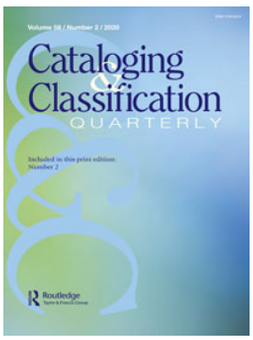

Cataloging & Classification Quarterly

ISSN: 0163-9374 (Print) 1544-4554 (Online) Journal homepage: https://www.tandfonline.com/loi/wccq20

The Public Library Metadata Landscape, the Case of Norway 2017–2018

Michael Preminger, Ingebjørg Rype, Marit Kristine Ådland, David Massey & Kim Tallerås

To cite this article: Michael Preminger, Ingebjørg Rype, Marit Kristine Ådland, David Massey & Kim Tallerås (2020) The Public Library Metadata Landscape, the Case of Norway 2017–2018, Cataloging & Classification Quarterly, 58:2, 127-148, DOI: 10.1080/01639374.2020.1711836

To link to this article: https://doi.org/10.1080/01639374.2020.1711836

© 2020 The Author(s). Published with license by Taylor & Francis Group, LLC

Published online: 10 Feb 2020.

Submit your article to this journal

Article views: 268

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=wccq20

CATALOGING & CLASSIFICATION QUARTERLY
2020, VOL. 58, NO. 2, 127–148
https://doi.org/10.1080/01639374.2020.1711836

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

The Public Library Metadata Landscape, the Case of Norway 2017–2018

Michael Preminger[a], Ingebjørg Rype[b], Marit Kristine Ådland[a], David Massey[a], and Kim Tallerås[a]

[a]Department of Archivistics, Library and Information Science, Oslo Metropolitan University, Oslo, Norway; [b]Department of Acquisition and Bibliographic Services, National Library of Norway, Oslo, Norway

ABSTRACT

The aim of this paper is to gauge the cataloging practices within the public library sector seen from the catalog with Norway as a case, based on a sample of records from public libraries and cataloging agencies. Findings suggest that libraries make few changes to records they import from central agencies, and that larger libraries make more changes than smaller libraries. Findings also suggest that libraries catalog and modify records with their patrons in mind, and though the extent is not large, cataloging proficiency is still required in the public library domain, at least in larger libraries, in order to ensure correct and consistent metadata.

ARTICLE HISTORY
Received November 2019
Revised January 2020
Accepted January 2020

KEYWORDS
Metadata; cataloging; metadata practices MARC formats; boundary objects

Introduction

The international cataloging landscape is changing, with consequences for local metadata practices. In 2019, Norway replaced the long-lasting cataloging standard *Anglo-American Cataloging Rules* 2nd edition (AACR2) with the newly translated *Resource Description and Access* (RDA). Public, academic, and national libraries are experimenting with – and in some cases building – new systems based on linked data principles, for example at the Oslo public library. As a result of new standards and organizational models the aging, manifestation-oriented MARC formats are under pressure.

As early as 2006, Roy Tennant, at that time User Services Architect at the California Digital Library, envisaged a future where "[..] the modern cataloger will one day be a software enabled specialist who can gather, subset, normalize, and enrich piles of records for a specific audience or purpose."[1]

The traditional trades of librarianship like cataloging, classification, reference work, etc. are still taught at library and information science (LIS) schools, but to a lesser extent.[2] Additional trades, such as teaching,

CONTACT Michael Preminger michaelp@oslomet.no Oslo Metropolitan University, Oslo, Norway.
© 2020 The Author(s). Published with license by Taylor & Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

leadership, information and communication technologies, and the like, have gained relatively higher prominence in the curriculum. This trend has been apparent for over a decade both internationally and in Norway. Dahl, Knutsen, and Tallerås[3] have advocated for broadening the understanding of the metadata life-cycle when educating catalogers, something that has had consequences for the cataloging education at the LIS department of Oslo Metropolitan University, Norway's largest LIS school. Research carried out in Norway indicates that librarians have been working on records locally, both cataloging and classification, but to a lesser extent than previously,[4] as the National Library of Norway has restricted allowed modifications to centrally cataloged records that are made available to the central library search portal Biblioteksøk to mainly adding additional subject headings or Dewey codes.[5]

Within this landscape of public library cataloging, we started with this project, exploring the following research questions:

- What has characterized metadata and cataloging practices in Norwegian public libraries in 2017–2018?
- What are the central quality attributes that affect these practices?
- To what extent is there need for cataloging competencies in public libraries?

Using metadata as a lens through which to regard practice, the main research data used for addressing these questions are 116,029 NORMARC[6] records downloaded from the catalogs of 49 public libraries of varying sizes, as well as 21,275 records from two central cataloging agencies. The record sampling criterion is expressed by the search index dc.date with value 2017. The publication date was our only sampling criterion; this guaranteed that both print and non-print resources are present in our sample. Here we were studying the two central agencies, as well as the client libraries as communities of practice interacting with these records, and to a lesser extent also the implications of various practices on usability.

In the remainder of the paper, after a discussion of the theoretical background of the work, we describe the Norwegian public library metadata landscape, and the data sample representing it. We follow with a thorough data analysis accompanied by examples illustrating important aspects of cataloging practices in light of the theory. We conclude the paper by drawing some implications for education and practice and suggesting paths for further research.

Theory and related work

MARC and metadata quality

In recent years, metadata quality has often been discussed in a digital library context. Greenberg and Robertson[7] define metadata quality as "accurate,

consistent, sufficient, and thus reliable." Robertson comments that the sufficiency idea is double-edged, referring "back to the primary and overriding definition for quality in any setting: fitness for purpose — as true for metadata as it is for designing a car or boiling an egg."[8] In this context, Robertson says that "there is an acknowledgement that a nearly perfect record is possible. There are also mechanisms which allow libraries to buy or exchange this agreed 'perfect' minimal record from external sources to reduce the volume and cost of in-house cataloguing."[9] A key remark here is that "[m]echanisms such as this can exist because the library community has shared purpose and conception of metadata quality, which allows an agreed 'level'" for exchange."[10]

Bruce and Hillmann examine seven of the most commonly recognized characteristics of metadata quality: completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness and accessibility.[11] The provenance aspect is particularly interesting in our context, applicable not only to the creating agency but also "what transformations have been applied to the data since it was created."[12] Conformance to expectations is also relevant.

The future of the MARC format has been under discussion for over a decade. Critics point to the fact that the format, originally intended for the production of printed catalog cards, is by nature strict and not very hospitable to exchange between local and union catalog systems.[13] Tennant has put at least some of the blame for the problems of the library catalog on the MARC format and its limited coverage.[14] The phrase "MARC must die" has been heard in library conferences for nearly two decades.[15]

But MARC has also had its proponents. Guenther[16] claims that many of the caveats Tennant lists are not due to the format itself, but a result of the policies of cataloging institutions. Indeed, MARC has exhibited a remarkable resilience, still fueling catalogs worldwide. Smith-Yoshimura et al.[17] carried out a study of MARC field usage in WorldCat, attempting to gauge the implications of field usage on library practices. The aim of the study was to "[..] inform more efficient and effective MARC metadata creation practices."[18] Possible criteria for "good enough cataloging," as they see it, may be discovery of known or unknown items? Machine or human matching? Discovery of all manifestations of a given work? Interpreting the potential value of an item for a user's needs? Limiting or faceting search results? Delivering content? Facilitating machine processing and manipulation?[19]

One interesting understanding of the demands for modern library metadata is that "[w]ith more text being indexed by search engines, focus should be on the authorized names, classifications, and controlled vocabularies that keyword searching of full-text will not provide."[20] Hjørland[21] seems to agree here, making the claim that, "[b]oth automatic and manual indexing systems are often supposed to be neutral and the best one for any purpose! It is this

positivist epistemology that I believe we need to challenge. We need to prove that Google, for example, is not always good enough."[22]

In the years 2013-2018, with reference to Smith-Yoshimura et al.[23] OCLC were "Ground- truthing" MARC Usage across all WorldCat libraries.[24] to find "which elements and subfields have actually been used, and more importantly, how?" This work was carried out mainly to inform decisions about "where we go from here," which among other activities should support decisions about transitions to new formats.

Educating catalogers

After being a core competency in LIS education institutions worldwide, the need to educate catalogers and classifiers has been under debate. While there may not be many voiced opponents to such an education, the development "on the ground" shows for example that many ALA- accredited LIS schools have, through the 2000s reduced cataloging education from being a core competency required for all students to introductory courses in bibliographic control. Proponents like Gorman[25] have emphasized the understanding element of such an education, and its significance for better bibliographic control. In Norway Dahl, Knutsen, and Tallerås[26] observed the need to broaden the perspective of cataloging into metadata education. The Bachelor curriculum at Oslo Metropolitan University has in recent years "lifted" the perspective of its former "cataloging" education away from "handcraft" into understanding of metadata practice in context, particularly involving more formats/rule-sets and types of description.

The NORMARC standard as a boundary object

In this paper the NORMARC standard on which creation and manipulation of bibliographic records in Norway are based, is regarded as a boundary object among different stakeholders, in our case cataloging agencies, individual libraries and end-users (See Figure 1). Boundary objects were introduced and defined by Bowker and Star in connection with classification systems and standards, referring to different communities usage of these.[27] Using boundary objects to research LIS related practices has previously been done, particularly for classification[28,29] but also for other artifacts, like the FRBR group 1 entities.[30]

Though the satisfaction of end-users is (or ought to be) the most relevant indicator for the success of the standard, various stake-holders develop their own way of following the standard that may or may not be directly related to user satisfaction. "For a boundary object to be most useful, it needs to serve as a mechanism for translation between all the different

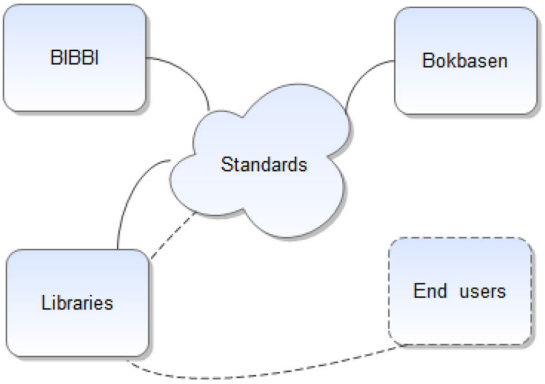

Figure 1. Standards as boundary object among stakeholders.

communities of practice who are using the boundary object."[31] But serving as such a mechanism does not necessarily presuppose consensus. "Boundary objects mediate between different groups; they do not provide a common understanding, or consensus between participants. Instead, they serve a dual function similar to that of geographic boundaries."[32]

To approach this, some trends have been identified in the data sample. These include differences of practice between the central agencies as well as differences of practices among libraries based on their size in terms of FTE (full time equivalents), when it comes to modifying imported records and producing new records. We have also tried to look at field usage differences between different types of records.

In this article, quality is defined as "Fitness for use." An original definition by Juran states that "the user of a product or service should be able to count on it for what they needed or wanted to do with it."[33] Users here means end-users, but in our study a somewhat more general interpretation is applied that also encompasses libraries and cataloging agencies.

To this end, the sample has been analyzed from the following points of view:

- What is the "typical" MARC record seen from the point of view of libraries and cataloging agencies (what fields are being used, and how often do they occur in relative terms)?
- What changes do libraries make to records and why (what "fitness for use" ideal do they aspire) do they make them?

The Norwegian public library metadata landscape

Cataloging actors

According to Statistics Norway,[34] in 2017 there were 674 public library departments in Norway, with a total of 1752.6 FTEs. The number of

distinct catalogs is estimated to just above 400 (we found no historical statistics for 2017 covering this).

Norwegian public libraries have traditionally purchased centrally cataloged records from the BIBBI database of Biblioteksentralen,[35] which is a public company cooperatively owned by Norwegian municipalities and the library association, delivering books and other library related products in addition to metadata. In 2016 the National Library, acting as a directorate under the ministry of culture, entered a cooperation with Bokbasen AS[36] (hereafter referred to as Bokbasen), for the purchase of centrally cataloged records of books published by Norwegian publishers.[37] Bokbasen has its origins in the Norwegian book database established in 1984 by the Norwegian publisher organization and contains information on all books published in Norway. It is now a corporation, majority-owned by a number of publishers, providing metadata and services based on the book database to a range of customers. Bokbasen also delivers other types of records to subscribing libraries through their integrated library system (ILS) vendors.[38] Libraries using the ILS Bibliofil share Dugnadsbasen, a common repository of catalog records they may contribute to and consume from.

Our data samples

Our original goal was to download one calendar year (2017) of *metadata production* from the agencies as well as from a representative sample of libraries. Unfortunately, the metadata production date (control-field 008, pos. 0-5) showed not to be reliable (blank for many records). We therefore resorted to selecting one year of production by the *publication date* of the cataloged manifestation. We ended up downloading:

- One year of metadata growth from a sample of 49 Norwegian public libraries, represented by the value 2017 of the search index *dc.date* (equivalently *260 $c 2017)
- Central agency production with the same criterion, including (but not limited to) for each imported library record its source record

This allowed us to analyze both the differences between the central agencies Biblioteksentralen and Bokbasen, and the differences between the agency records and the respective library records.

Libraries

The 49 libraries from which records were downloaded, comprise a sample of the Norwegian public library landscape, both geographically and by

Table 1. Our library sample breakdown by FTE classes, total #records pr. class and mean #records pr. library.

Class(FTEs)	#Libraries in class	tot. #records	mean #records pr. library
0–2	15	25,815	1,721.0
2–5	12	21,692	1,807.7
5–10	10	23,912	2,391.2
10–20	8	25,462	3,182.8
>20	4	19,148	4,787.0
Total	49	116,029	–

size[39] (see Table 1). Since the intention was to download all applicable records from the selected libraries, processing costs dictated limiting the number of libraries to around 50, of various sizes.

Since 2017, the requirements from records included in the National Library search portal, practically discourage libraries from changing the agency-provided records let alone local data like shelf marks. Although in small volume, libraries seemingly still manipulate the downloaded records, and our hypothesis is that this is done to satisfy a fitness-for-use ideal they see, which is not always met by the quality standard imposed and followed by the central agencies. As automatic detection of record modification is encumbered by lack of confidence, our estimate is therefore conservative.

Cataloging sources

For the central agencies, all the dc.date = 2017 records were downloaded, producing the distribution detailed in Table 2.

Data analysis

In this section some traits of the material are reported, from which one can extrapolate some core information about the 2017-2018 public library metadata landscape. The presentation first treats the central agencies and the library catalogs separately, and then presents and discusses field usage statistics where the actors are treated jointly.

Central agencies

We have downloaded all dc.date = 2017 records from both central agencies, giving us 6,072 pairs (same manifestation) and 9,130 disjoint records (Table 2). Of these 6,072, 5,815 were directly comparable. The remaining records consisted of records with multiple ISBN numbers that we chose to leave out of the direct comparison.

Table 2. Central agency record sample. The last column "#Imported from" gives the total number of library records imported from the agency.

Agency	# Disjoint records	# Parallel records	Total	# Imported from
BIBBI	6,629	6,072	12,701	89,153
Bokbasen	2,501	6,072	8,573	16,149
Total	9,130	12,144	21,274	

Table 3. Classification by agencies.

(a) Summary of classification comparison between agencies

Different main classes	Similar main classes	Similar divisions	Similar sections (and beyond)	Similar	Total
142	114	170	624	1,477	2,527
5.6%	4.5%	6.7%	24.7%	58.4%	100%

(b) Differences of subject code between agencies

Code		VS		#
900–999	History	310–390	Social Science	26
900–999	History	other classes		24
300–309	Social sciences, sociology & antropology	other parts (part. 900)	000–999	21
362.1–4	Social medicine	610–618	Medicine health	13
650	Management & public relations	100, 300, 400	Philosophy and psychology, Social sciences, Language	12
030, 080	Encyclopedia & books of facts, Quotations	other classes		6
370	Education	other classes		4
Others				35

Dewey Decimal Classification

In Table 3(a), *different main classes* mean that the agencies interpret the main class differently. *Similar main classes* mean agencies agree on the main class but diverge on the secondary class. *Similar section* means a divergence in the secondary class (e.g., under 320 for political science, 321 for system of governments and states vs. 324 for the political process). In as many as 24.7% of the records, the section interpretation is similar, but diverges further down the hierarchy. Many of those cases exhibit different levels of specificity rather than different interpretations.

Divergent classification may be due to either different interpretations of the tables or a failure to adhere to the standard on one or both sides. As seen from Table 3(b), few subjects stand for a majority of the differences. Some differences seem as lack of adherence to rules, for example that textbooks in certain subjects should be classified by the subject of the book itself, not by 372.3–372.8. Other differences express divergent interpretation. An example of such a difference is whether the focus is on social provision of services to people with physical illnesses (360), or on technology of medical services (610).

As seen from Table 3, it is 300 (Social sciences) and 900 (History) that stand out when it comes to determining the main class. One example is "Dagbøker fra Jerusalem" (Diaries from Jerusalem), classified as 320.956 and 956.9422, respectively, where the question is whether this is a book about Israeli politics or a multidisciplinary book about Israel. Biographies also exhibit a difficulty. A biography of people associated with a specific subject should be with the subject, but sometimes a person is associated with multiple subjects.

We have not identified any particular patterns of interpretation distinguishing the agencies.

The agencies seem to have comparable practice regarding classification. We do not know whether this is a consequence of the tender referred to (see note 37) or similar practices inherent to the agencies. A comparison of an older set of parallel records might have shed some light on this.

Added entries

Added entries are access points that in NORMARC, just like in MARC 21, are recorded in 7XX field tags.

The agencies differ in their treatment of added entries for short story collections. BIBBI catalogers seem to use detailed analytical entries, whereas Bokbasen catalogers tend toward a monolithic treatment. This is apparent in Table 4(a) and (b) where Bokbasen does not use 700 $*t* at all, but a significantly higher number of relator codes (700 $*e*) instead. It is worth mentioning here that RDA permits both types of practice.[40]

One example of these apparent differences of practice, is the collection of folk-tales by Asbjørnsen and Moe (Table 5), where the BIBBI catalogers (to the left) assigned an analytic added entry to each tale, including the title ($*t*) and the responsible person, whereas Bokbasen only assigned a non-analytic added entry to Jørgen Moe (interpreting him as co- author) and the illustrator Theodor Kittelsen. Another interesting feature of the

Table 4. Use of added entries by agencies – Difference in registering added entries between agencies (across 5815 books cataloged by both agencies).

(a) total occurrences of subfields

Agency	700 $*a*	700 $*e*	700 $*t*	740 $*a*
BIBBI	5,844	4,574	1,272	775
Bokbasen	5,632	5,623	0	693

(b) Records with at least one occurrence of subfield

Agency	700 $*a*	700 $*e*	700 $*t*	740 $*a*
BIBBI	3,354	3,262	179	571
Bokbasen	3,667	3,664	0	417

Table 5. The collection of folk-tales by Asbjørnsen and Moe, BIBBI (left) vs. Bokbasen.

```
700 ^0 $a Moe, Jørgen
       $d 1813-1882
       $e medforf.
       $j n.

700 ^0 $a Kittelsen, Theodor
       $d 1857-1914
       $e illustr.
       $j n.
700 ^2 $a Asbjørnsen, Peter Christen
       $d 1812-1885
       $j n.
       $t Gullslottet som hang i lufta
700 ^2 $a Asbjørnsen, Peter Christen
       $d 1812-1885
       $j n.
       $t Smørbukk
700 ^2 $a Asbjørnsen, Peter Christen
       $d 1812-1885
       $j n.
       $t Herreper
```

```
700 10 $a Moe, Jørgen
       $d 1813-1882
       $e forf.
       $j NO
       $8 m
       $0 (NO-TrBIB)90053857
700 10 $a Kittelsen, Th.
       $d 1857-1914
       $e illustr.
       $j NO
       $8 m
       $0 (NO-TrBIB)90055908
```

Â(hat)-symbol after the field tag denotes unused (blank) indicator.

Table 6. Short stories by Tolstoy, excerpts. BIBBI (left) vs. Bokbasen.

```
700 ^2 $a Tolstoj, Lev Nikolajevitsj
    ^2 $d 1828-1910
    ^2 $j r.
    ^2 $t Streiftoget
700 ^2 $a Tolstoj, Lev Nikolajevitsj
    ^2 $d 1828-1910
    ^2 $j r.
    ^2 $t To husarer
700 ^2 $a Tolstoj, Lev Nikolajevitsj
    ^2 $d 1828-1910
    ^2 $j r.
    ^2 $t Albert
700 ^2 $a Tolstoj, Lev Nikolajevitsj
    ^2 $d 1828-1910
    ^2 $j r.
    ^2 $t Polikusjka
700 ^2 $a Tolstoj, Lev Nikolajevitsj
    ^2 $d 1828-1910
    ^2 $j r.
    ^2 $t Ivan Iljitsj_ død
```

```
740 0^ $a Streiftoget
740 0^ $a To husarer
740 0^ $a Albert
740 0^ $a Polikusjka
740 0^ $a Ivan Iljitsj_ død
```

Bokbasen record, is the use of a name authority ID in $0. This is apparently a choice taken by the agency, here "bending" the standard as the subfield is not a part of the NORMARC format. Another example of the differences regarding analytical entries is the cataloging of Tolstoy's Short Stories (Table 6), where BIBBI (left) provides analytical Personal name (700) entries for each of the stories, whereas Bokbasen provides Non-analytical title (740) entries.

Libraries

The analysis of the library records consisted mainly in looking at modifications done by libraries to agency records, relating the extent of

Table 7. Our library sample breakdown by FTE classes, total #records per class and mean #records per library.

Class (FTEs)	#Libraries in class	tot. #records	mean #records pr. library
0–2	7	25,815	1,721.0
2–5	7	21,692	1,807.7
5–10	5	23,912	2,391.2
10–20	3	25,462	3,182.8
>20	2	19,148	4,787.0
Total	24	50,971	–

Table 8. Changes introduced to a sample of fields.

(a) Changes introduced by libraries to agency records

FTE	# libraries in group	average #changes pr. library	average percent of recs change
0–2	7	46	3,1
2–5	7	32	2.1
5–10	5	48	2.4
10–20	3	205	6.2
>20	2	689	14.2

(b) Import vs. changes from agencies

changes	total import	agency
2,446 (88%)	44,523 (93%)	BIBBI
337 (12%)	3,431 (7%)	Bokbasen AS

modification to library size. We present examples from two types of modifications. Additionally, we were gauging in-house record production and editing, looking at records that are not branded by either of the central agencies.

Sampling modifications

As some of our libraries did not have a ”BIBBI” subscription, and thereby received a smaller portion of each BIBBI record upon purchasing an item (something that was brought to our attention late in the process), the modification analysis needed to rely on material from libraries possessing a subscription. Thus, the data is somewhat indicative in nature. We have attempted to maintain a more or less equal number of records between the size classes, but the totals are smaller (see Table 7), making the derived statistics somewhat less confident.

Whereas for repeatable fields modification counts would be unreliable (and would demand great manual effort), counting modifications to non-repeatable fields provided us with more reliable modification statistics. The juxtaposition of agency record vs. respective library record relied on parallel field-tag, indicator-values, subfield codes and subfield value match. Automatic record modification routines were filtered according to information we had gathered from the cataloging agencies, so that we, as far as

Table 9. Changes introduced to Target audience (019 $*a*).

FTE	# libraries in group	average #changes per library	average percent of recs change
0–2	7	3	0.2
2–5	7	2.8	0.18
5–10	5	6.2	0.3
10–20	3	38.3	1.16
>20	2	456	9.42

possible, were only left with modifications attributable to the effort of library staff. In Table 8(a), we list the number of changes introduced to subfield $*a* of a selection of non-repeatable fields (019, 100, 240, 245, 260, 300, 520, and 546). For the 5xx (note fields) we set a minimum threshold Levenstein's distance of 10% of the longest of the imported and the modified subfield value, which is rather conservative. When it comes to the distribution of source agency in these changes, records imported from Bokbasen stand for 12% of the source records to which those changes were applied, whereas it only stands for 7% of the source records in this part of the material (Table 8(b)). The over-representation of Bokbasen in modification statistics may be explained by the agency being a newcomer as a cataloging source, and therefore less aligned with the libraries as a community of practice.

Example: Modification to Target Audience

One area in which we observed modifications is the target audience of the cataloged work, (019 $*a*). In this part of the material we identified 1,310 confident occurrences of such a change (real number probably higher), of which 1,100 were changes to existing code, 97 addition, and 113 removals. The largest libraries (larger than 10 FTEs) make more such changes per record than the smaller libraries (see Table 9). Though small numbers in absolute terms, this may be seen as an example of library staff adapting the catalog for the user group as known to the library, facilitating improved retrieval. Most of the modified codes used by the libraries appear to be valid.

Classification codes

Libraries occasionally make changes to the classification portion of the records they import from a central agency.

Classification codes are found both in the record's designated DDC field (082), as well as in the shelf mark (090). The shelf mark assigned by the cataloging agencies normally follows the *Abridged Dewey Decimal Classification*, 15th edition. Here the libraries do not have any instructions as to what changes may be made, so many variants are in use.

Table 10. Examples of abridgements of accessory table code.

From		To	
796.3340835	Soccer-Young people twelve to twenty	796.334	Fotball
305.90691409481	Refugees (Norway)	305.906	Refugees

Table 11. Examples of changes in classification.

(a) Nordic history	
example: 1945-1969	
DDC23 948.1043	DDC5 948.1054
(b) Travel manuals	
example: travel manuals	
Krakow 914.386204	Poland 914.3804

In 2015, Norway implemented a complete translation of DDC, substituting the adapted Norwegian edition used prior to this. As a consequence, a deeper classification has been applied, and changes were made to adapt to the American edition, e.g., regarding the division into periods of Nordic history. Libraries subsequently needed to adapt to changes in the depth of classification, and different shelving. This reflects in almost all the observed changes libraries make in the classification. After ignoring changes that are probably caused by changes to the agency's code not yet captured by the library, only eight of the libraries in our sample had made more than 50 changes to the imported records. Seventy percent of these were pure abridgements of Dewey (see Table 10), the remaining 30% were changes into a DDC5-number (Table 11(a)) or other changes adapting the number to the libraries' own practice.

Some libraries have kept the DDC 5 class number for Nordic history. This is a domain about which libraries have relatively many manifestations, and reclassifying must have been a practical challenge, because here there is not only a change, but also a displacement.

Another example where libraries chose to keep the DDC 5- number is travel manuals (see Table 11(b)).

In conclusion, libraries mostly endorse the choice the agencies have taken in field 082. When it comes to the shelf mark (090), a further abridgement to the original number is often made.

Import vs. in-house production/editing

The degree and the nature of in-house record production in libraries is ill-defined, and counts are encumbered by some uncertainty. Apart from import from BIBBI and Bokbasen with small local changes, practice ranges

Table 12. Import vs. In House practice. The right-most column presents the share of books of the in-house records.

FTE	# libraries in group	Total	Imported	In-house prod./ editing	Own production % of total	Book % of in-house
0–2	15	25,817	25,243	564	2.2	43
2–5	12	21,695	21,301	375	1.7	34
5–10	10	23,913	22,511	1,396	5.8	23
10–20	8	25,463	23,228	2,235	8.8	34
>20	4	19,151	15,710	3,441	18	69

Table 13. Fields occurring in more than 20% of the library resp. agency records. The emphasized columns (marked %C) compare the subsets of records describing manifestations common to the agencies. Local fields omitted.

Library	(all)		Library (IH)		BIBBI			Bokbasen		
Description	Tag	%	Tag	%	Tag	%C	%	Tag	%C	%
Title Statement	245	100	245	100	245	**100**	100	245	**100**	100
Publication data	260	97	260	94	260	**100**	100	260	**100**	100
Acquisition Nr.	025	(18)						025	**100**	100
Cataloging Source	040	(13)						040	*__100__	*100
Physical descr.	300	95	300	76	300	**100**	100	300	**99**	97
ISBN	020	91	020	48	020	**100**	91	020	**100**	100
Main-entry pers.	100	88	100	51	100	**89**	83	100	**89**	89
Summary	520	79	520	21	520	**72**	56	520	**99**	98
Index Term genre	655	70	655	30	655	**66**	71	655	**68**	69
Topical term	650	61	650	31	650	**75**	69	650	**87**	85
Added-entry pers.	700	52	700	32	700	**58**	51	700	**63**	62
Uniform title	240	33			240	**35**	*28	240	**35**	32
Classification code	082	33	082	51	082	**45**	38	082	**48**	45
Original title	574	32			574	**34**	*27	574	**35**	31
General note	500	32	500	20	500	**(18)**	(15)	500	**25**	23
Subj. a.e. Geogr.	651	27			651	**26**	26	651	**43**	44
Accessibility note	532	27								
Target audience	521	26			521	**26**	37	521	**25**	21
System & file detail	538	24			538	**(6)**	(13)			
Content note	505	21			505	**36**	31	041	**40**	35
Series Statement	440	20			440	**28**	32	440	**25**	27
Language note	546	(16)			546	**11**	*53			

between copy-cataloging and editing of records from external sources, such as Library of Congress, WorldCat, and Bibsys, and creating records from scratch. The sample here relies on the last (40th) position of the 008 control field having neither the value "2" nor "9", as BIBBI and Bokbasen brand records produced by them. We see that these latter records (labeled "IH" in Tables 13 and 14) are different from the former in the distribution of field occurrences, but the effort attributable to library staff in their preparation is uncertain. Having said that, the figures in Table 12 indicate that larger libraries do significantly more in-house metadata production than the smaller libraries, whereas among smaller size classes the differences are less significant.

Field usage (agencies and libraries)

Inspired by work done by Smith-Yoshimura et al.[41] on OCLC records, field-occurrence counts have been performed. We compared the agency

Table 14. Fields occurring in more than 10% and less than 20% of the library records resp. agency records. Local fields omitted.

Library (all)		Library (**IH**)		BIBBI		Bokbasen	
Tag	%	Tag	%	Tag	%	Tag	%
025	18	740	15	503	18	532	16
546	16	041	15	500	15	511	11
040	13	546	15	250	13		
090	12	505	15	740	12		
041	11	022	14				
740	10	090	14				
		511	14				
		651	13				
		503	13				
		521	12				
		599	11				
		240	11				
		440	11				
		025	10				

records (see Tables 13 and 14), studying overall usage pattern, and some specifics. For each record, the field is counted if at least one $a subfield is present. These numbers characterize cataloging practices by the sheer number of distinct fields occurring in more than 20% (10%) of the material. Looking at agency records, one observes a slightly richer usage of subject headings on BIBBI's side, as well as a higher usage of specific note fields. The usage pattern of the 025 field also differs. Bokbasen uses it to store (duplicate) the ISBN, whereas BIBBI uses it almost exclusively to store European Article Numbers (EAN) for non-book material. Additionally, we observed a smaller usage of the general note field and a slightly higher usage of the specific note fields. The in-house records exhibit a lower number of fields occurring in more than 20% of the records, but, for example, a higher percentage of records with a DDC classification (082), indicating a higher percentage of nonfiction materials cataloged by libraries (see also Table 12). We also observed that the In House produced (IH) library records are somewhat poorer on average as measured by the diversity of fields.

One peculiarity is field 538, which occurs in 24% of all library records but in very few agency records, indicating the popularity of electronic resources delivered by Biblioteksentralen (accompanied by BIBBI records).

Summary

In this section, the varying practices from the communities of practice represented in this paper are presented and analyzed. We have seen three main sources for different practice:

- Varying size, man-power, and mandate between libraries and agencies, probably responsible for differences of field usage between IH and imported records.
- Different "fitness for use" ideals and adaptation of records to serve local patrons. This explains modifications done to the 019 $*a* (target audience) subfield, and the enrichment of local history and culture-related records (further discussed in the "Local History and Culture" Section).
- Different interpretations of the standards, consistent with lack of "common understanding, or consensus,"[42] exemplified by the different implementations of added entries by the agencies.

Further aspects and examples

Reuse and interoperability

Reuse in a metadata context can be regarded both in terms of reducing labor, and in terms of avoiding redundancy, both locally and globally. Principle 4.3 of IFLA's Statement of Cataloguing Principles[43] states "[..] all efforts should be made to provide open access to published mappings between the standard used and the *International Standard Bibliographic Description* (ISBD), to foster better interoperability and accurate reuse of information."

The term interoperability is used in many domains (often referred to as a capability of a system), but particularly in the sense of exchanging information. It is also defined as "the effective exchange of content between systems."[44]

Regarding redundancy, Smith-Yoshimura et al. encourage catalogers to "Consider how best to take advantage of linked data and avoid creating the same redundant metadata in individual records. Consider sources outside the traditional library environment."[45] Tillett, goes even further, advocating the augmentation of records by "information from a rights management organization, a publisher or manufacturer or distributor, further enhanced by a cataloger to provide a classification and/or subject terms to help find that resource, while others may add more content information. or expand the relationships, and maybe even later the data can be enhanced by a scholar with some special expertise or knowledge about the resource or the topic it covers."[46]

In this context it is interesting to observe the Bokbasen records, and their extensive use of external authority codes for persons ($0 subfield of 100 and 700).

The labor-saving aspect of reuse is manifested in acquisition of centrally produced records from agencies, and in the copy-cataloging of records from other sources. Copy cataloging is a clear case of reuse of previously

invested labor, but seems to represent a problem, as traces of the original record irrelevant to the intended new record occasionally remain after copying. As hypothesized by the head of Bibliotek-Systemer (email correspondence), these may be parts of the record less easily visible to the cataloger (position-based values of control fields are typical), that are left unchanged, creating inconsistencies. Bibliotek-Systemer also hosts and maintains a database "Dugnadsbasen", where libraries can reuse each other's record production, updating and improving records that are not provided by the central agencies.

The case of the multilingual library

The multilingual library (Det Flerspråklige Bibliotek [DFB]), formerly a department within the Oslo Public Library, and since 2017 a department in the National Library, is an interesting in between case of being both a library and an agency. Whereas BIBBI and Bokbasen AS are commercial actors, DFB operates directly under a public mandate, committed to supplying libraries with foreign language media and associated records. Media are distributed to the libraries either through interlibrary loan (ILL) or as depots, where book packages are lent to, and are kept by the customer library for a limited period. In this period, the books are a part of the library's own collection, and the associated records are a part of the library's catalog.

One example reported by Lazzarini[47] on the side of DFB, is, for books originally written in a foreign alphabet, to allow a repetition of field 245 (originally non-repeatable in NORMARC), to facilitate search in the original alphabet rather than the transcribed form. Table 15 shows an example record describing a Persian book. This feature required a modification of the ILS program code that DFB had to commission. This can be seen as an example of a willingness to bend the standard to facilitate fitness for a certain purpose, which is consistent with the "lack of consensus" aspect of boundary object theory.

Local history and culture

As stated in the *IFLA Public Library Service Guidelines*,[48] public libraries have a responsibility for "providing a focus for cultural and artistic development in the community and helping to shape and support the cultural identity of the community." *Lokalsamlingen* (the local collection) is a traditional Norwegian way to partially fulfill this role by facilitating access to other materials relevant to the local history and culture. The following two example records from our material show the attitude toward local history and culture.

In Table 16, we juxtapose a part of a library record (left) against the parallel part of the BIBBI agency record (right). We see that the subject

Table 15. A DFB record with a repeating 245 field.

*000 c
*0010603344
*008171205 a 01ara
*015 $a10787705$bBibliofilID
*015 $bDFB
*019 ab,bubl
*020 $a978-614-442-230-4$cNKr 175,-
*082 $a344.0412$zh$25
*0823 $223/nor
*090 $cARA b 344.041$dADA
*10010$aAḍāilah, Nawwār$332268500
*24510$aAmir al-aṭibā'$bsīrat: Abū Bakr al-Rāzī$cNawwār Aḍāilah ; [illustr. av] Widād Arshīd
*24510$aأمير الأطباء$c9$aنوار عضايلةara
*260 $aBeirut$bAsala$c2017
*300 $a27 s.
*500 $aTittelen betyr: Emiren av legene

headings are modified. The 610 field is truncated to leave only subfield $*a*, whereas as many as six subject heading fields (650) are added. For the Stavanger Library, this is an important local history document, an anthology about the 1981 disaster of the Alexander Kielland marine oil platform, and it is conceivable that the library staff wishes to facilitate for retrieval by allowing a richer set of search entries. Table 17 shows an excerpt of a record describing a *bygdebok*, best translated as ”rural chronicles”, which is a documentation of a Norwegian rural community, juxtaposed with the relevant part of its source record from the Bokbasen agency (right). The library record to the left has three content notes, a personal name added entry, and a main series added entry.

Conclusion and further research

In this paper we have gauged the metadata landscape of Norwegian public libraries 2017–2018, considering the NORMARC format as a boundary object to which libraries, cataloging agencies and users relate. To this end

Table 16. The Alexander Kielland Oil Platform, library (left) vs. agency record.

```
610 ^1 $a Alexander L. Kielland (plattform)
650 ^^ $a Arbeidsulykker
650 ^^ $a Nordsjøulykker
650 ^^ $a Oljeplattformer
650 ^^ $a Plattformhavari
650 ^^ $a Pårørende
650 ^^ $a Ulykker
651 ^^ $a Norge
700 ^0 $a Myhre, Aslak Sira
       $d 1973-
       $e medarb.
       $j n.
700 ^0 $a Paulsen, Tord
       $e red.
       $j n.
700 ^0 $a Smith-Solbakken, Marie
       $d 1959-
       $e red.
       $j n.
740 00 $a Alexander Kielland-ulykken
740 00 $a Alexander L. Kielland-ulykken
```

```
610 ^1 $a Alexander L. Kielland
       $q oljeplattform
       $1 363.119622338190916336
       $2 BS
       $9 nob
       $0 ulykker og
          sikkerhetstiltak
700 ^0 $a Paulsen, Tord F.
       $e red.
       $j n.
700 ^0 $a Smith-Solbakken, Marie
       $d 1959-
       $e red.
       $j n.
740 00 $a Alexander L.
          Kielland-ulykken
740 00 $a Alexander Kielland-ulykken
```

Table 17. The Fjell Rural Chronicles, library (left) and agency.

```
505 ^^ $a Bibliografi: s. 690-703
505 ^^ $a Namneregister: s. 705-717
505 ^^ $a Notar: s. 642-704
650 ^^ $a fjell
       $q kommune
       $x bygdebøker
700 ^0 $a Skurtveit, Halvor
       $d 1972-
       $j n.
760 ^^ $t Fjell bygdebok
       $w 117128
```

```
500 ^^ $a Over tittelen: Mellom havet og byen
520 ^^ $a Mellom havet og byen, det sjette
          bandet i serien Fjell bygdebok, [...]
650 ^0 $a Lokalhistorie
       $2 Bokbasen AS
       $9 nob
650 ^0 $a Lokalhistorie
       $2 Bokbasen AS
       $9 nno
651 ^^ $a Fjell
       $2 Bokbasen AS
       $9 nob
```

we selected 49 libraries of different sizes, for which we downloaded all the records satisfying dc.date = 2017. We also downloaded the cataloging agency records using the same parameter.

The cross section chosen for the analysis, spanning materials published in 2017, thereby cataloged during 2017 and 2018, is interesting as it is immediately after the introduction of Bokbasen AS as a record supplier to the National Library for distribution to public libraries, along with Biblioteksentralen's BIBBI database to which many libraries still subscribe.

We have found some slight differences in practice between the two agencies. Particularly the treatment of added entries is an example of following a standard without necessarily agreeing on the details. Also, the number of changes made to Bokbasen records is relatively higher than its share of the imported records, indicating a "newcomer effect."

When it comes to the consumer libraries, the sample that we downloaded shows that libraries make few changes to the records that they import from central agencies, and the larger the library, the higher the change rate. The changes appear to have a local user profile, facilitating exposure of materials to target groups as the libraries know them, rather than adopting the agency's (or publisher's) assignment. We have also

detected changes seemingly done to better facilitate access to local history / local culture materials. One path of further research could be to investigate the user benefits of these local changes. A linked data or graph-based data model offers the potential of combining central and local catalog data, potentially enhancing these benefits.

We have also observed erroneous records, with a potential of reducing metadata quality. One example is control-fields inconsistent with the rest of the record. We suspect this is due to a copy cataloging error, or creation of a record based on an old record.

The above aspects imply that, particularly in larger libraries, cataloging training is still needed.

In accordance with our hypothesis, our analysis indicates that there is a greater element of "following standards" in the cataloging agencies, and a greater willingness among public library staff to facilitate for their patrons.

Apart from classification code assignment, we have not found space in this paper to do an in-depth analysis of subject headings, or other intellectual-labor-intensive metadata fields, but this is highly desirable to do in a follow-up to this project, as it is this type of metadata that represents a benefit of the library catalog as opposed to Google, for example.

Further research should also address the use of catalogs by library patrons and other users, to trace these effects in the search process.

Acknowledgments

The authors thank Biblioteksentralen, Bokbasen AS, Bibliotek-Systemer AS and also the librarians in various libraries for help and timely responses to our enquiries during the writing of this paper.

ORCID

Michael Preminger http://orcid.org/0000-0003-3524-4374

Notes

1. Roy Tennant, "The New Cataloger," *Library Journal* 131, no. 7 (April 15, 2006): 32.
2. Jane M. Davis, "A Survey of Cataloging Education: Are Library Schools Listening?" *Cataloging & Classification Quarterly* 46, no. 2 (2008): 182–200.
3. Tor Arne Dahl, Unni Knutsen, and Kim Tallerås, "Mellom tradisjonen og weben: katalogis- ering, metadata og bibliotekarutdanning" [in nob], in *Krysspeilinger: Perspektiver på bibliotek- og informasjonsvitenskap* (Oslo: ABM-media, 2011).
4. Ingebjørg Rype and Unni Knutsen, "Bruk av Dewey i norske bibliotek" [in nob], *Bok og bibliotek* 79, no. 4 (2012): 36–7.
5. Nasjonalbiblioteket / National Library of Norway, *Biblioteksøk* [in nob], accessed December 14, 2019, https://kunnskapsbase.bibliotekutvikling.no/ressurser/biblioteksok/.

6. The Norwegian NORMARC resembles MARC 21 and has local "dialects." Field description will be provided where necessary.
7. Jane Greenberg and W. Davenport Robertson, "Semantic Web Construction: An Inquiry of Authors' Views on Collaborative Metadata Generation," in *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities*, 2002 (Firenze: Firenze University Press, 2002), 45–52.
8. R. John Robertson, "Metadata Quality: Implications for Library and Information Science Professionals," *Library Review* 54, no. 5 (2005): 296.
9. Ibid., 297.
10. Ibid.
11. T.R. Bruce and D.I. Hillmann, "The Continuum of Metadata Quality: Defining, Expressing, Exploiting," in *Metadata in Practice*, eds. Diane I. Hillmann and Elaine L. Westbrooks (Chicago: American Library Association, 2004), 243.
12. Ibid., 244.
13. Leif Andresen, "After MARC - what then?," chap. 5 in *MARC and Metadata - METS, MODS, and MARCXML : Current and Future Implications*, ed. Bradford Lee Eden (Oxford: Emerald Publishing Limited, 2004), 40–51.
14. Roy Tennant, "Library Catalogs: The Wrong Solution. (Digital Libraries)," *Library Journal* 128, no. 3 (February 15, 2003).
15. Roy Tennant, "MARC Must Die. (Digital Libraries)," *Library Journal* 127, no. 17 (October 15, 2002).
16. R Guenther, "MARC: Not dead yet," *Library Journal* 128, no. 1 (2003): 12–12, issn: 0363-0277.
17. Karen Smith-Yoshimura, Catherine Argus, Timothy J. Dickey, Chew Chiat Naun, Lisa Rowlinson de Oritz, and Hugh Taylor, *Implications of MARC Tag Usage on Library Metadata Practices* (Dublin, Ohio: OCLC Research, 2010), https://www. oclc.org/content/dam/research/publications/library/2010/2010-06.pdf.
18. Ibid., 8.
19. Ibid.
20. Ibid., 27.
21. Birger Hjørland, "Is Classification Necessary after Google?," in *20 Años del Capítulo Español de ISKO: Actas del X Congreso ISKO Capitulo España (Ferrol;. 2011),* eds., María de los Ángeles Pérez Pais María del Carmen and González Bonome (España: Universidade da Coruña, 2012), 19–30.
22. Ibid., 25.
23. Smith-Yoshimura, "Implications of MARC Tag Usage on Library Metadata Practices."
24. OCLC Research, "Ground Truthing the Use of MARC," accessed December 14, 2019, http: //roytennant.com/proto/groundtruthing/.
25. Michael Gorman, "Why Teach Cataloguing and Classification?," *Cataloging & Classification Quarterly* 34, nos. 1–2 (2002): 1–13, https://doi.org/10.1300/J104v34n01_01.
26. Dahl, Knutsen, and Tallerås, "Mellom tradisjonen og weben: katalogisering, metadata og bibliotekarutdanning."
27. Geoffrey C Bowker and Susan Leigh Star, *Sorting Things Out: Classification and its Consequences* (Cambridge, Massachusetts: MIT Press, 1999).
28. Hanne Albrechtsen and Elin K. Jacob, "The Dynamics of Classification Systems as Boundary Objects for Cooperation in the Electronic Library," *Library Trends* 47, no. 2 (1998): 293–312.
29. Ingebjørg Rype, "Dewey i norske folkebibliotek" [in nob] (Master's thesis, Oslo Metropolitan University, 2014).

30. Ingbert Floyd, "Multiple interpretations: Implications of FRBR as a Boundary Object," *Proceedings of the American Society for Information Science and Technology* 46, no. 1 (2009): 1–8, https://doi.org/10.1002/meet.2009.14504603110
31. Ibid., 4.
32. F. Harvey and N. Chrisman, "Boundary objects and the social construction of GIS technology," *Environment and Planning A* 30, no. 9 (1998): 1683–1694.
33. Richard Y. Wang and Diane M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems* 12, no. 4 (1996): 5–33.
34. Beate Bartsch et al., "Kulturstatistikk 2017" [in nob], *Statistiske analyser* 160 (December 2018), https://www.ssb.no/kultur-og-fritid/artikler-og-publikasjoner/_attachment/372015?_ts=167a1eae4a8.
35. Biblioteksentralen SA, "Metadata" [in nob], accessed December 15, 2019, https://www. bibsent.no/metadata/metadata-til-bibliotek.
36. https://www.bokbasen.no/
37. Prior to the agreement between the National Library and Bokbasen, the delivery was put out for tender where both BS and Bokbasen participated. The minimum requirement the National Library placed on records contributed to a more similar application of NORMARC between the agencies.
38. Fride Fosseng, "Tilleggsprodukt metadata til bibliotek" [in nob], March 2019, accessed December 15, 2019, https://support.bokbasen.no/hc/no/articles/115001697434-Tilleggsprodukt-metadata-til-bibliotek.
39. As measured by FTE-data for libraries taken from the Norwegian Bureau of Statistics public library statistics for 2016
40. Joint Steering Committee for Development of RDA, "General Guidelines on Recording Relationships between Works, Expressions, Manifestations, and Items, Section 24," April 2017, accessed December 15, 2019, http://access.rdatoolkit.org/rdachp24.html.
41. Smith-Yoshimura, "Implications of MARC Tag Usage on Library Metadata Practices."
42. Harvey and Chrisman, "Boundary objects and the social construction of GIS technology."
43. IFLA Cataloguing Section and IFLA Meetings of Experts on an International Cataloguing Code, "Statement of International Cataloguing Principles (ICP), 2016 edition with minor revisions, 2017" (Den Haag, IFLA, 2016), https://www.ifla.org/files/assets/cataloguing/icp/icp_2016-en.pdf.
44. Jenn Riley, *Understanding Metadata: What is Metadata, and What is it for?* (Baltimore, MD: National Information Standards Organization, 2017), 6, https://groups.niso.org/apps/group_public/download.php/17446/Understanding+Metadata.pdf.
45. Smith-Yoshimura, "Implications of MARC Tag Usage on Library Metadata Practices," 14.
46. Barbara Tillett, "RDA and the Semantic Web, Linked Data Environment," *JLIS.it* 4, no. 1 (2013): 140, https://www.jlis.it/article/view/6303
47. Eleonora Lazzarini, "Flerspråklig litteratur i norske folkebibliotek: metadataproduksjon og gjenfinning" [in nob] (B.S. Thesis, Oslo Metropolitan University, 2019).
48. Christie Koontz and Barbara Gubbin, *IFLA Public Library Service Guidelines*, 2nd ed., IFLA Publications 147 (Berlin, De Gruyter Saur, 2010), Chapter 1.3.5.