

RDA based Data Model of the Finnish Memory Organizations

RDA in the Wider World (August 11, 2016)

Marja-Liisa Seppälä - National Library of Finland (presenter)

Leena Furu-Kallio - Finnish Museums Association

Miia Herrala - National Archives of Finland

Contents

- Starting points for the metadata co-operation
- Phases in planning a shared (meta)data repository
 - Phase 1: creating a data model
 - Phase 2: planning metadata exchange
 - Phase 3: steps to implementation
- Other initiatives in the national metadata co-operation

Starting points

- In 2014, Finnish memory organizations (libraries, archives and museums) decided to align their metadata, in order to
 - prevent any overlap in cataloguing
 - improve the quality of metadata
 - support better the user interface *Finna* (shared by the Finnish memory organizations)
- Besides libraries, archives and museums, the whole public sector in Finland strives for interoperable metadata and common data architecture

Starting points (2)

- RDA was chosen to one of the recommended standards for libraries, archives and museums
- Archive and museum sectors decided to implement the RDA instructions of agent metadata, i.e.
 - Chapters 8-11: attributes of person, family and corporate body
 - Chapters 29-32: relationships between person, family and corporate body
 - Appendix K: relationship designators for person, family and corporate body

Common ground

- Comparing the concept models of the memory organizations:
 - libraries focus on resources whereas archives and museums take into account context and temporal dimension
 - easiest to start alignment with agent (person, family and corporate body) metadata
- Aiming to create a shared (meta)data repository for agent metadata (by 2019)

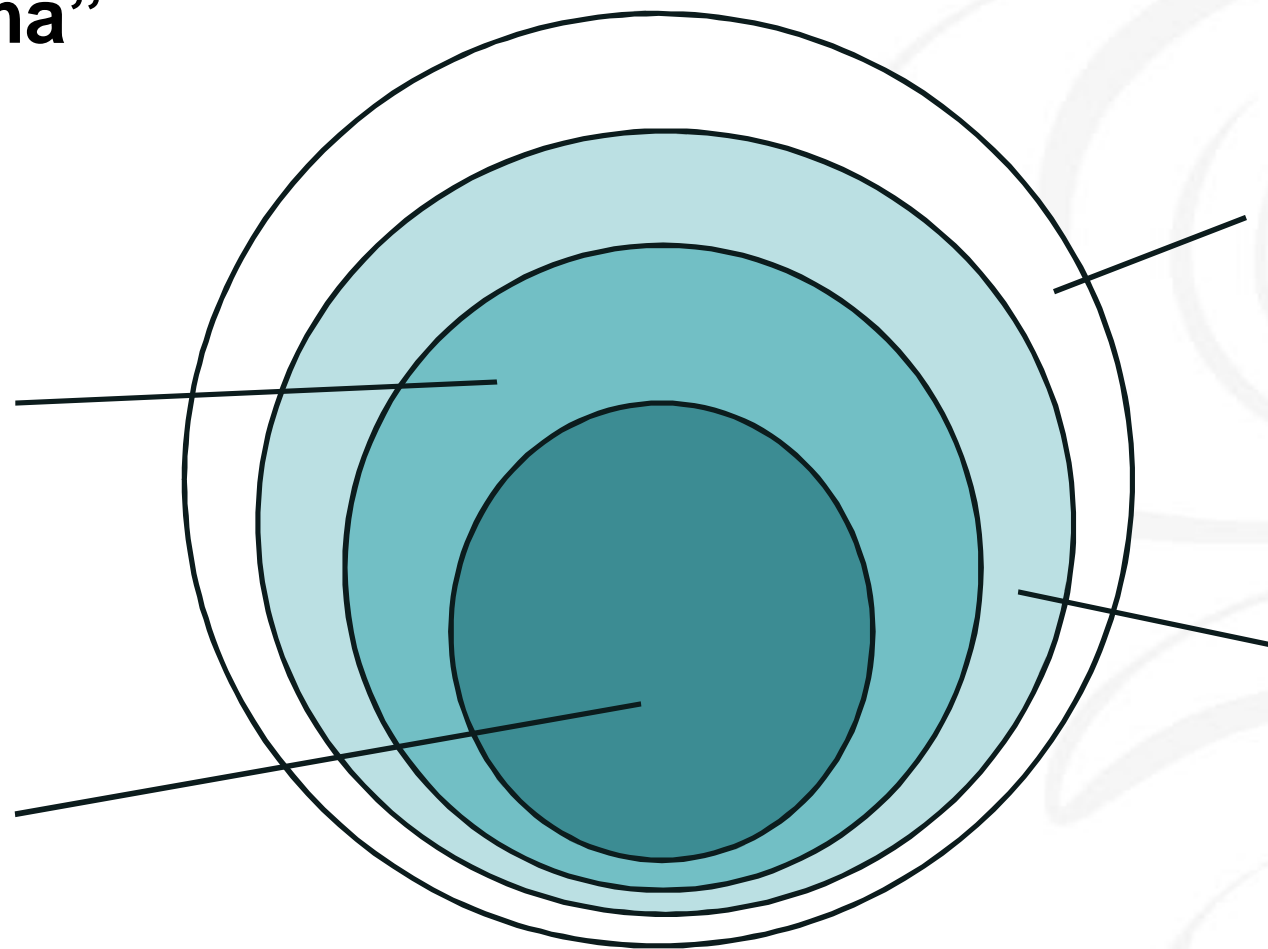
First phase: a data model

- Creating a RDA based data model for agent metadata (in 2015-2016), including
 - all the RDA (agent) elements: core and optional
 - most of the ISNI (International Standard Name identifier) elements
 - most of the EU (European Union) Core Vocabulary elements
 - elements specific only to archives, museums or libraries
- ∅ "Master schema" = all the metadata elements useful for libraries, archives and museums

"Master schema"

Elements
common to all
sectors - optional

Elements
common to all
sectors -
mandatory



Sector specific
elements -
not available to
others

Sector specific
elements -
available to
others

Sector = archive, museum or library sector

Layout of the data model

- Lists
 - of identifying and describing elements
 - of elements of administrative metadata (metametadata), e.g.
 - status of identification and maintenance history
- Lists include:
 - name, identifier and short description of an element
 - information of an element being repeatable or mandatory
 - mappings between standards (like RDA, ISNI, EU core, EAC-CPF, MARC21, Spectrum)

Layout of the data model (2)

- Detailed specifications of some elements:
 - relationship, time, place, identifier, name and display restriction
 - specifications are about:
 - type, time span, language, attributes and values
 - e.g.
 - what are the types of an element
 - which language codes are used
 - are values of an element included in the data model or are they in a separate ontology/vocabulary

The main issues in the datamodel

- Differences between memory organizations in the following matters:
 - defining an agent
 - choosing a preferred name of an agent
 - authorized access point
 - level of granularity, e.g. about time, name, type

Defining an agent

- In RDA, one (real life) agent may have several entities, e.g.
 - when a corporate body changes its name
 - when a person has several public identities
- Archives and museums create only one record for one (real life) agent including the whole history of agent in it
- ∅ Compromise:
 - ∅ follow RDA whenever a shared record is created or used by libraries
 - ∅ otherwise, follow the practice of archives and museums in the shared data repository

Preferred name of an agent

- Preferred source of information (e.g. a title page or similar source) is not always available at archives and museums
- Preferred name determined by archives and museums might be different than preferred name in RDA
- ∅ Compromise:
 - ∅ if possible, archives and museums try to record a RDA preferred name
 - ∅ if not, the element *RDA preferred name* is left unrecorded and the name is recorded to the element *Preferred name*
 - ∅ Libraries will add the *RDA preferred name* later to the record

Preferred name of an agent (2)

- Consequently:
 - lots of various types of names (preferred and others) in a record
 - the same name in different elements and different names in one element
 - no deleting names but tagging names by the sectors

Authorized access point (AAP)

- AAP not relevant to archives and museums but they are ready to record elements needed for AAP
 - e.g. *place associated with the family* which is mandatory only when needed to distinguish a family from another family with the same name.
- Instead of AAP, metadata as a whole identifies an agent at archives and museums
 - not only attributes but also relationships support identifying of an agent

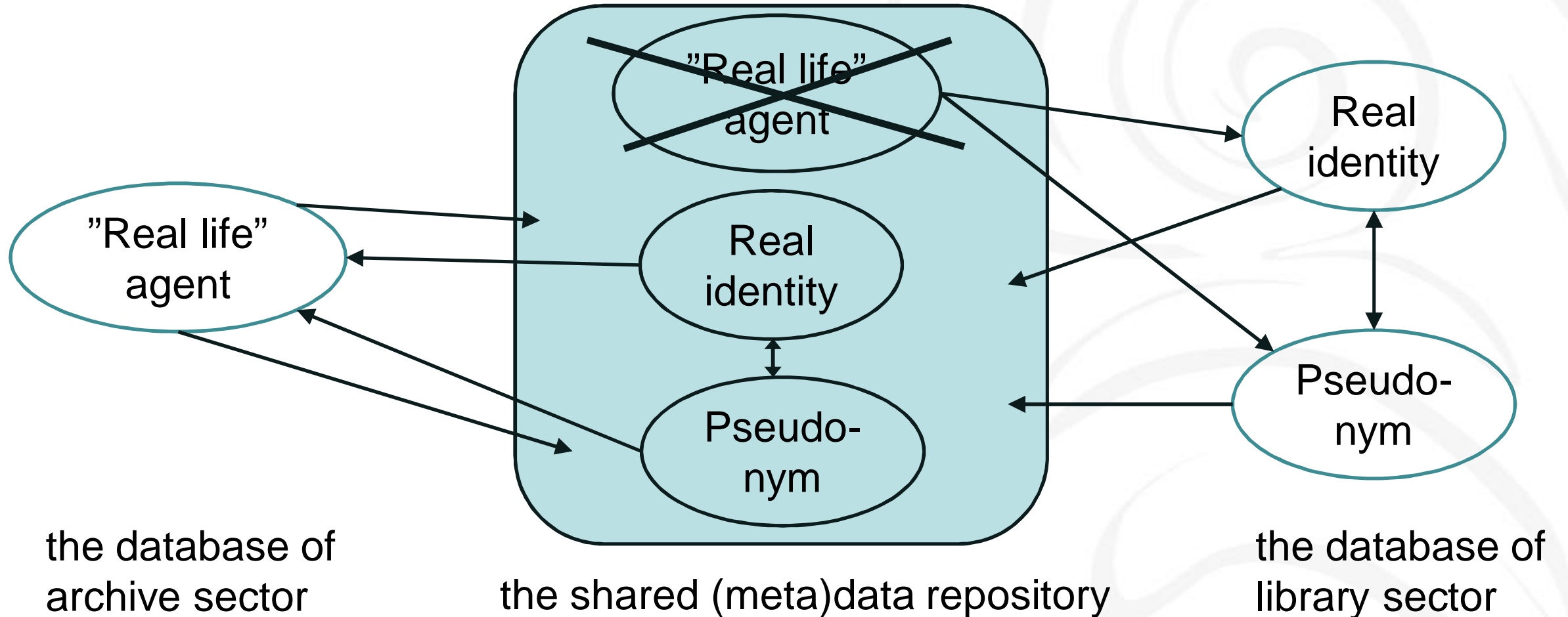
Level of granularity

- Archives need to define a type of agent more precisely than RDA
 - a separate ontology of agent types will be created and linked to the shared (meta)data repository
- Time:
 - archives and museums record months and days in addition to years
 - time elements in the data model consist of both start and end dates: not separate elements for e.g. date of birth and date of death
- Name is divided into subelements *given name* and *family name*

Second phase: metadata exchange

- Planning of metadata transfer between the shared data repository and databases of each sector (Autumn 2016 – Spring 2017)
- Differences in the level of granularity between the data model and e.g. libraries' MARC21 require detailed planning of conversions, replication rules and double control of records
- Aim is to avoid any data loss, e.g.
 - the element *official name* (in the repository) is converted to the element *variant name* (in the library database) but the *variant name* doesn't override the *official name* when it is converted back to the data repository

Example of metadata transfer: "Real life" agent vs. RDA entity



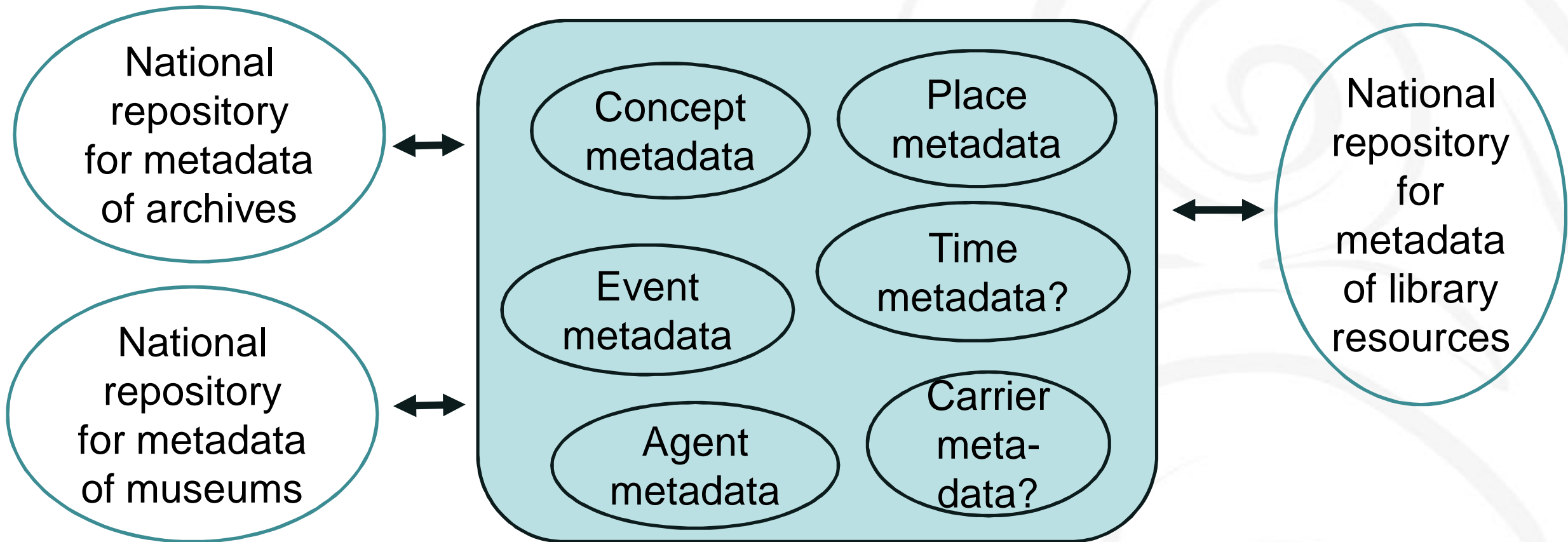
Third phase: implementation

- Planning of implementation (in 2017-2018)
- The 4 years' vision is to create a production system which could be fully integrated to cataloger's interface
 - need for a new information system free from MARC21 or other old formats
 - in addition, closed ILSs of the library sector must be renewed
- Responsibility issues:
 - “editorial board” accepts new proposals for agents and solves disagreements in cataloguing
 - every sector takes responsibility over larger variety of agents than before

Other initiatives in metadata co-operation

- National Metadata Vocabulary:
 - RDA based registry and vocabulary for libraries, archives and museums
 - includes links to the international RDA Registry
- Place ontology:
 - planning has started: co-operation through the whole public sector
 - the first challenge is to determine the entity *place*
- Event ontology?

Data architecture in future?



Shared metadata repositories in public sector

Useful links

- EU core vocabulary: <https://joinup.ec.europa.eu/category/glossary/core-vocabulary>
- ISNI standard: http://www.iso.org/iso/catalogue_detail?csnumber=44292
- User interface *Finna*: <https://finna.fi>
- National Metadata Vocabulary: <http://finto.fi/mts/en/>
- Finnish Thesaurus and Ontology service *Finto*: <http://finto.fi/en/>
- RDA Registry: <http://www.rdaregistry.info/>

Thank you!

marja-liisa.seppala(at)helsinki.fi