



Digitisation at NLF and the National Digital Library

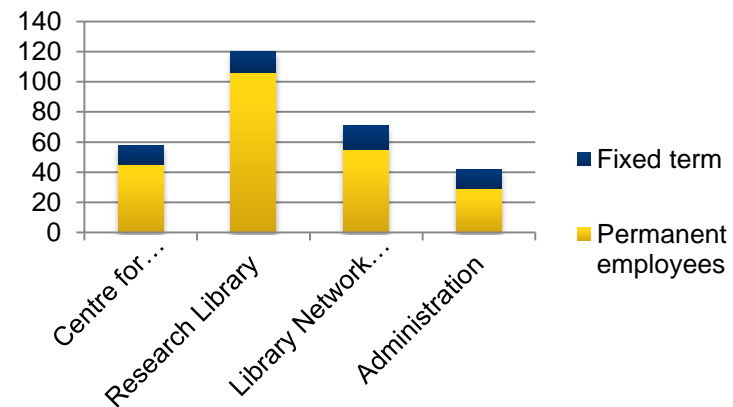
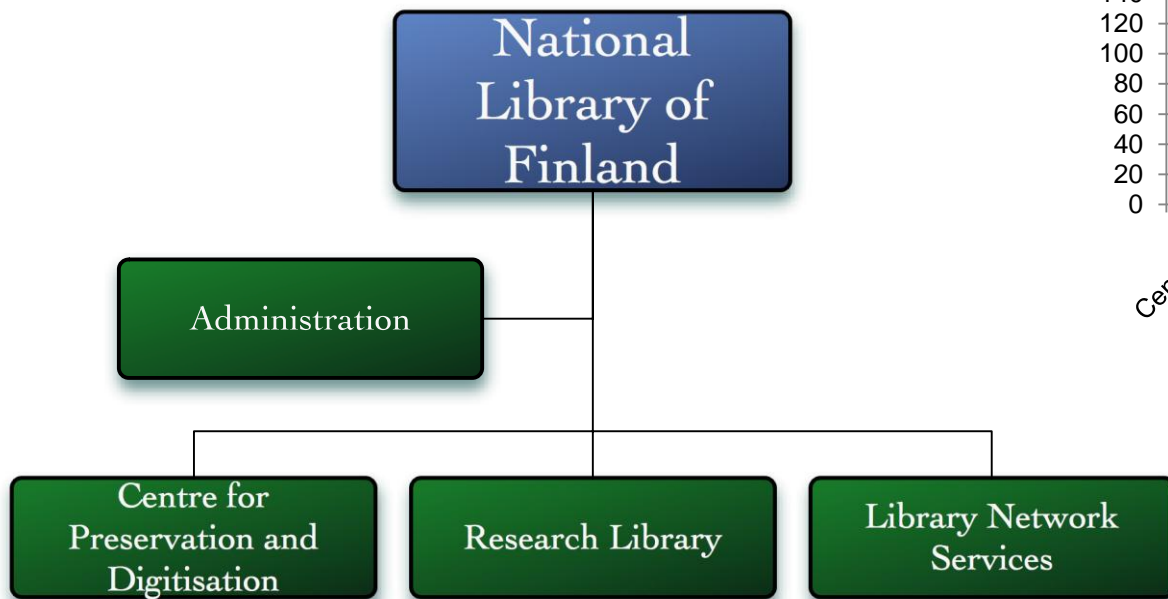
Karo Salminen

March 15th 2012

Sofia

Overview

- National Library of Finland as a organisation
- Digitisation at NLF
- Joint approach of libraries, museums and archives to access and preservation for digital material – the National Digital Library



Many roles of NLF

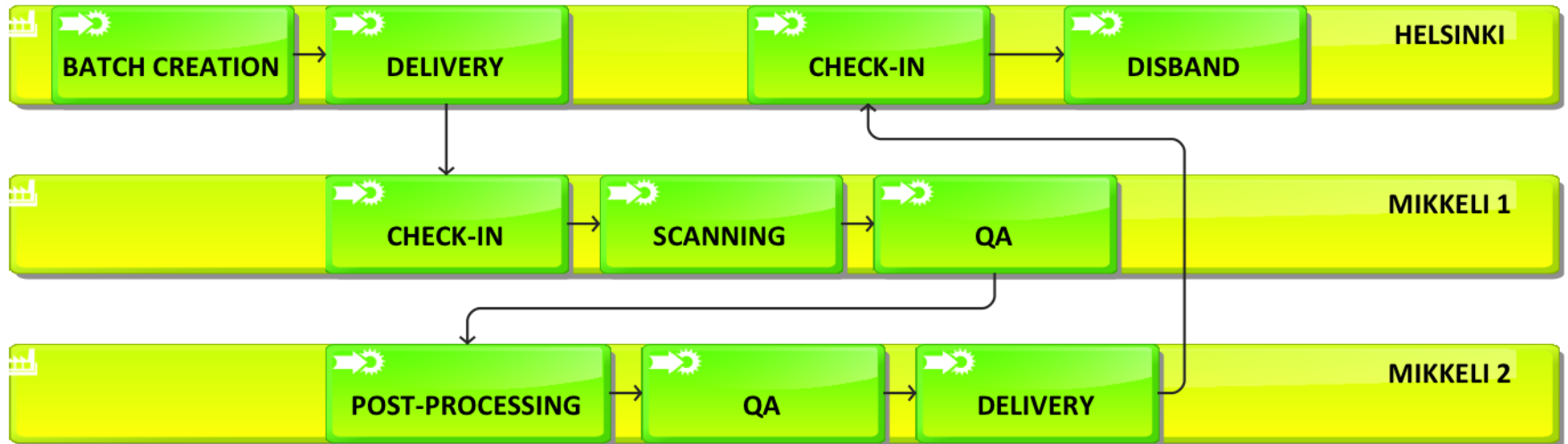
- A research library – part of University of Helsinki
- A national library:
 - Receiving legal deposits since 1707: all written publications deposited to NLF
 - Collecting web sites and electronic publications since new law came into force in 2007
- A rather new role: A national service and development centre for the Finnish library sector

Digitisation at NLF

Centre for Preservation and Digitisation

- Rather large organisation unit of NLF
 - Majority of the permanent employees carrying out tasks on microfilming of newspapers and magazines, and conservation
 - All the Finnish newspapers are microfilmed
 - Investigating new possibilities for replacing the traditional workflow:
 - Piloting digital depositing of newspapers with various publishers
- In-house digitisation
 - Better control over quality and handling of the physical originals
 - Structural markup is labour intensive

Digitisation workflow – from and back to Collections



Tools for mass digitisation

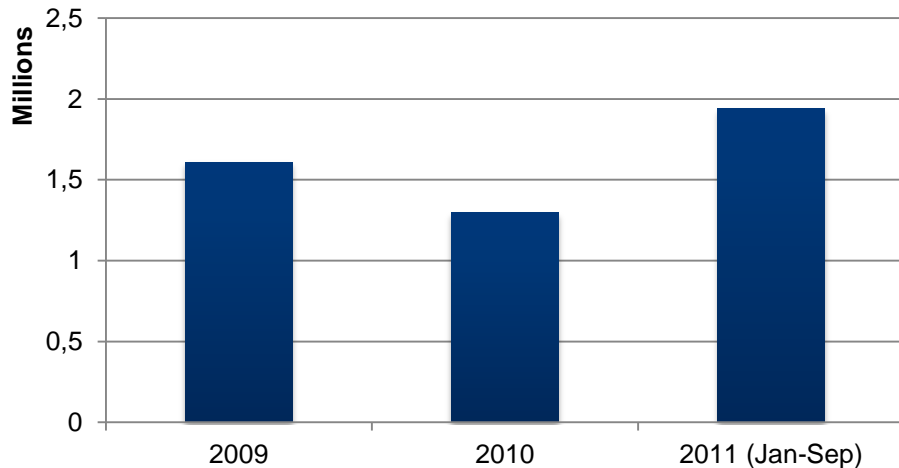
- An automatic scanner
- Software (docWorks by CCS) for postprocessing:
 - layout analysis
 - text recognition (OCR)
 - structure analysis
 - quality control
- Operators for controlling quality and marking up document structure





Mass digitisation in numbers

Digitised pages



- Mainly newspapers and magazines
- In 2011 majority of magazines published between 1810-1944 had been digitised
- Possible to increase the rate of annual digitisation to 4 million pages – only (long-term) funding is missing

Challenges in the digitisation workflow

- Distance between locations
 - Logistics and item tracking
- Cooperation over organizational boundaries

Digitisation and preservation

- Digitising for access and to protect the physical originals, not for preservation per se
- Original copy is the target of preservation
- However, why not treat the digital copy as a target of preservation?
 - In the end, only digital copy can, in principle, last forever
 - Practical reasons: avoid redigitisation
- Not forgetting born-digital material

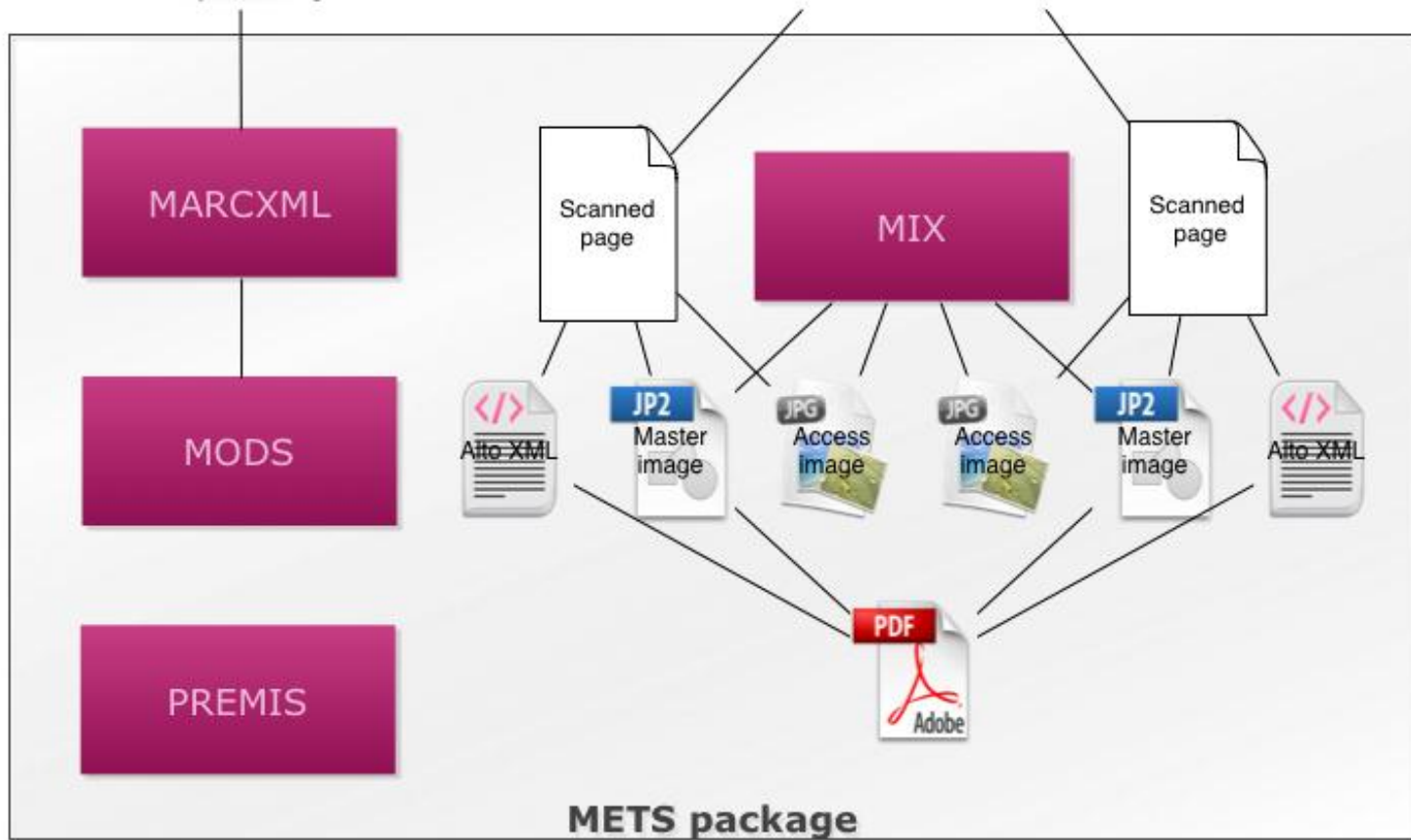
Criteria for digitisation - policies

- Critical masses – coherent collections
- Protection of the physical original
- Demand
- Content: national or international value

- Authentic and high-quality digital surrogates for research use and public reuse
- Enhancing visibility, accessibility and usability of unique collections

Digitisation planning

- Public sector is the primary source of funding; cooperation when possible to produce larger collections and to reduce overall costs
- Digitising volume is dependent on resources, possibility to cooperate and level of quality
- A three-year plan, focus on:
 - Newspapers & magazines (fragile & widely used source material, but difficult to provide access to original items)
 - Cassette tapes (brittle tape)
 - Ephemera (popular)



User access

- Currently, two independent solutions exist:
 - One especially for newspapers and magazines (bespoke software)
 - Another for monographs (DSpace with a browser module developed in-house)

HISTORICAL NEWSPAPER LIBRARY

Quick search

Search

[Main page](#) [Browse](#) [Search](#) [Advanced search](#) [Article index](#) [Feedback](#)
[Titles](#) > [Laatokka](#) > 15.06.1882 Laatokka no 24

[1 <<](#) [< 1](#) [2](#) [3](#) [4 >](#) [>> 4](#)

[Show 1:1](#) [Download issue](#) (PDF 5,6MB)

LAATOKKA

Sanomia Sortavalasta ja Itä-Karjalasta.

N:o 24.

Torstaina 15 p. Kesäkuuta.

1882.

Tilaukshinta

Sortavalasta: kelo kuukautta	3 m. 50 p.
toisenaikana	2 " 80 "
puoli	2 " 20 "
kuukausi	1 " 25 "
kolme kuukautta	4 m. 34 p.
kuusi kuukautta	7 m. 50 "
vuosi	12 m. 50 "
kolme kuukautta	3 m. 55 "
kuusi kuukautta	6 m. 45 "
vuosi	12 m. 51 "

Wenäjällä: kelo kuukautta 2 ruplaa, puoli kuukautta 1 rupla.

Wastuun-alainen toimittaja K. Suomalainen.

Zakoni ja -pailla:

Seitti joutuu joka arki toistaa kello 1 pöytäkirja Helsingin kirjastoissa, jossa kaupunkilainen tilaukset otetaan vastaan ja josko myös syytysjärjestämistä on määrätty 10 pöytäkirja.

Ilmoituksia otetaan vastaan kirjapainossa. Hinta on 10 pöytäkirja pieneltä riviltä tai sen sisältämältä sijalta toistokissa ja 15 pöytäkirja etuisuilla. Kuotitiedot suomennetaan eri maissa. Ilmoitukset ovat annettuat kirjapainon viimeistään toistokissa kello 4.

Hänen Keisarillisen Majestetinsa

Armollinen

Huhe

Suomenmaan Säädhyille

marfinalia waltioapäiviä lopetettaisa

Helsingissä 6 päivä heinäkuuta 1882.

Suomen Suuriruhtinaan-
maan Edusmiehet!

Koska se aika, joka oli määrätty teidän kooaja-olollenne, nyt on loppuun kulunut, Minä teitä kiitän sekä teidän uutterasta työstänne maan hyväksi, että myöskin

tukseen määrättyjä lahjoitusvaroja, waloitettaihin opettamaan myöskin weraisa kielisä;

i) että ne warat, jotka esittelijäsihteri wainaja Henrik Reinhold Beck testamentilla on määrännyt entisen Turun hitappan koululaitokselle, käytettaihin aiottuun tarkoitukseen.

Mitä miuhin pyyntöesityksissä tawattawiin anomuksiin tulee, ehdottaa waltiofuntta:

F) että ne eiwät antaisi aihetta miuhin toimeen waltiofunttien puolelta.

Wastalauseita on kolme.

Great J. Pippingföhd y. m. m. esittämät säätäjien hyväksyttäväksi: 1:o että waltiofuntta...

wikfi ten waltiofunttien järjestelmän mukaan, jota waltiofunttien hyväksyä; 2:o että waltiofunttien sen lisäksi anowat, että perustukseksi aiottuille muuttokissa koululaitoksissa ajettaihin keisarillisen senaatin 1879 wouuden koulukomitean mientinnön johdolla tekemä suunnelma, säilyttämällä sen pääpiirteet:

a) Ominwaraiset pohjakoulut mahdollisuuden mukaan päättywillä oppimäärillä, joiden tarkoituksena on sekä valmistaa korkeampiin jatkokouluihin ja siten formata nuorisia n. k. oppikoulujen alempia luokkia, että myöskin astua nuhisten alhaisempien realkoulujen sijaan, joissa sen tahden ei lueta latinan kieltä ja jotka kielellisissä suhteissa etupäässä tarwittawat

maamme koululaitoksen perin pohjaisista uudesta järjestämistä, kun S. M. Keisari waltioapäivän awauksessa suoraan ilmoitti, että Hän ei ole woinut hyväksyä juuri samaa uudesta järjestämistä, niin katomme walmollisuudellemme maan waltiofunttien kunnioitussella ehdottaa,

että waltiofunttien hylkäsiwät yleisen waltiofunttien loppulauseen n:o 2 sekä sen johdosta myöskin loppulauseen n:o 3 ja ettei anomusehdotus tässä osessa antaisi aihetta miuhin toimeen. Niinkuin arwatakin sopia, on waltiofunttien hyväksyttävä pappis- ja talonpoikaissäädhyissä, mutta sitäwastoin hyväksyttävä aatelis- ja pormarisäädhyissä sillä poikkeuksella, että ettei mitään walmollisuutta

Nimeke The history of Lapland :

Tekijä Schefferus, Johannes
Johannis,

Vuosi 1674.

Sisällysluettelo

[THE PREFACE](#)[THE HISTORY OF LAPLAND](#)[CHAP. I.](#)[CHAP. II.](#)[CHAP. III.](#)[CHAP. IV.](#)[CHAP. V.](#)[CHAP. VI.](#)[CHAP. VII.](#)[CHAP. VIII.](#)[CHAP. IX.](#)[CHAP. X.](#)[CHAP. XI.](#)[CHAP. XII.](#)[CHAP. XIII.](#)[CHAP. XIV.](#)[CHAP. XV.](#)[CHAP. XVI.](#)[CHAP. XVII.](#)[CHAP. XVIII.](#)[CHAP. XIX.](#)[CHAP. XX.](#)[CHAP. XXI.](#)[CHAP. XXII.](#)[CHAP. XXIII.](#)[CHAP. XXIV.](#)[CHAP. XXV.](#)[CHAP. XXVI.](#)[CHAP. XXVII.](#)[CHAP. XXVIII.](#)[CHAP. XXIX.](#)[CHAP. XXX.](#)[CHAP. XXXI.](#)[CHAP. XXXII.](#)[CHAP. XXXIII.](#)*Of the temperature of the Air, and soil of Lapland.*

7

I proceed next to the disposition and nature of the Country, having first given you a Map of it.

CHAP. III.

Of the temperature of the Air, and soil of Lapland.

WE have seen how *Lapland* is situate; let us next proceed to other particularities of it. That 'tis very near the Pole appears from its latitude, inso much that for some months in the Summer the Sun here never sets, and on the contrary in the Winter it never rises, which *Herbersten* says is but forty days, and the three hours in the night the body of it is something darkned, so that his rays appear not, yet there is so much light, that they continue their work all the while. Indeed the same account is not to be taken of the whole Country, since part of it lies nearer, and part further distant from the Pole; and of these too some parts are more East, and some more to the West, from whence 'tis that with some of them the Sun is scarce above the Horizon for so many daies as he pretends. And altho in the Summer it never sets and goes below the Earth, yet neither does it rise much above it, but as it were kisses and gently glides along the Horizon for the most part; as likewise in the Winter when lowest it is not much beneath it: which is the reason that tho they have one continued night for some months; yet every day the Sun comes so near, that it makes a kind of twilight. *Job. Magnus* saies that in the absence of the Sun there are two twilights, one in the morning, the other in the evening, in which those poor remainders of day provide that the night should not be utterly destructive. And by how much the Sun is farther absent, the light of the Moon is clearer. Hence *Petr. Claud.* saies that when the Moon shines they go a fishing, and dispatch all other necessaries that are to be done without doors; and when it does not, if the air be clear, even the light of the Stars so much abates the darknes, that the horror of the night is much lessened, and there is light enough for the dispatch of severall busineses; which is farther assisted by the whiteness of the Snow. The Air of *Lapland* is cold, but fresh and clear, and consequently very wholesome, being much purified by the winds which are here very frequent and violent. It has bin attested to me by eye-witnesses, that there rises a certain wind out of the Sea, which beginning to blow raises presently such thick and dark clouds even in the midst of Summer, that they utterly hinder the sight, and in the Winter drives the snow with such force and quantity, that if any were


Preservation of digital objects

- Preservation copies stored securely and independently of the access copies
- Relying on the national effort for a more complete solution

Crowdsourcing some of the manual work

- Structural analysis and other post-processing tasks require manual work
- OCR'd text is not proofread by anyone in the postprocessing process, however
- Launched a pilot in early 2011, in attempt to crowdsource proofreading, with a game developed by Microtask


DISH Award 2011 – Best crowdsourcing project: Digitalkoot



SCORE: 12345 **SAVED: 4/10**

LOG IN TO PLAY!

In Mole Bridge you write the words where the computer failed. Watch the tutorial!



LEVEL: 2 **SCORE: 12500**

LOG IN TO PLAY!

In Mole Hunt you try to identify words the computer has misread. Watch the tutorial!

Crowdsourcing: next steps

- Engage researchers and individuals to fix the OCR'd text while reading the actual articles
- Sequel to Digitalkoot project: Kuvataalkoot, funded by European Social Fund 2011-2013
 - Marking of images and/or articles
 - Indexing them

Collaboration on national level: National Digital Library

Libraries, archives and museums together

- Funded by Ministry of Education for years 2008-2011 and 2011-2013; the most extensive cooperation project between libraries, museums and archives so far in Finland
- Goals:
 - Single point of access to all digital material, including catalogues – a joint public interface (in operation 2011/2012)
 - Digitisation of key materials in libraries, archives and museums
 - Digital preservation solution for digital cultural material
 - Central aggregator for the European Digital Library Europeana

NDL Governance 2011-2013

Executive group / appointed and chaired by the Ministry of Education and Culture

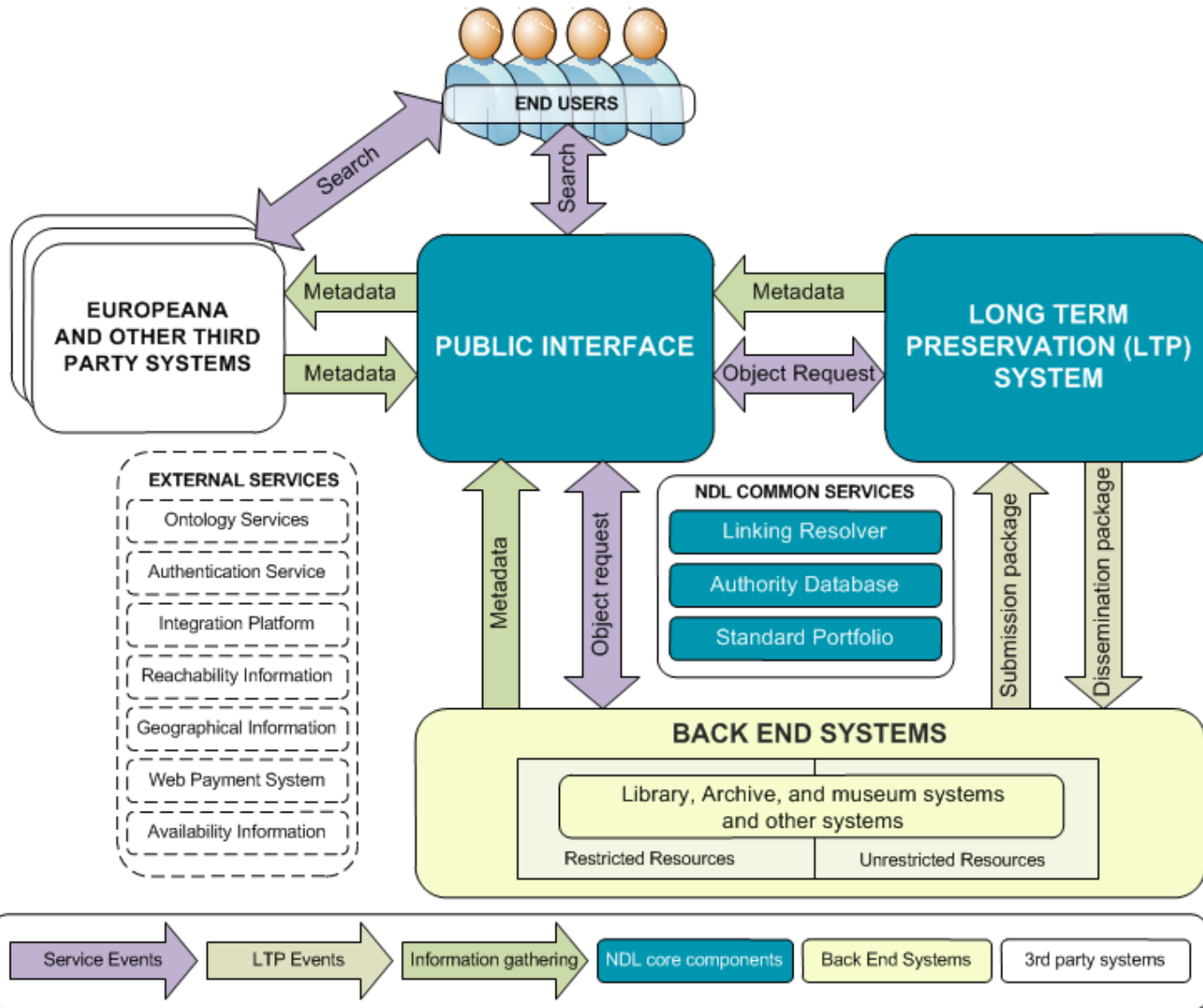
Steering group / appointed and chaired by the Ministry of Education and Culture

User Interface:
maintenance and
development: the
National Library

Consortium of
libraries, archives
and museums /
chaired by the NL

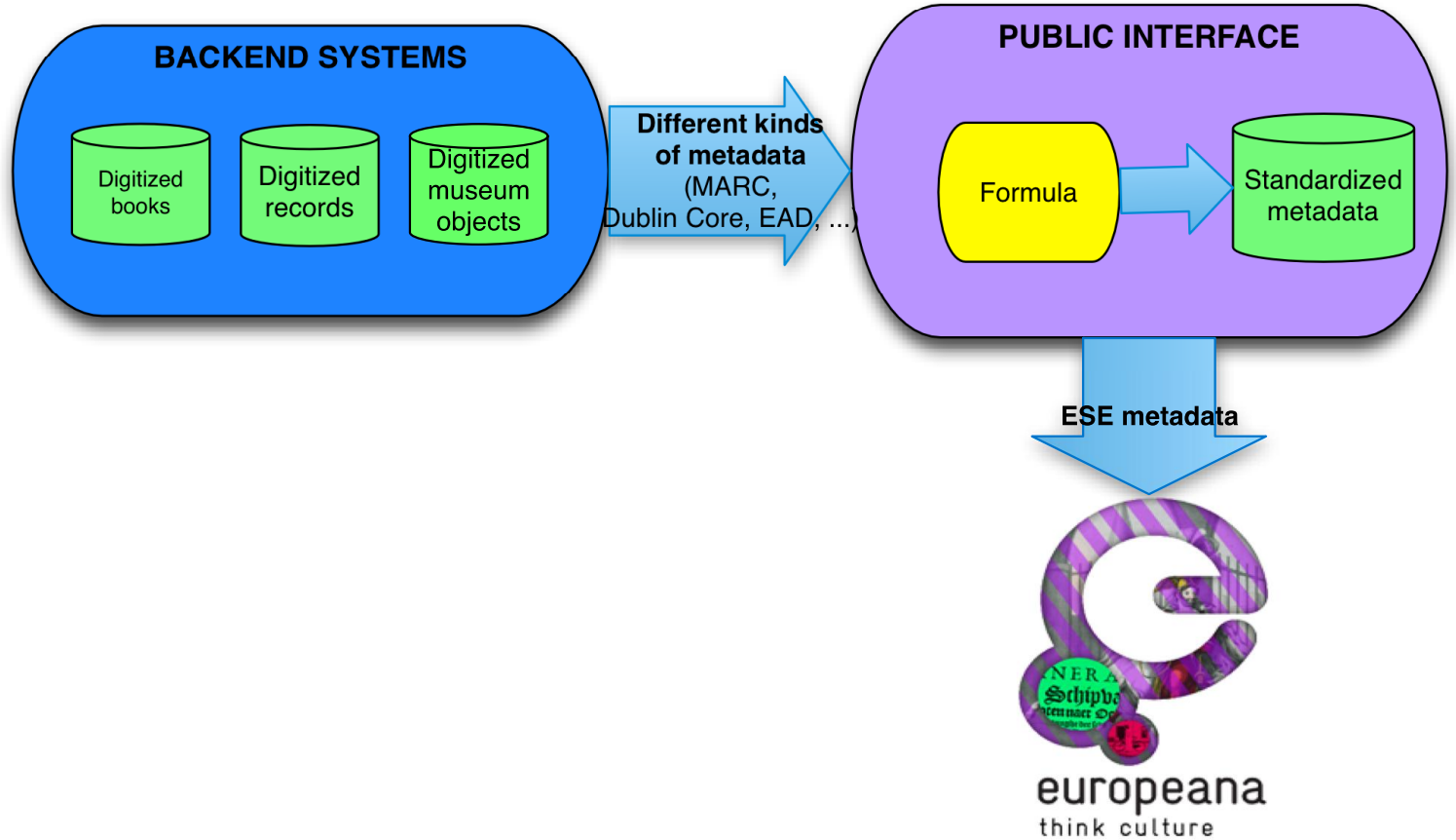
Support Group
of libraries,
archives and
museums /
chaired by CSC

Long-term
preservation Project
/ CSC (IT Center for
Science)



Single point of access

- Bringing the wealth of information sources to one, single public interface
- Services of libraries, museums and archives integrated into one system
 - Renew loans, buy pictures and order materials
 - Central indexing of the databases and information sources provides quick and versatile searches
 - Emphasis on usability
- Suffered from major delays, but being currently implemented



What has been digitised?

- During 2008-2011 16 million objects were digitised, for 2011-2013 goal is to digitise 7 million objects
- Photos and maps, newspapers, parish registers, diaries, museum objects, herbariums
- Within NDL several seminars on digitising audio, photographs, and digital preservation have been held

Preserving the digital heritage

- A national, shared solution for preserving digital heritage
- What's been done:
 - Prerequisites for member organisations
 - Process modelling
 - Specifications of metadata
 - File types based on media type
 - Identifiers
 - In OAIS terminology: SIP and AIP specifications

Lessons in cooperation

- At first, conceptual boundaries between institutions hindered efficient cooperation
- Discovered needs to upgrade and existing systems and standardise information (especially in archive and museum sectors)
- In the process, growing interoperability between systems and improving information coherence

Further information

- [Digitisation Policy of NLF](#)
- [Preservation Policy of NLF](#)
- NLF's digitised material online (not subject to copyright):
 - <http://www.doria.fi/handle/10024/69173>
 - <http://digi.nationallibrary.fi>
- Crowdsourcing:
 - <http://www.digitalkoot.fi>
- National Digital Library: www.kdk.fi/en



16 40

THE NATIONAL
LIBRARY
OF FINLAND