Metadatabrønn (Metadata Well)

National Library of Norway

Request for Information

November 2021

Table of Contents

About the National Library of Norway

Tasks and responsibility

The operations of the National Library of Norway (NLN) are governed by the Legal Deposit Act [(Pliktavleveringslova)](). The objective of the act is to ensure that all documents made publically available in Norway are deposited with NLN, so that records of Norwegian culture and society are preserved and made available as source material for research and documentary purposes.

In its role as a research library and custodian of the national collection of deposited material, NLN is a key part of Norway's research infrastructure. NLN also plays a role in cultural policy by preserving cultural heritage for posterity and by commemorating prominent Norwegian authors. Its collection constitutes a unique knowledge base for the benefit of present and future generations.

NLN operates a repository library (Depotbiblioteket) for inter-library loans.

NLN aims to make its entire collection digitally available at nb.no. Also, NLN has been given the task of digitising the audio-visual collections of archives and museums in Norway. Hence, digitisation of analogue material in any medium is a continuous and highly prioritized activity.

NLN is mandated by the Norwegian Government to function as a competence center and service provider for the whole cultural heritage sector, including libraries. This involves guiding and supporting the Norwegian libraries in their aimed-for role as active and relevant agents in their respective communities. To a large extent, NLN's duties in this respect are outlined in the policy document National strategy for Libraries 2020-2023.

Among other tasks, NLN has the responsibility to provide useful bibliographic standards and tools to other libraries, as well as free metadata of good quality and in a timely fashion. The objective is to fulfil the latter by establishing a single authoritative source of bibliographical data - a so-called Metadata Well, accessible for all via open and standardized interfaces.

Contact information

The contact point at the National Library of Norway for this procurement is Mercell.

Request for information

This RFI asks for information about possible candidate systems, infrastructures/technical solutions through which the Metadata Well may be realized.

No information contained in this RFI, or in any communication made between NLN and any Potential Provider in connection with this RFI, shall be relied upon as constituting a contract, agreement or representation that any contract shall be offered in accordance with this RFI. NLN reserves the right, subject to the appropriate procurement regulations, to change without notice the basis of, or the procedures for, the competitive sourcing process or to terminate the process at any time. Under no circumstances shall the National Library of Norway incur any liability in respect of this RFI or any supporting documentation.

The National Library of Norway will not reimburse any costs incurred by Potential Providers in connection with preparation and submission of their responses to this RFI.

The Norwegian Library Landscape

The library sector in Norway has a long history of resource and data sharing. Through establishing a shared, authoritative metadata resource, we hope to support and even strengthen that culture.

The Norwegian library landscape includes libraries of the following types:

The libraries

Public libraries: According to the Act of Public libraries, each individual municipality in Norway must provide a public library within its boundaries, its operation and collection fully funded by the municipality.

County libraries: Library services at the regional level, mostly providing services to the public libraries within the county. Range of services and collections varies greatly. Funded by the county.

School libraries: Libraries in primary, secondary or high school, for educational as well as recreational purposes. In some cases, school libraries double as public libraries.

Special libraries and research libraries: Libraries connected to companies, research institutes or universities and colleges. Varies greatly in size and nature, - ranging from 1 person services in specialized companies to large university libraries with more than 100 employees.

The National Library of Norway provides free access to high-quality bibliographic data for all libraries. The purpose is to free up staff resources in the libraries and ensure consistent, quality-assured metadata in the libraries' catalogues. The delivery consists of metadata for Norwegian publications of printed books, audiobooks, textbooks, e-books and language courses in addition to book cover images. A preliminary description is available no later than two working days after the publication has been announced and a complete metadata description including classification number, subject headings and a summary no later than two working days after publication.

At present, this is solved by procuring metadata from a commercial supplier.

Metadata suppliers and library system vendors

At least two types of corporations outside the library sector should be considered in the Metadata Well context:

Metadata suppliers: There are two commercial metadata producers in Norway, namely Biblioteksentralen AB and Bokbasen AS. Both companies sell metadata to libraries, Biblioteksentralen also sells content, for example books and films.

Library system vendors: Norwegian libraries use various library systems. The majority of academic libraries, including NLN, are gathered in the BIBSYS consortium, sharing one catalogue residing on the Alma Library Platform by Ex Libris, with Unit as the national system manager. Public libraries and small special libraries on the other hand, usually have their own local catalogue provided by one of a handful Norwegian vendors. The systems currently in use are Bibliofil by Bibliotek-Systemer AS, Micromarc and Quria by Axiell, Tidemann by Bibliotekservice AS as well as the open-source system Koha and Oslo Public Library Deichman's in-house developed system.

Negotiations on protocols, standards and interoperability issues are generally conducted within a collaboration group called Biblioteksystemleverandørene, comprising all vendors/managers of systems used in Norwegian libraries, as well as Oslo Public Library Deichman and the National Library.

The Metadata Well

Purpose

The establishment of the Metadata Well is based on measures specified in the [National strategy for libraries 2020–2023](#) from The Ministry of Education and Research and the Ministry of Culture.

The Metadata Well shall provide libraries and others with a single authorised source of uniquely identified metadata descriptions. The Metadata Well will also constitute a hub for the reuse of metadata between libraries. Suppliers of metadata and service providers can use the infrastructure provided by the metadata well to provide additional data or services.

The Metadata Well - essential features

The Metadata Well will store and exchange quality-assured and uniquely identified metadata descriptions for publications and archive material in Norwegian libraries and the LAM sector (libraries, archives, and museums). The entities referenced in metadata descriptions must be assigned globally unique and persistent identifiers so that they can be recognized throughout their life cycle and across systems.

All metadata produced by the National Library for the Norwegian National Bibliography will also be included in the Metadata Well. Other metadata providers can connect to the infrastructure, and offer supplementary services such as key words, descriptions of content, reviews etc. linking them to metadata descriptions in the Metadata Well. The Metadata Well will not include holding information.

The metadata descriptions must be in accordance with current guidelines for metadata and content description, for instance Resource Description and Access (RDA).

It must be possible to import metadata to the Metadata Well from The National Library's catalogue, from metadata providers and libraries. It must also be possible to create and edit metadata descriptions directly in the Metadata Well.

The Metadata Well must be accessible via open and standardized interfaces in several formats, including MARC 21 and linked data according to one or more ontologies.

The metadata descriptions must contain information about its provenance and degree of completeness.

Biblioteksøk ("Library Search") (see Environment and Stakeholders) can be based on data in the Metadata Well and can retrieve holdings information from the local library catalogues where the metadata descriptions are encoded with unique IDs from the Metadata Well. This will eliminate the current challenge of having to combine different entries.

Metadata in the Metadata Well are to be openly available for any user through search, download or other available means of delivery. An access control mechanism must be provided for users allowed to make changes in the data.

Environment and stakeholders

The Metadata Well will be part of an environment of existing norms, processes and systems, both internally in the National Library and in a wider context. It is important to identify elements in the environment that The Metadata Well must be able to relate to or interact with.

A metadata well as outlined in this document, represents in many ways a new method for managing shared metadata. Hence, when implementing such a resource into an existing environment, focusing solely on adapting it to this environment may not be sufficient. On the contrary, best results may be achieved when the new resource is allowed to influence and possibly cause changes to said environment and its activities.

The Metadata Well's environment, including specific stakeholders, and their potential use and benefit of the metadata well are outlined below.

Stakeholders and their potential use and benefit of the Metadata Well

Metadata descriptions in The Metadata Well must be open and accessible to everyone, ie. anyone who wants to, can develop and offer services based on the data, either for their own use or a market.

In the library sector, the following stakeholders are particularly relevant:

Libraries

- will have a single authorised source of quality-assured metadata,
- will be able to reuse metadata from other libraries,
- will have metadata automatically updated from a central source,
- will be able to purchase services related to the Metadata Well from external service providers, for example, reviews.

Metadata providers

- can offer additional data or services related to metadata descriptions in the Metadata Well, such as reviews and subject headings.

Bookstores

- can link metadata from The Metadata Well to books they offer.

Library system vendors

- can develop services based on data in The Metadata Well.

Service providers

- can provide services that are based on metadata descriptions in The Metadata Well, such as book covers, subject headings, reviews, and other additional services.

Universities, colleges and public bodies

- will have a single authorised source of quality-assured metadata on publications in one place,
- can also make metadata descriptions about their own publications available to users in the Metadata Well.

The National Library

- will have a central place of delivery of metadata to libraries and other stakeholders,

- must integrate the production of metadata for the Metadata Well with the import of data from publishers and the production of metadata for its own catalogs and bibliographies,

- can use the metadata well in its own services, for example by allowing Biblioteksøk ("Library Search") and the National Bibliography to build on quality-assured metadata from a single source and receive automatic updating of authorized data elements.

Special services

Biblioteksøk ("Library Search")

The National Library's Bibliteksøk ("Library Search") provides a point of access to all Norwegian library catalogues as well as allowing the placement of orders for loans. Norwegian library staff and borrowers with a national library card (Nasjonalt lånekort) can order books and other material available for lending. The books and material ordered can be collected at the nearest library. The order will be processed based on criteria such as availability, geographical location and the number of previous loans. The National Library, particularly through its repository library (Depotbiblioteket) is a key supplier to Bibliteksøk ("Library Search"). Bibliteksøk ("Library Search") also provides access to books in the National Library's digital collection.

BaseBibliotekBaseBibliotek is a comprehensive register of all Norwegian libraries (and their foreign interlibrary loan connections) that have been assigned a Norwegian library number. The library number is used to identify libraries in services that libraries use and can also be used as identification against the Metadata Well.

Requirements to the Metadata Well

Basic functionality

Identity management
The Metadata Well must have an authentication and authorisation service to control rights to deliver, update and use metadata.

Search interface
The Metadata Well must have a search interface that allows for free text search, advanced search and search refinement.

Metadata editor
The Metadata Well must have a metadata editor where new metadata descriptions can be created and existing ones deleted or corrected, e.g., by merging duplicates.

Interaction with authority files and vocabularies
The Metadata Well must be interoperable with authority files and vocabularies. When using authority files and vocabularies managed by NLN, full integration should be aimed for, see also Use of authorities below.

Batch update
The Metadata Well must support batch update. See also Validation.

Support for data conversion
The Metadata Well should support data conversion between an entity-based format and MARC 21, and the other way around.

Data access and ingest
A variety of mechanisms by which to access and ingest data to the Metadata Well should be offered:

- Search and retrieval through a graphical user interface
- Dedicated APIs for search, retrieval, updating, creation and deprecating of data
- Data dumps for download
- Facilities for harvesting sets of data
- Facilities for uploading sets of new and modified data

For information about data formats and exchange protocols to be supported, see the relevant subsections of Formats and standards.

Validation

Validation of incoming data is essential, - the validating mechanism must be *flexible* yet *reliable*, and preferably support validation of incoming data according to expressed constraints/rules, see also Expressing quality requirements below.

Specific attention should be paid to preventing duplicate entities, e.g., that same publication exists as more than one entity.

Data representation in the Metadata Well

Entity based representation aligned with RDA/LRM

The internal representation in the Metadata Well should preferably be entity based as opposed to record based, but we are also open to other solutions. The entity model to be used in the Metadata Well must be well aligned with that of RDA and thereby Library Reference Model (LRM). Hence, the Metadata Well should allow for *explicit representation of* or *reference to* bibliographical entity types like items, manifestations, expressions, works, agents and other entities, like the ones defined in RDA and LRM.

Use of authorities

The entities in Metadata Well must reference authorities from designated authority files, where possible. Several authority files will be made available, including:

- For agents, the Norwegian Authority File: Persons and Corporate Bodies, to be used as the primary source of information about agents connected with documents in the Metadata Well.
- For original works (as defined by RDA), National Authority Files for Works, to be used as the primary source of information about original works manifested by documents in the Metadata Well. Note: This is still work in progress.
- For subject headings etc., vocabularies available through nbvok, NLN's vocabulary service, among others Norwegian Thesaurus on Genre and Form
- For classification, Dewey Decimal Classification as serviced through Norsk WebDewey

Entities in the Metadata Well may also reference authorities from external sources, like National Place Name Register (SSR) from Norwegian Mapping Authority, Humord from University library of Oslo, certain Marc vocabularies a.o.

Any reference to entities from some authority service must also refer to the authority source itself.

Global identifiers

The metadata must contain persistent, globally unique identifiers for all referenced entities, including

- manifestations, works and other LRM-like entities in the Metadata Well entity model
- authorities hosted in separate registries
- the descriptions themselves (depending on the entity model, see The concept of descriptions below)

Identifiers for entities generated within the Metadata Well must comply with the URI syntax specified by NLN.

Provenance information and quality issues

At any given time, the Well's bibliographical data will exist in many different degrees of completeness and quality, - about resources at different stages of their life cycle. Below are outlined some important considerations in that respect.

Prepublication metadata

The Metadata Well must allow for prepublication metadata, that is, metadata issued by commercial publishers about their publications-to-come. Prepublication metadata constitute initial bibliographical descriptions, to be corrected and expanded as soon as the described documents are published.

The concept of descriptions

Although bibliographical *entities* - e.g., manifestations - constitute the Metadata Well's primary content objects as opposed to bibliographical *records*, the notion of (bibliographical) *descriptions* is still an important one, in the sense that it must be possible to represent information *about the descriptions as such*. Thus, the data model must allow for explicit representation of "data about metadata", like provenance, lifecycle status (e.g., new/updated/deleted), completeness (e.g., short, full...) and other administrative information.

The Metadata Well must also enable representation of *subsets* of descriptions. For example, a manifestation might be described in the Metadata Well using a large number of properties, not all of which are relevant in all usage contexts. In particular, being able to declare a core set of information elements (about some manifestation) as eligible for the national bibliography, will be essential. Whether this should be solved by representing such descriptions explicitly, or by other means (e.g., application profiles and/or dedicated APIs) is a matter for further investigation and discussion.

Expressing quality requirements

The data in the Metadata Well will most likely originate from many sources. Consequently, validation of incoming data becomes crucial. On the other hand, not all incoming data can be made subject to the same quality requirements. Hence, a way of expressing quality requirements explicitly must be provided, for example some language or scheme by which constraints and conditions to the data may be expressed and updated as needed.

Formats and standards

NLN is mandated to regulate and provide crucial infrastructure elements for Norwegian libraries, including cataloguing rules, classification system, as well as metadata exchange format and -protocols. Our main policy in the metadata field is to follow international standards as much as possible.

Cataloguing rules

Currently, RDA is used as cataloguing rules, while MARC 21 is the chosen metadata exchange format.

Hence, bibliographic descriptions in the Metadata Well must comply with the rules given in the RDA standard, or RDA adaptations applicable to specific material types, like FIAF for cataloguing moving images. Note that the phrase "RDA rules" also includes any national guidelines for Norway.

Data formats

Our basic assumption is that MARC 21 will be used in (some) Norwegian libraries for many years to come. At the same time, an increasing number of libraries are investigating the opportunities involved in migrating from MARC 21 to linked data. Therefore, the Metadata Well must be able to import and export both MARC 21 records, as well as metadata in the form of linked data, although decision on

which ontologies or application profiles to support (e.g., BIBFRAME, RDA Vocabulary, Schema.org) is yet to be taken.

Internationally, considerable effort has been put into developing linked data-based replacements for MARC. BIBFRAME, originating in the Library of Congress, is beginning to gain some usage, especially by libraries in North America. RDA has defined its own linked data metadata vocabulary, which is more stringent and comprehensive than BIBFRAME, and is more compatible with the IFLA Library Reference Model (formerly FRBR). Outside the library world, BIBFRAME and RDA are less known, whereas schema.org and Dublin Core are well known and used for metadata purposes in many contexts.

Data exchange protocols

For data ingest to and delivery from the Metadata Well, both OAI-PMH, SRU and preferably also SPARQL should be supported.

Developing the Metadata Well - main approach

The Metadata Well as described in this document will probably have strong effects on the library metadata ecosystem in Norway. Realizing such a service is a large endeavour and must be performed in a controlled and incremental fashion. Although the realization process is not yet detailed, we foresee 4 main phases, reflecting an approach of starting with the simpler metadata:

1. Establish a minimal version of Metadata Well, handling metadata created by running cataloguing. This involves developing a minimum viable product (MVP, described below) for testing in NLN and a limited number of metadata producers and users outside NLN.
2. Expand the minimal version to obtain a functional level perceived as acceptable for production. Based on the results and experience from phase 1, amendments and expansions are made, to include all necessary facilities for production in a full-scale user/stakeholder environment.
3. Import legacy metadata into the Metadata Well. This involves converting legacy bibliographical descriptions from NLN and other libraries to Metadata Well's entity model.
4. Maintenance and further developments according to future needs.

Minimum viable product (MVP)

The MVP of the Metadata Well must be able to provide metadata for all publication types covered by NLN's present metadata service, that is Norwegian publications of printed books, audiobooks, textbooks, e-books and language courses, published from a given point of time onwards.

Furthermore, the MVP must comply with the following requirements:

- The metadata must contain persistent, globally unique identifiers for all referenced entities, as described in Global identifiers above.
- The MVP must enable export of bibliographical descriptions in MARC 21 as well as linked data according to one given ontology (to be specified)
- The MVP must offer facilities for manual editing of metadata
- The MVP must offer APIs for retrieving as well as inserting/updating metadata
- The MVP must provide authentication and authorisation services, advanced enough to enable a few selected external metadata providers to participate in phase 1.

Appendix: Process diagram

Non-disclosure agreement

The National Library of Norway will treat all information submitted as confidential. However, ideas submitted may be used by the National Library of Norway in connection with future procurement processes or project work, and may also be presented in an anonymized format at conferences and dialogue with market actors, or in tender documents, etc. If necessary, the National Library of Norway reserves the right to contact persons or organisations who respond to this RFI to ask follow-up questions or to obtain clarifications on feedback submitted.