

# The road to the future of Finnish library knowledge

## Finnish linked library data project

Jarmo Saarikko & Matias Frosterus

National Library of Korea 2023-10-06

# About the authors

## ▪ Jarmo Saarikko

- is an information specialist at the National Library of Finland. He has a long experience in special libraries and metadata management. He is currently working on a model of linked library data based on Bibframe and RDA. He has a Ph.Lic. degree in Zoology from the University of Helsinki.

## ▪ Matias Frosterus

- is an information systems manager at the National Library of Finland. His main responsibilities include advancing of the metadata services of the National Library. He has a D.Sc. (Tech.) degree in Media Technology from Aalto University.

# The National Library of Finland

in brief





# The National Library of Finland



# The National Library of Finland

Ministry of Education and Culture

University of Helsinki

National Librarian  
Kimmo Tuominen

Research Library

Library Network Services



approx. 230  
employees

## Preserving Cultural Materials

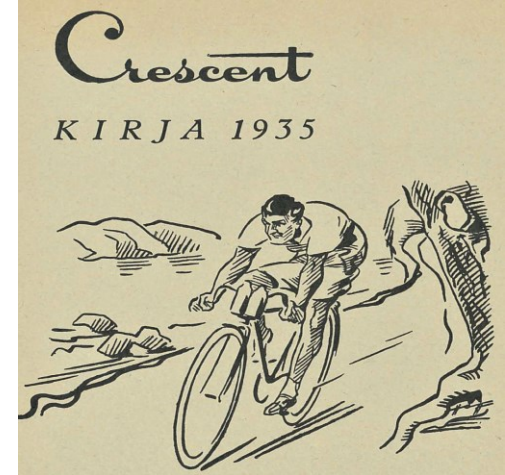
- Legal deposit services: physical and digital legal deposit copies
- National bibliography and discography – the Finnish national imprint
- Identifiers, standardization, cataloguing standards and cooperation services

## Research library for the academic community

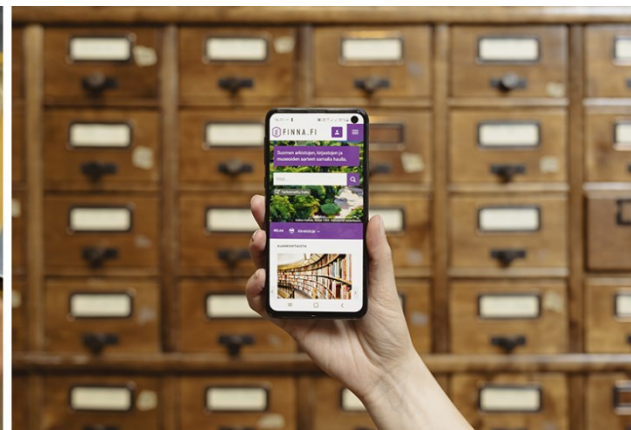
- History, literary science, philosophy, music and Russian studies
- Digitization. Also as a service. Facilities in the city of Mikkeli.

## Library Network Services for organisations

- Finna discovery portal services for libraries, also for archives and museums
- Services for describing and managing library and other materials and collections (Melinda, Koha, Finna, Finto, Annif and institutional repositories)
- Providing e-resources and open science services for consortium members (FinELib)



# Strategy 2021 - 2030



## *Cornerstones of operations*

- Openness
- Renewal
- Bildung

## *Bildung at the heart of society*

- Cultural heritage for the common good
- National Library at the heart of academic community
- National Library as a resource of Bildung and learning
- Creating a powerful hub of expertise through network cooperation

# The Finnish Strategy

- Strategic plan of the National Library of Finland 2021–2030
  - <https://urn.fi/URN:NBN:fi-fe2020042822711>
- We will develop the description of information resources in accordance with the principles of **open, linked data**.
- Digital service solutions will primarily be based on **open-source code, open interfaces** and broad Finnish and international **collaboration**.



# FAIR data principles

- Findable
  - Accessible
  - Interoperable
  - Re-usable
- The data produced by the National Library will be managed in accordance with the FAIR principles and we will instruct users in the responsible use of data - <https://www.go-fair.org/go-fair-initiative/>

# Linked data

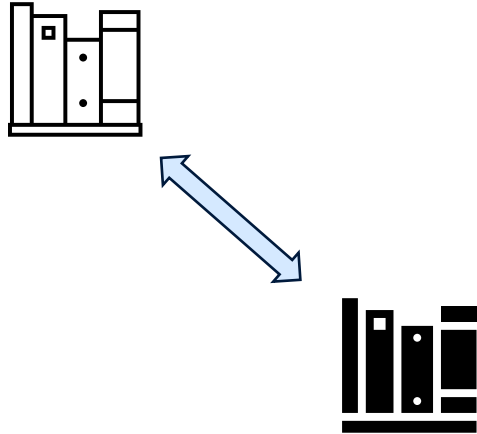
in library context

# Linked data

- Metadata with machine-processable **relations** to other metadata
- Uses
  - Identifiers that are dereferenceable
  - RDF (Resource Description Framework – W3C standard for metadata)
  - Ontologies and vocabularies
  - Descriptive metadata
  - Cataloguing according to the entity-relation style
- Linked library data is supported with RDA description rules

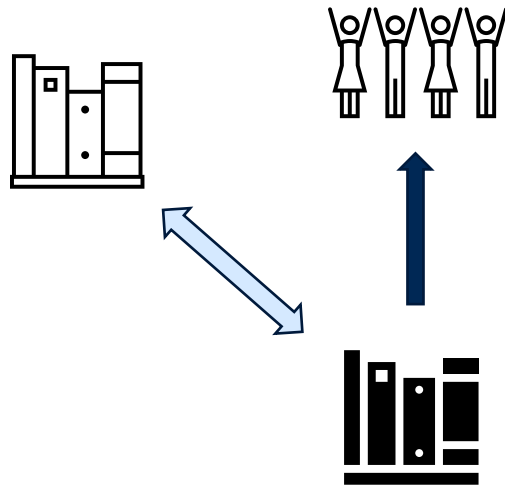


# Motivation for linked library data - interoperability



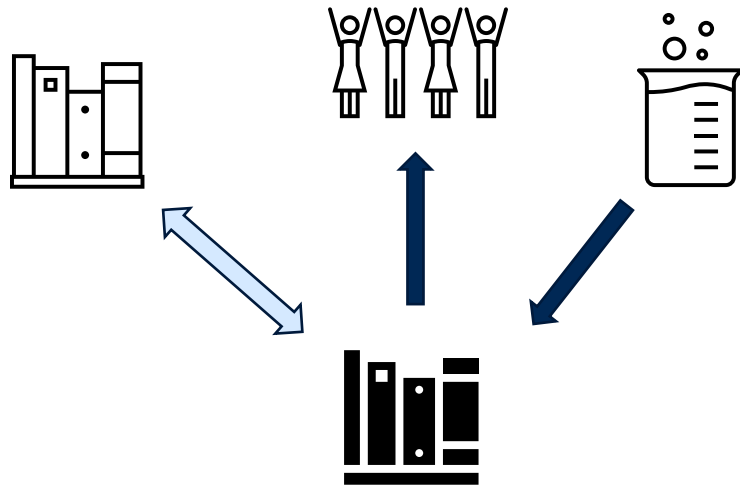
- **Interoperability** with several types of library data
  - A starting point

# Motivation for linked library data – wider use



- Non-libraries can make use of library data
  - Better re-usability for our data
  - Makes our data more valuable

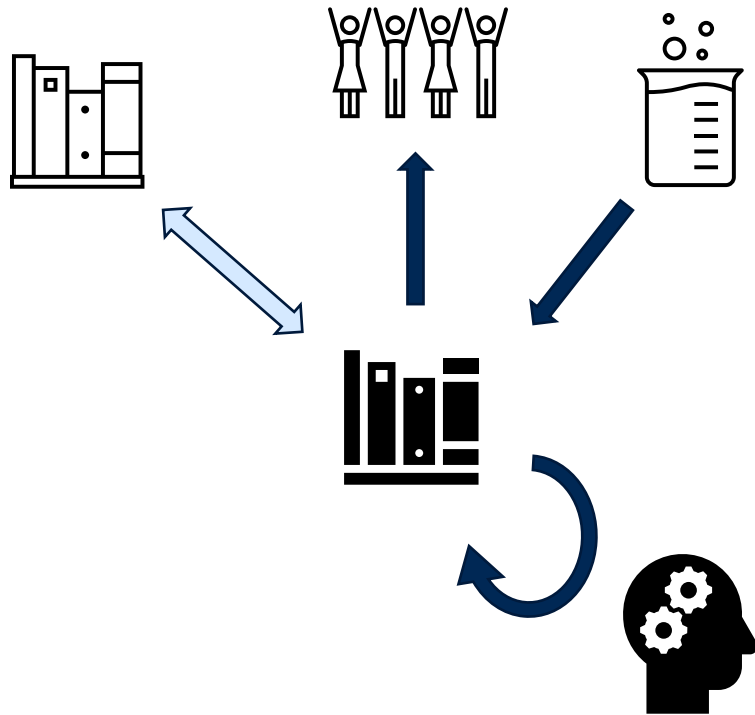
# Motivation for linked library data - enrichment



- Libraries can enrich their own data
  - For example by linking to research data
  - Public sector data
  - Wikidata

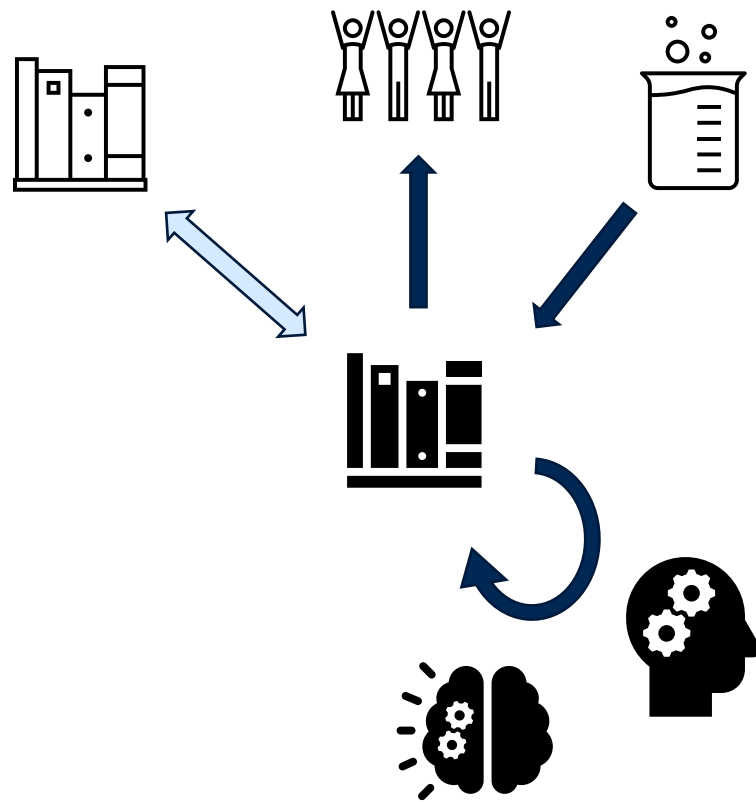


# Motivation for linked library data – re-usability



- Cataloguers can better utilize metadata that has already been produced
  - For example, when cataloguing a new edition, the old Work-level annotation can be re-used

# Motivation for linked library data - RDA



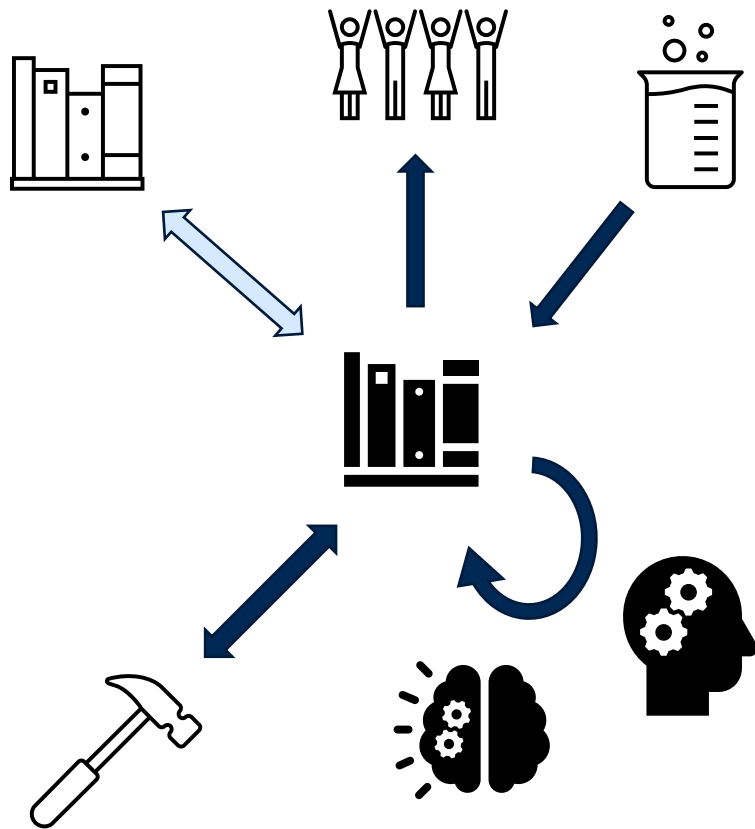
- Enables wider and smoother utilization of RDA properties
  - RDA cataloguing rules designed for description of linked data

## Research Description & Access

<https://access.rdatoolkit.org/>

- A paid service
- National license

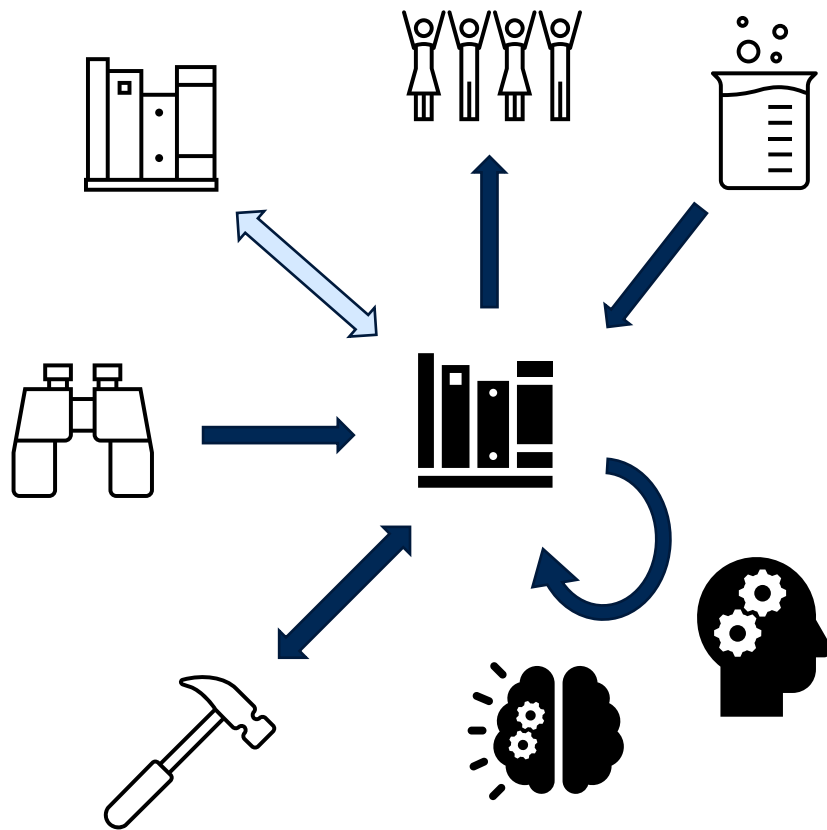
# Motivation for linked library data – new tools



- Makes it easier to employ new systems and tools for library data
  - Machine-readability of data
  - Automation and AI

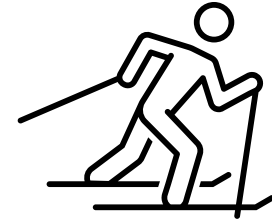


# Motivation for linked library data – new views



- Enables new ways and points of view for accessing the data for end-users
- New discovery services built on top of linked data

# Our linked data continuum



Union  
Catalogue  
Melinda

Adoption of  
Original RDA

Identifiers

Concept-  
based SKOS  
vocabularies

Official RDA

New  
Melinda

Finto  
Thesaurus and  
Ontology  
Service

BIBFRAME

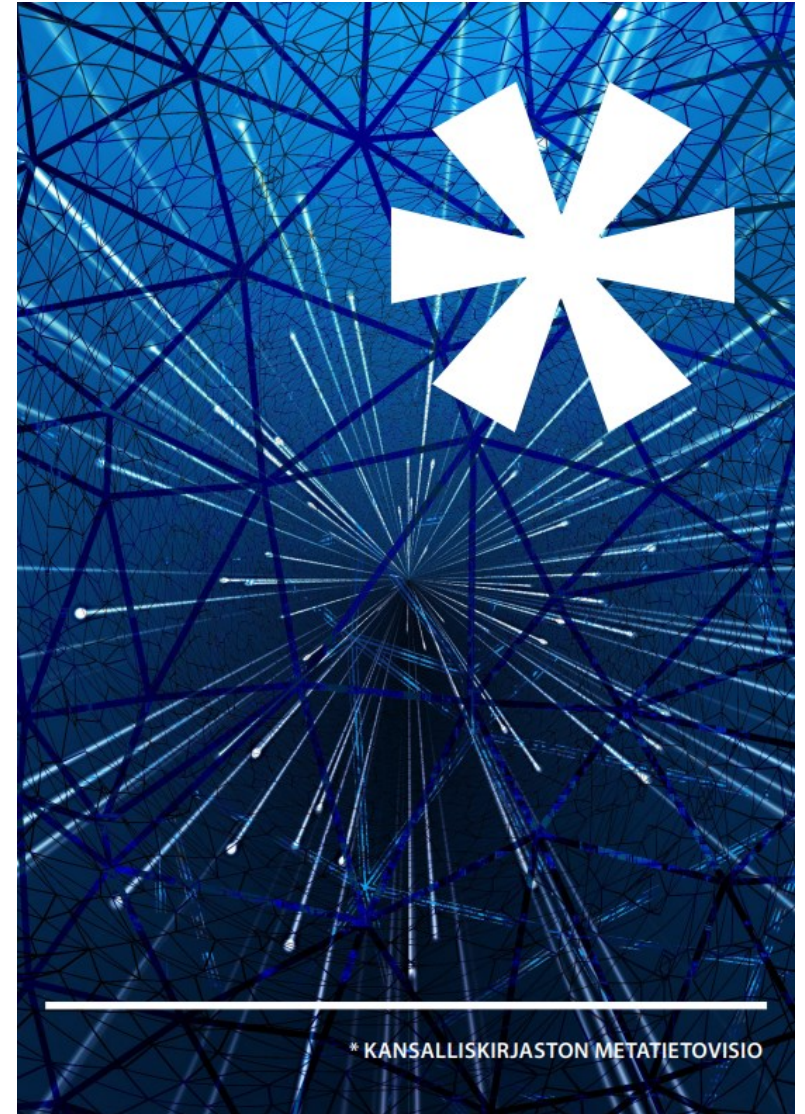
Works and  
expressions

# The Metadata Vision

of the National Library of Finland  
aimed for the GLAM sector

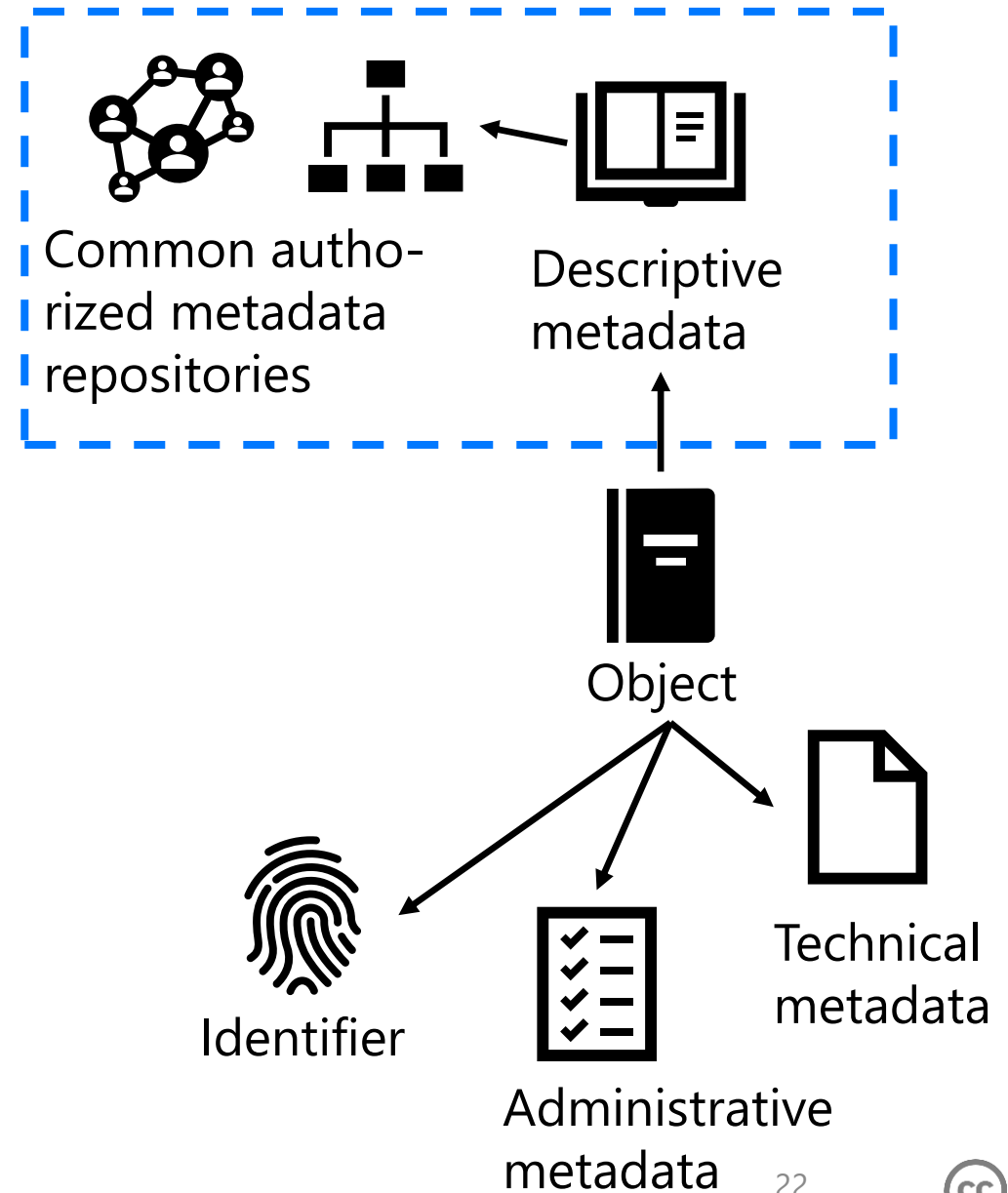
# Metadata vision

- White paper style document of how metadata is going to develop during this decade in the GLAM sector
  - Published by the National Library of Finland in 2020
  - <http://urn.fi/URN:ISBN:978-951-51-5905-2>
  - (Unfortunately only in Finnish)
  - A beginning for discussion
- **The National Library is committed to producing quality and rich metadata.**



# Metadata vision

- Customers and users
- Development of the metadata itself
- Information architecture
- Systems and tools
- The ecosystem of cataloguing (annotation)
- The role of the National Library of Finland



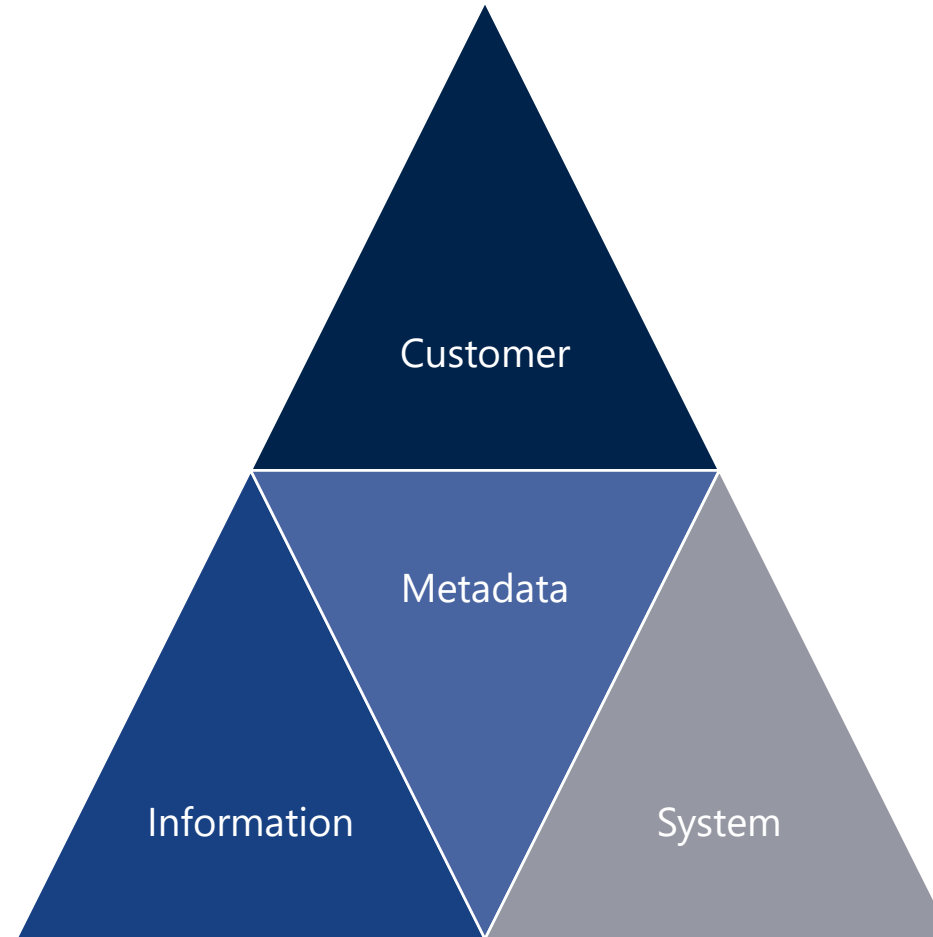
# Key messages of the Metadata vision

- The importance of identifiers
- Common authorized metadata repositories
- International interoperability
- Multilinguality
- Semantic interoperability





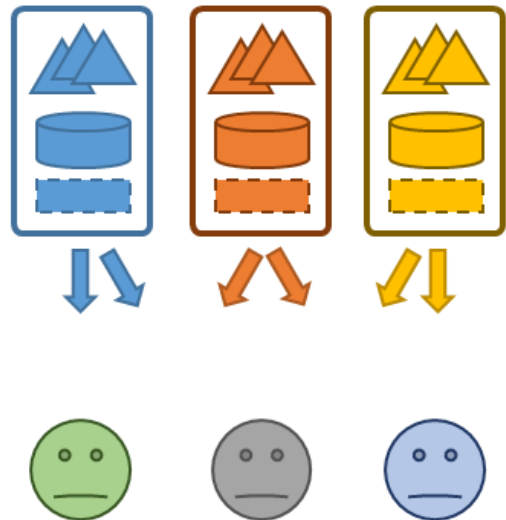
# Key messages of the Metadata vision



Metadata as a channel

# Key messages of the Metadata vision

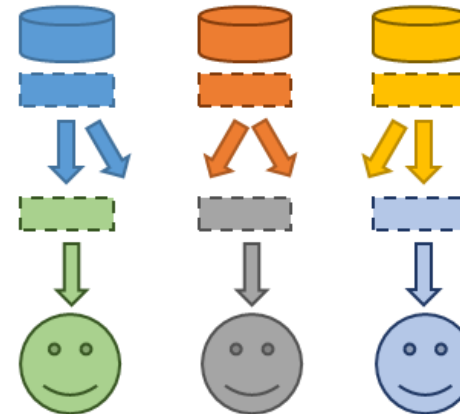
Overlapping information  
Siloed systems  
Non-modular



Common Authorized  
metadata repositories



Interoperable  
metadata



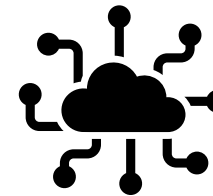
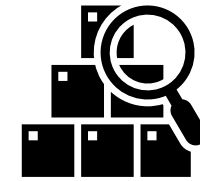
Services tailored for  
specific user groups

# Key messages of the Metadata vision

- Metadata as the basis for **AI**
  - Open metadata makes for more transparent AI systems
  - The causes and effects in the functionality of AI more understandable
- **Automation** is going to get more prominent
  - Frees up human resources to be able to contextualize and link the information in order to serve the more complex information needs of the users
  - Human control in semi-automatic systems
  - The ability to hide complexity but also to show it when necessary
  - Linking should be easy
- **Provenance** data

# Key messages of the Metadata vision

- Development possibilities of the **discovery interfaces** for the end-users
  - Search expansion
  - Explainability
    - The need to be able to answer the question "**Why** this set of results?"
    - However, it can be very difficult!
- Powerful interfaces
- Modularity of systems



# The reception of the Metadata vision

- The goal was to form the beginning of a conversation in Finland
- The Metadata Vision is a useful for reference
  - Made for our own use, but
  - many others have been discussing and referring to it
  - May help in funding applications
- A complementary document to the National Library Strategy

# The reception by public libraries

- Public libraries launched a project called Towards the Metadata Vision (2022-2023)
  - A deep dive into what the linked data future holds for the public libraries
  - Interviews with library personnel in Finland as well as a trip to Sweden
- The final report of the project was published in April 2023
  - An overview of the landscape and the incoming changes, the attitudes towards it, and a rough timeline prediction
  - [https://www.kirjastot.fi/sites/default/files/content/ykn\\_kohti\\_metatietovisiota\\_1\\_oppuraportti2023.pdf](https://www.kirjastot.fi/sites/default/files/content/ykn_kohti_metatietovisiota_1_oppuraportti2023.pdf) (only in Finnish)
  - For building a clear roadmap and preparing the architectural changes a continuation project is being planned



# Linked data in practice

at the National Library of Finland

# Shared authorized metadata repositories for

- Subjects
  - Classifications
  - Genre/Form
  - Agents
  - Places
  - Time
  - Metadata value and code lists
- All vocabularies are available at the National Thesaurus and Ontology Service Finto –  
[finto.fi/en/](https://finto.fi/en/)

from all ▾

English ▾

✕

Search

# finto

## Welcome to Finto.fi!

Finto.fi is a centralized service for interoperable thesauri, ontologies and classification schemes for different subject areas. You can use Finto.fi to browse vocabularies or integrate the vocabularies into your own system using the open APIs.

Finto AI provides a user interface and API for automated subject indexing.



### Available vocabularies and ontologies

#### GENERAL

- Allärs - General thesaurus in Swedish
- FGF - Finnish genre and form vocabulary
- HCLCS - Helsinki City Library Classification System
- KANTO - Kansalliset toimijatiedot
- KOKO Ontology
- Metatietosanasto
- PLC - Finnish Public Libraries Classification System
- Pondus categories
- UCUM - The Unified Code for Units of Measure
- UDC Summary
- YSA - General Finnish thesaurus
- YSO - General Finnish ontology
- YSOn käsite-ehdotukset
- YSO-time

### Tweets from @Fintopalvelu



**Nothing to see here -**

# The shared authorized metadata repositories

## ▪ **Subjects**

- Going beyond the term level into the **concept level** helps multilinguality immensely

## ▪ **Agents**

- Plans for a shared agent metadata service for the whole GLAM sector
- Data model based on RDA is far along but no funding for the pilot project yet

## ▪ **Places**

- Challenges: Modeling the temporal changes of places is difficult!

## ▪ **Time**

- Authorized labels for periods – may have cultural differences which affect mappings

## ▪ **Metadata vocabulary**

- Provides identifiers to various phrases, values, etc. which are needed in cataloguing - mapped with RDA registry - [www.rdaregistry.info](http://www.rdaregistry.info)

# The Linked Library Data Project

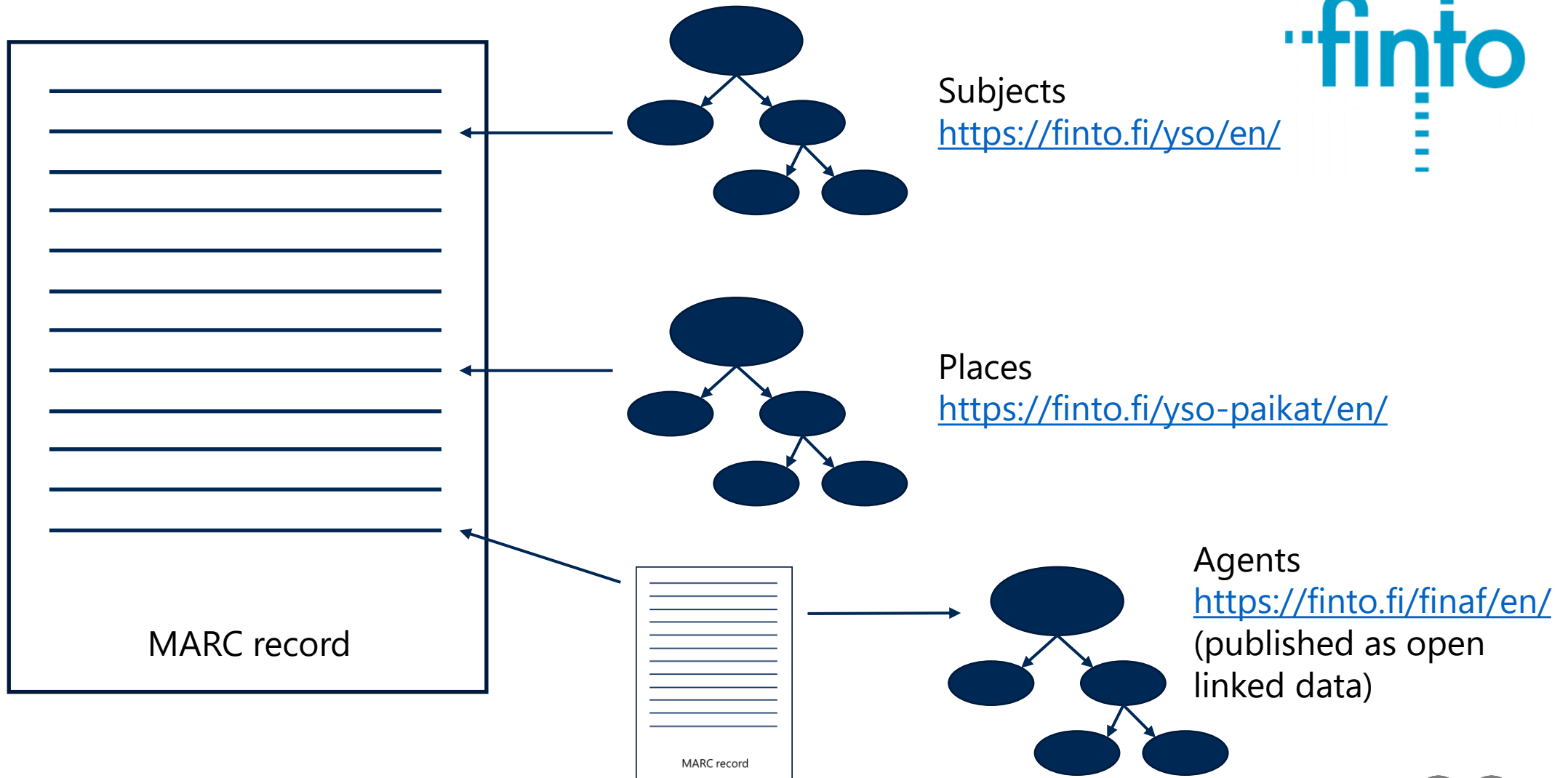
- Prepares the adoption of linked data for bibliographic cataloguing in the National Library and in part for the whole Finnish library sector
- 2022-2024
- A companion project of the platform change for the Finnish union catalogue Melinda running simultaneously

# Main goals of the Linked Library Data Project

- BIBFRAME based **data model** for Melinda
  - Including support for Works and Expressions (WEMI model)
  - Deep integration with RDA properties
- Further development of the “**Metadata vocabulary**”
- **Training** for catalogers in linked data for the library sector
- Preparation of the **data conversion**
  - Conversion rules from MARC21 to linked data and back
  - A plan on what will be done in the pre-processing steps and what in the conversion itself
  - Mining the works and expressions from the MARC records
  - Support for MARC21 will continue for many years to come

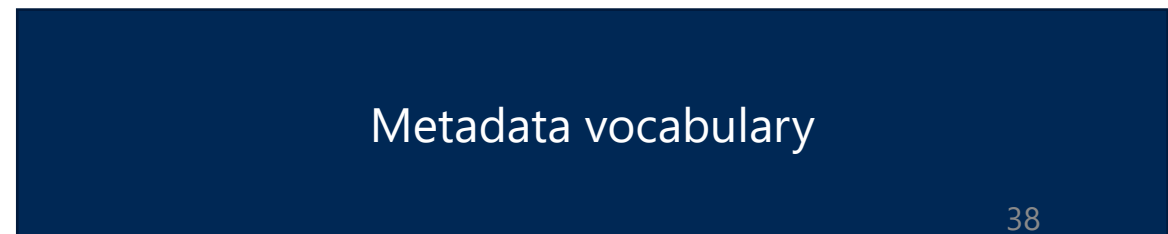


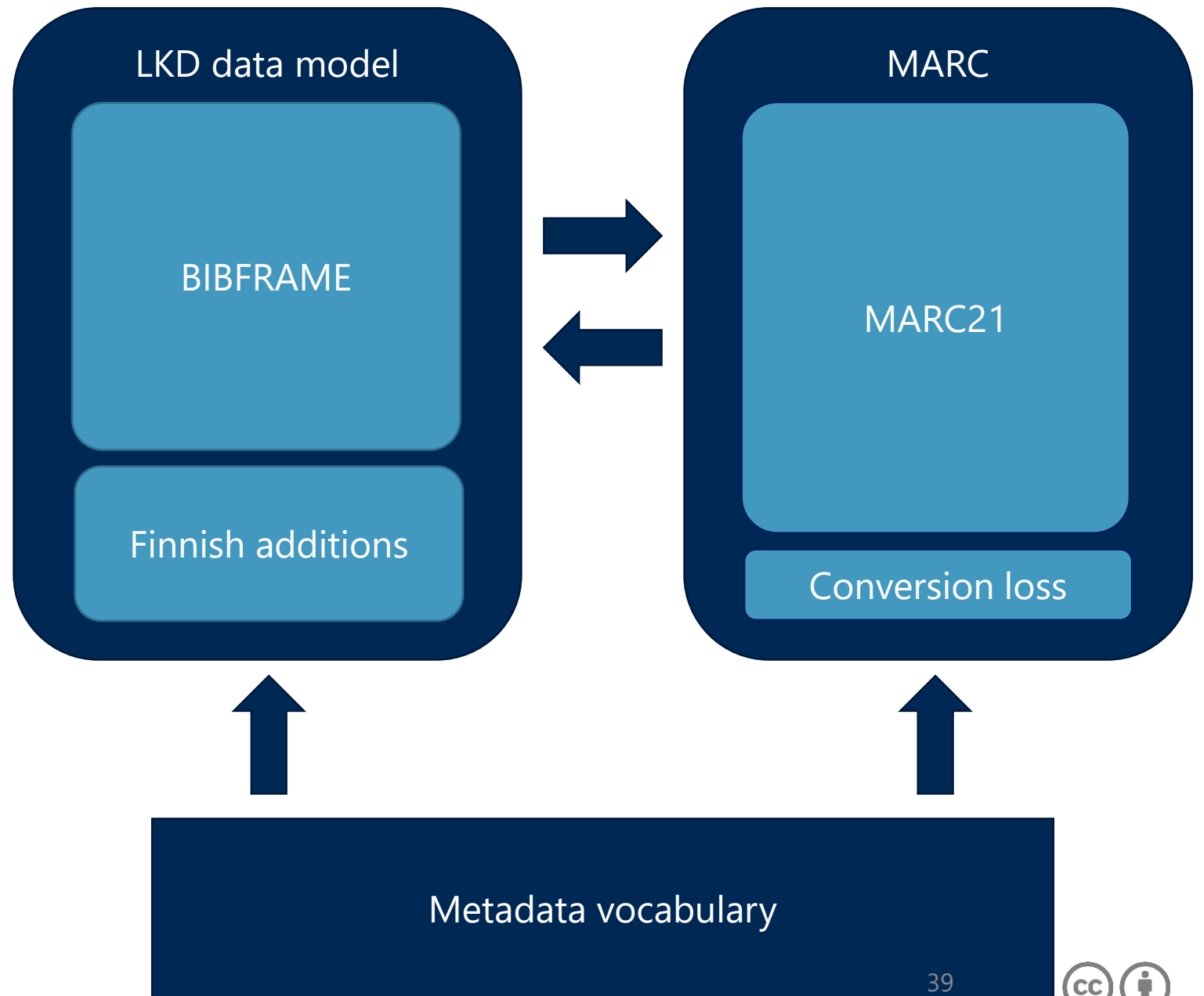
# Enriched bibliographic records in MARC

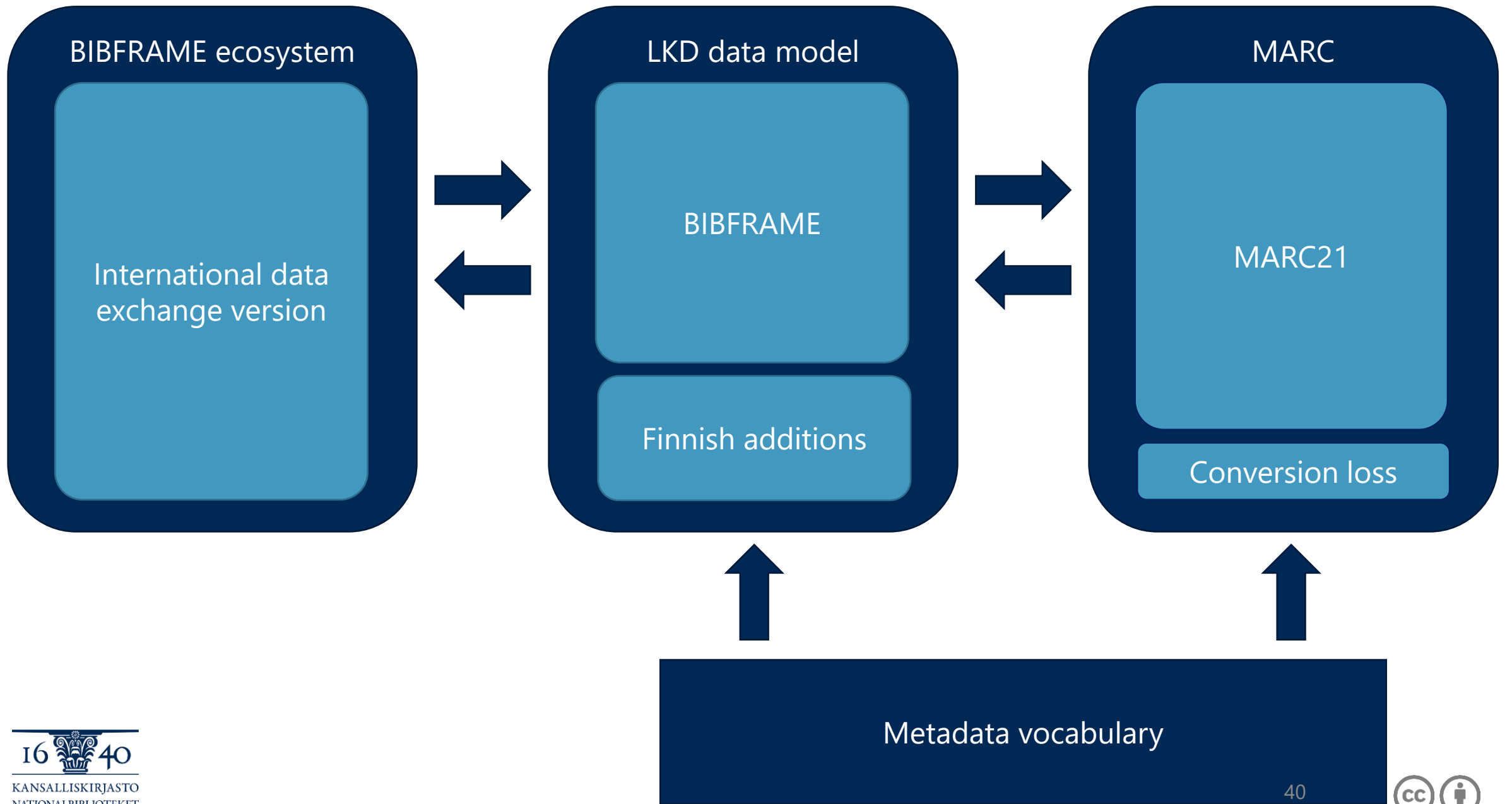


# The Metadata Vision

of the National Library of Finland  
aimed for the GLAM sector







# Resources of the LKD project

- Project manager (1)
- Systems specialist (1)
  - coding in Python
- Information specialists (2.5)
  - MARC, cataloging and modeling specialists
- RDA specialist (0.3)
  - 30% workload for this project

Six people working on the project

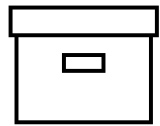


# International collaboration of the LKD project

- BIBFRAME Interoperability Group
- RDA collaboration
- Participation in ShareVDE
- BIBFRAME Workshop in Europe Organizing Committee

# Our approach

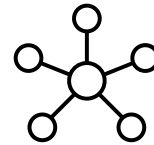
Sample MARC data



Conversion to BIBFRAME



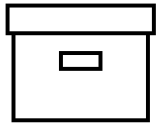
Analysis of the results



Conversion rules and pre-processing steps based on the nature of our data

# Our approach

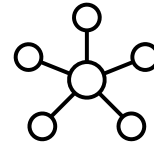
Sample MARC data



Conversion to BIBFRAME



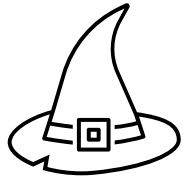
Analysis of the results



Conversion rules and pre-processing steps based on the nature of our data



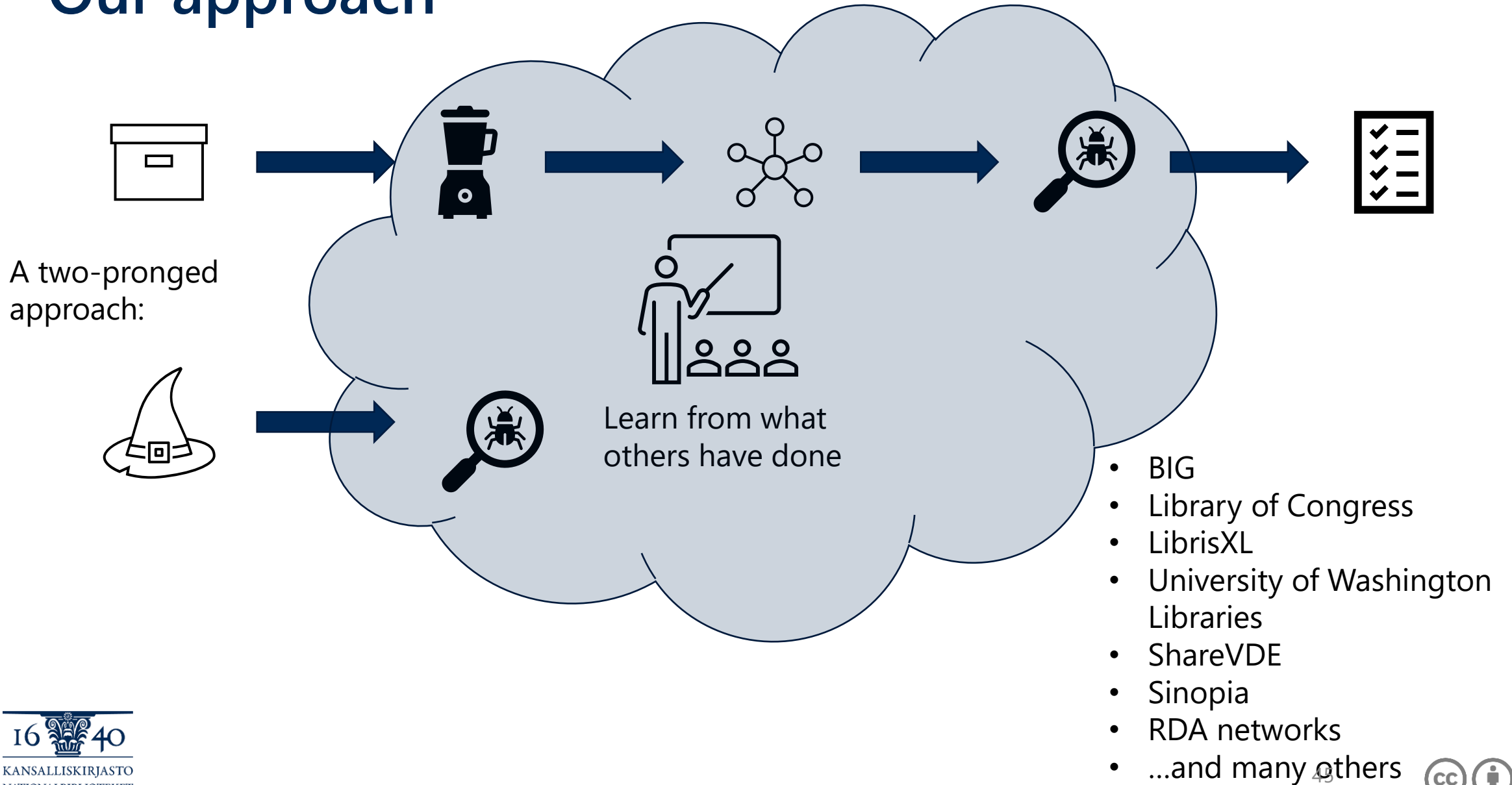
A two-pronged approach:



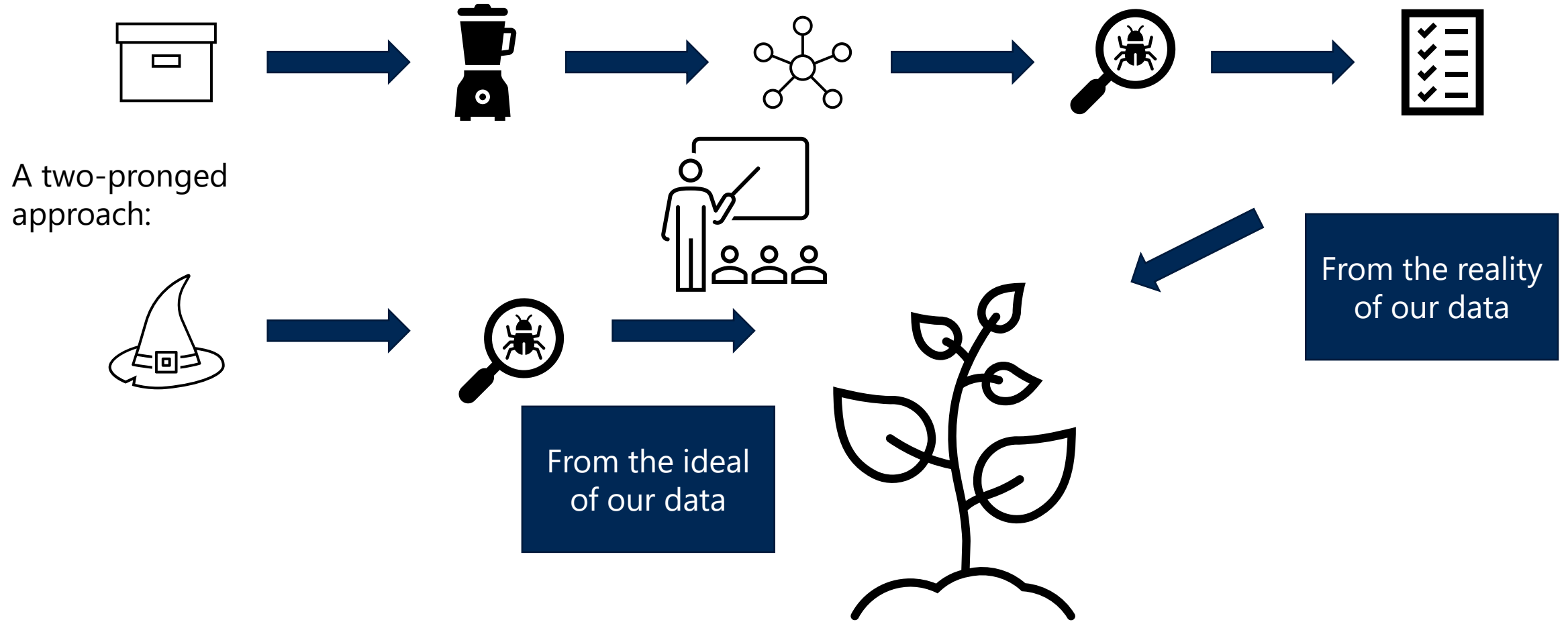
RDA as another starting point

Mapping RDA to BIBFRAME

# Our approach



# Our approach



# Where we are at v0.1

- We published the first raw version of the datamodel in January
  - 0.1.0 version contained the nationally **mandatory elements** for the narrow and broad description levels according to the **Official RDA**
  - Temporary namespace and identifiers
- Differences from Bibframe
  - We divided **bf:Work** into **lkd:Work** and **lkd:Expression**
  - We want to utilize our Metadata vocabulary for the controlled values of properties as opposed to free literal strings used in many Bibframe properties
    - using **owl:ObjectProperty** instead of **owl:DatatypeProperty**
    - values selected from [finto.fi/mts/en/](http://finto.fi/mts/en/)

# Where we are at v0.2

- Version 0.2.0 was published in June 2023
  - <https://schema.finto.fi/lkd/>
  - A huge number of fixes
  - Better translations of the terms to Finnish
  - Tackled **:title** – created a new solution with `lkd:TitleType`
  - Linking between datamodels strangely difficult
    - Still not happy with the semantics - see v.0.3

# What's next at v0.3

- Handling subclasses of bf:Work to our model
  - Continues handling the split to lkd:Work and lkd:Expression
- New way to map with RDA using annotation properties
- Numerous minor fixes
- Planned for October 2023
  
- By the end of 2023
  - Start testing the model in practice by the end of the year
    - Using a linked data catalogueing tool
  - RDA-BIBFRAME mapping
    - Mapping of RDA Work (rdac:C10001) will be ready



# Melinda

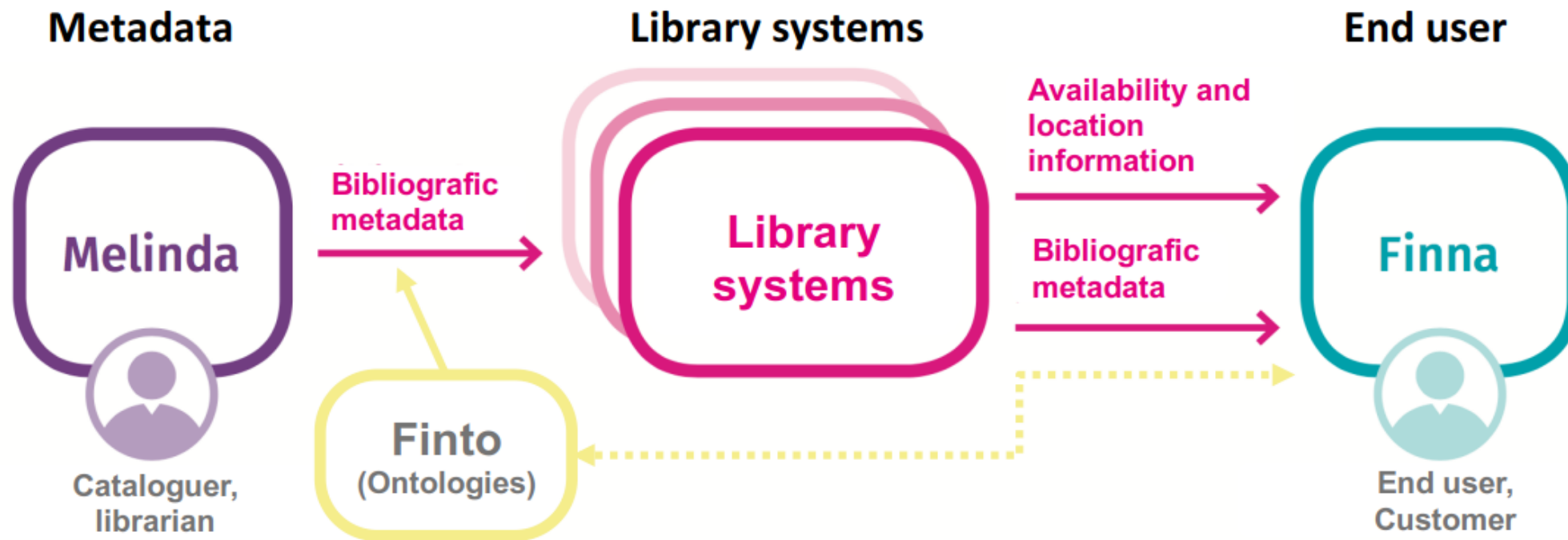
The national metadata repository  
A union catalogue

# Melinda – new platform

- Melinda is a national metadata repository
  - Shared metadata repository service for Finnish libraries
  - An essential part of the **data description ecosystem**
  - Both bibliographic and authority records
  - Most but not all libraries use this platform for cataloguing
  - Currently using the Aleph library system.
- Aims to collect all bibliographic metadata of all Finnish library collections into one database to be used by everyone
- A shared metadata repository increases the efficiency of data description work and provides better information retrieval services

# Melinda for data description – Finna for searching

Big centralized services: **Melinda** for cataloguers, **Finna** for end user



# New infrastructure for Melinda

- Strategical premises:
  - **open-source system,**
  - **supports LOD**
- Original target: the new system is in production during 2025
  - (will be postponed)
- Extra resources for planning from 2022
- In coordination with Linked Library Data Project
- Investigated open-source solutions 2021-2023
  - Libris XL (National Library of Sweden)
  - FOLIO
  - Separate BIBFRAME cataloguing tools: Sinopia, Marva

# Investigations for the New Melinda so far

- In-house groups to investigate the two most promising options in different areas (2021-)
  - Data description tool
  - Data processing
  - Integrations and APIs
  - Technical platform
- Deeper analysis of Libris XL (Spring 2022-)
  - Local test instance of XL
  - Deeper investigations in code level, github pull requests to KB
  - Deeper testing with *Libris katalogisering*
  - Discussions with Swedish National Library Libris team & management

# Current situation for New Melinda

- No decision as of yet
  - Funding negotiations with the Ministry for 2025-2028 in early 2024
- Challenges
  - Our extensive use of MARC record component parts
    - e.g. all music tracks of a music record are catalogued separately
  - Current integrations from the Melinda repository to the various local library systems are complex



# The future of library data is linked

- Great changes are expected
- Great possibilities to ensure our relevance in society
- Planning and directing are keys to success





[www.kansalliskirjasto.fi](http://www.kansalliskirjasto.fi)

**Linked Library Data project**

[github.com/NatLibFi/lkd](https://github.com/NatLibFi/lkd)

[www.kiwi.fi/x/uwHNEQ](https://www.kiwi.fi/x/uwHNEQ)

Contact information: [matias.frosterus@helsinki.fi](mailto:matias.frosterus@helsinki.fi)

