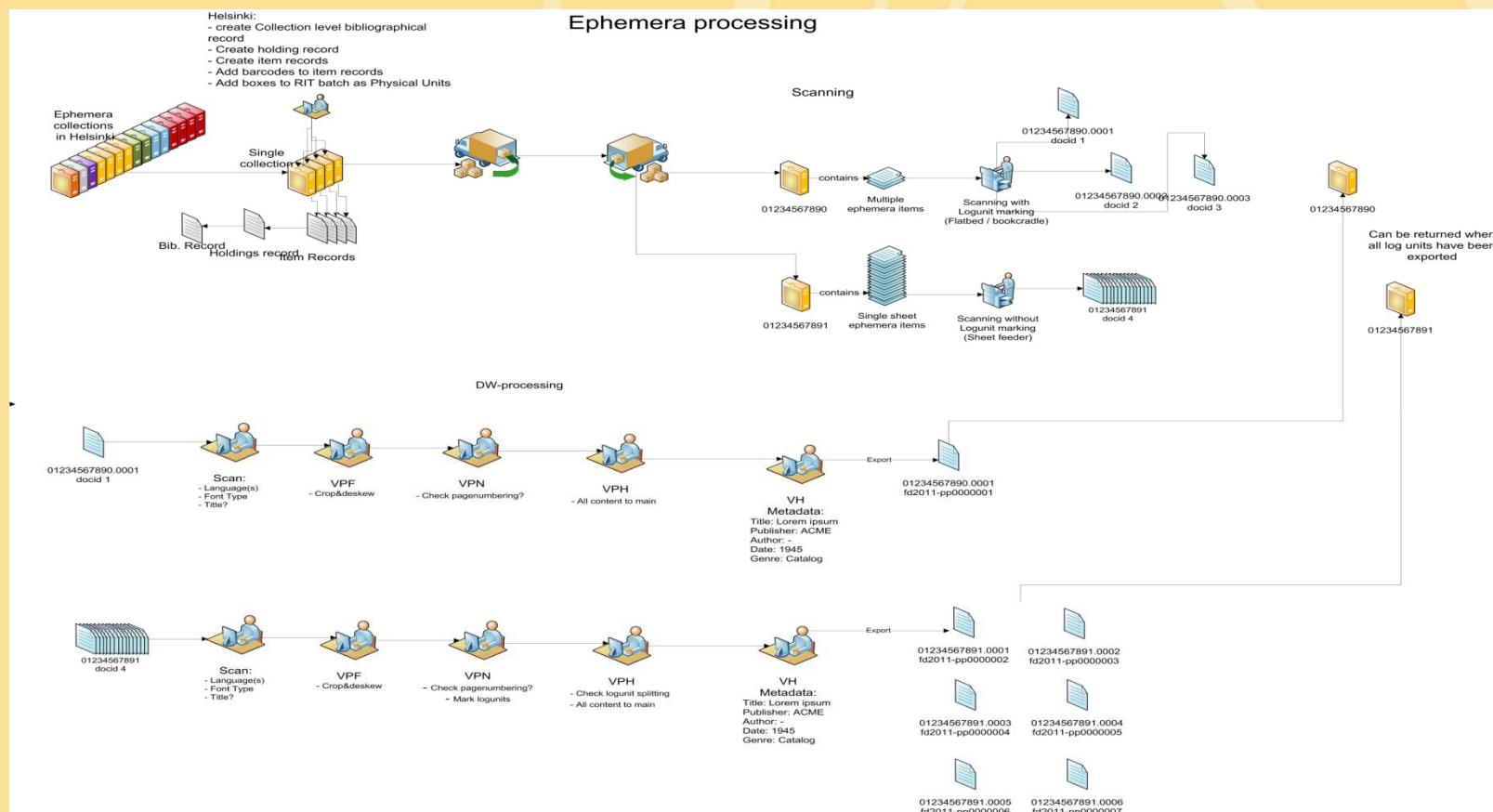


Figure: Jukka Kervinen, NLF

# Remote Item Tracking (1)

RIT system is tracking a:

Current function

One tab for each project

Main Frame

Quick launch menu for all functions

NLF Remote Item Tracking System

Teologia

Home Hello World kaunokirjallisuus RIT Teologia TravelEuropeana Vaitos Site

Welcome DW\administrator | 2010/

This Site

Project Status

Project Status Batch Status Phys.Unit Status Phys.Unit Overview

All batches currently active in this project

	Batch Name	Created By	Created On	Current Location	Current Status	Days in Production	Physical Units	To be scanned	To be processed	Cataloging needed	In Conservation	Ready Re
X	TEOLOG-2009021	DW\osmot	05.01.2010	Helsinki	Unloaded	37	51	0	0	0	0	0
X	TEOLOG-2010001	DW\pasiko	13.01.2010	Mikkeli	Processing	38	52	2	31	0	0	17
X	TEOLOG-2009020	DW\lampipee	13.01.2010	Mikkeli	Shipping back from production	42	47	0	0	0	0	0
X	TEOLOG-2009019	DW\osmot	14.01.2010	Mikkeli	Processing	41	61	27	11	0	0	22
X	TEOLOG-2009-0018	DW\osmot	15.01.2010	Mikkeli	Processing	41	57	23	0	0	2	34
X	TEOLOG-2009-0017	DW\osmot	15.01.2010	Mikkeli	Processing	41	63	0	13	0	4	50
X	TEOLOG-2009-0016	DW\osmot	18.01.2010	Mikkeli	Processing	37	43	37	3	0	0	3
X	TEOLOG-2009-0015	DW\osmot	19.01.2010	Mikkeli	Processing	37	52	44	5	0	0	3
X	TEOLOG-2009-0014	DW\osmot	19.01.2010	Mikkeli	Processing	37	67	67	0	0	0	0
X	TEOLOG-2009-0013	DW\osmot	19.01.2010	Mikkeli	Processing	37	64	20	16	0	0	28

# Remote Item Tracking (2)

## Source provider (library)

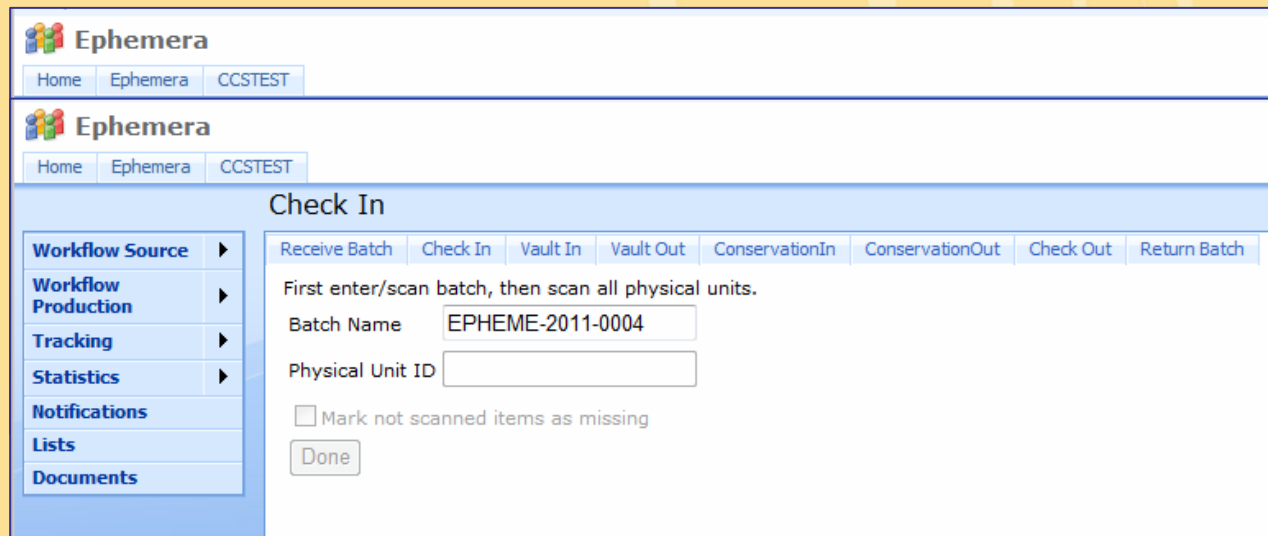
- Create Batch
- Edit Batch
  - add physical unit
  - finalize batch after
- Prepare Batch for
- Ship batch to digit
- Receive batch from
- Verify content of r

The screenshot displays the Ephemera web application interface. The main window is titled 'Edit Batch' and includes a sidebar with navigation options: Workflow Source, Workflow Production, Tracking, Statistics, Notifications, Lists, and Documents. The main content area shows a 'Create new Batch' button, a dropdown menu for selecting a batch (currently showing 'EPHEME-2011-0004'), and buttons for 'Add Phys. Units' and 'Finalize Batch'. Below this, a table titled 'All physical units in this batch' is shown, with columns for ID, Added, Quality Level, Comment Codes, Comment, Scanning Mode, Color Mode, Resolution, and Log. Unit Count. The table currently displays 'No rows returned...'. A modal window titled 'Add Physical Units to Batch EPHIME-2011-0004' is open, showing fields for Project (Ephemera), Batch (EPHEME-2011-0004), Phys. Unit ID (1201011122), Location (Customer), Signature, Quality Level, Comments, Comment Codes (BLP, CAT, CON, COP), and checkboxes for 'Valuable Item' and 'Create Barcode'. The modal also includes buttons for 'Get Bib. Data', 'Try all sources', 'Scanning Mode', 'Color Mode', 'Resolution', 'Origin', 'Reset Fields', 'Create & Next', and 'Close'.

# Remote Item Tracking (3)

## Digitization Center workflow

- Receive batch from Library



The screenshot displays the Ephemera web application interface. At the top, there is a header with the Ephemera logo and navigation tabs for 'Home', 'Ephemera', and 'CCSTEST'. Below this, a sidebar on the left contains a menu with options: 'Workflow Source', 'Workflow Production', 'Tracking', 'Statistics', 'Notifications', 'Lists', and 'Documents'. The main content area is titled 'Check In' and features a sub-header with tabs: 'Receive Batch', 'Check In', 'Vault In', 'Vault Out', 'ConservationIn', 'ConservationOut', 'Check Out', and 'Return Batch'. The 'Check In' tab is active. The main content area contains the following text and form elements:

First enter/scan batch, then scan all physical units.

Batch Name

Physical Unit ID

☐ Mark not scanned items as missing

# Remote Item Tracking (4)

## Monitoring

- View project status

The screenshot displays the Ephemera web application interface. The top navigation bar includes 'Home', 'Ephemera', and 'CCSTEST'. The main content area is divided into two primary sections: 'Project Status' on the left and 'Recent Statistics' on the right.

**Project Status Section:**

- Project Name:** A text input field.
- Remarks:** A text area.
- Filters:**
  - ☐ Where 'Current Status' =
  - ☐ Where 'Current Location' =
  - ☐ Where 'Created By' =
  - ☐ Show hidden(deleted) batches
- Apply Filter:** A button.
- Batch List:** A table with columns 'Batch Name' and 'Created By'.
 

Batch Name	Created By
EPHEME-2011-0003	DOCWORKS\pasikost
EPHEME-2011-0002	DOCWORKS\tiholtta
EPHEME-2011-0001	DOCWORKS\lampipee
EPHEME-2011-0004	DOCWORKS\lampipee

**Recent Statistics Section:**

- Workflow Source:** A dropdown menu with options: Workflow Source, Workflow Production, Tracking, Statistics, Notifications, Lists, and Documents.
- Recent Statistics:** A section with tabs for 'Current Statistics', 'Recent Statistics', and 'Scanning Statistics'.
  - Timeframe:**
    - Show data from:
      - ☒ last week
      - ☐ last month
      - ☐ total
      - ☐ individual timeframe
    - Timespan per row:
      - ☒ day
      - ☐ week
      - ☐ month
    - from: [text input]
    - to: [text input]
    - Refresh:** A button.
  - Filter:**
    - Filter by batch...
      - EPHEME-2011-0003
      - EPHEME-2011-0002
      - EPHEME-2011-0001
      - EPHEME-2011-0004
- Statistics Table:**

Date / Timeframe	Items created	Items shipped	Items scanned	Items processed	Items returned
Wed 31.08.2011				1	
Tue 30.08.2011			1	1	
Mon 29.08.2011					
Sun 28.08.2011					
Sat 27.08.2011					
Fri 26.08.2011			1		
Thu 25.08.2011			1		
Wed 24.08.2011			1		
Tue 23.08.2011				1	
Mon 22.08.2011					
<b>Sum:</b>	<b>0</b>	<b>0</b>	<b>4</b>	<b>3</b>	<b>0</b>
- Summary:**
  - returned: [text input] 0
  - missing: [text input] 0
  - TOTAL:** [text input] 7



# 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**

The screenshot shows a web-based form titled 'docWORKS - [ 491 / Ephemera / Ephemera / Ephemera\1201016915\0002 ]'. The form is divided into two main sections: a left sidebar with a tree view of MODS elements and a main content area on the right. The tree view includes categories like 'docContent', 'Hierarchy', 'MODS', and 'Content'. The main content area contains a form for 'Fishing in Finland' with fields for 'name', 'language', 'place', 'publisher', 'dateIssued', 'dateCreated', 'dateCaptured', 'dateValid', 'dateModified', 'copyrightDate', 'dateOther', 'identifier', 'relatedItem', 'extension', 'note', 'genre', 'subject', 'authority', 'Basic:topic', 'Basic:geographic', and 'Content'. A preview of the book cover 'FISHING IN Finland' is shown on the right side of the form.

## KOTELON SISÄLLÖN INVENTAARILISTA



NBN: FPP2011001

Signum: Valtionhallinto IA

Nimeke: Kansanhuolto, 1939-1949.

Kotelo: 43

Määrä: 221 kpl

# Conclusions behind the success of establishing the digitisation process for Ephemera

- The systematic formal documentation from the very beginning helps in manage 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.holtt@helsinki.fi](mailto:tiina.holtt@helsinki.fi)**

**Claus Gravenhorst**

**[c.gravenhorst@content-conversion.com](mailto:c.gravenhorst@content-conversion.com)**