

Metadatakvalitet i konsortiet

Unni Knutsen (UBO)

Johan A. S. Jørgensen (NMH)

Anne Munkebyaune (UNIT)

Ingebjørg Rype (NB)

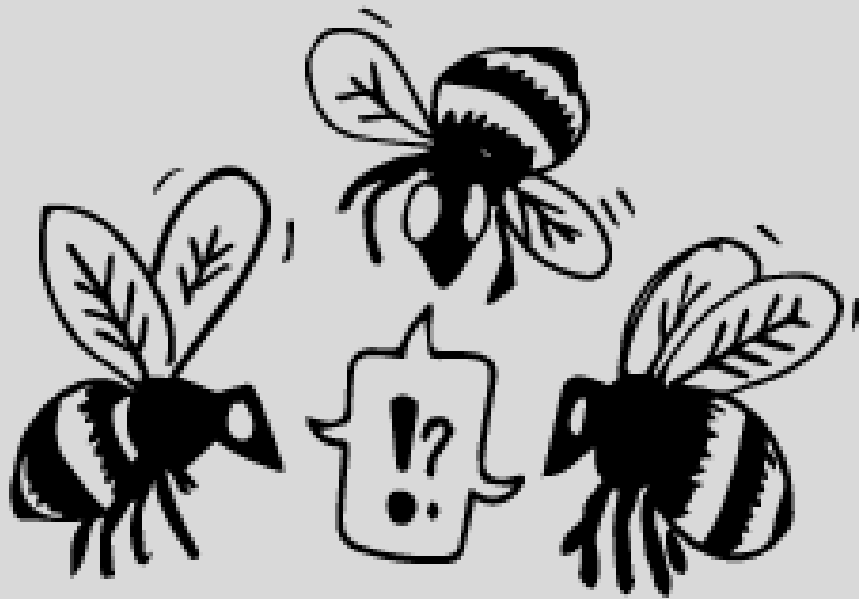
Rekkefølge

Sesjon 1

- Fellesklassifisering (Unni Knutsen og Ingebjørg Rype)
- Emneordspraksis (Unni Knutsen)

Sesjon 2

- Datakvalitet (Anne Munkebyaune)
- RDA og datakvalitet (Johan A. S. Jørgensen)
- Katalogspråk og skriftsystem (Unni Knutsen)



Kort tilbakeblikk - Dewey

- Norsk WebDewey (og Alma) ble lansert i august 2015
- Før lansering utarbeidet Norsk komité for emnedata (da NKKI) et sett med anbefalinger
- Verktøykassa for katalogisering reflekterer anbefalingene og lanserer også prinsippet om fellesklassifisering
- Det er imidlertid også mulighet for å ikke delta i fellesklassifiseringen

Fellesklassifikasjon – slik virker det i praksis

- Klassifikasjonen er felles, ikke ett biblioteks eiendom i likhet med den øvrige bibliografiske posten
- Klassenumrene dubleres ikke
- Første bibliotek som klassifiserer gjør dette på vegne av hele konsortiet
- Bibliotekene som klassifiserer må følge de retningslinjene som er skissert for registreringspraksis
- Hvis en mener at et klassenummer som er gitt av et annet bibliotek har for lav spesifisitet eller er feil, skal vedkommende bibliotek kontaktes. Hvis dialogen medfører endring i klassifikasjonen, skal det sendes melding til alle bibliotek som har beholdning på posten

Fellesklassifisering - ikke bare Dewey

- Gjelder alle felles felt:

060 NLM

080 UDC/UDK

Eksempel på hvordan dette er formulert i verktøykassa: Hvis du klassifiserer etter NLM, gjør du det på vegne av alle som bruker NLM i konsortiet

Til diskusjonen: Hvordan opplever dere som bruker disse systemene at dette fungerer?

Fellesklassifikasjon – Ikke bare Dewey

- 084 Annen klassifikasjon

For å kunne bruke dette feltet må klassifikasjonsskjemaet være registrert som gyldig klassifikasjonssystem i [LCs Classification Schema Source Codes](#).

Eksempler: loovs (Løøvs klassifikasjon, noterlyd (Norges musikkhøgskole)

- Klassifikasjonssystemer i 084 kan være brukt av flere

For de som ikke deltar i fellesklassifisering

- Bruk lokale felt:

092 Lokal Dewey

Verktøykassa: Dette er et lokalt Dewey-felt for hver enkelt institusjon. Det vil ikke vises for eller søkes opp av andre institusjoner i Alma. Også bibliotek som deltar i felleksklassifisering kan legge numre i 092 når de avviker fra retningslinjene

097 Lokal klassifisering (klassifiseringssystemer som ikke finnes på lista til LC ([LCs Classification Schema Source Codes](#)))

Dewey - fellesklassifisering

- Feltene 082 og 083 (alternativt Deweynummer)
- Alle klassenumre som legges i 082 og 083 skal være korrekt og fullstendig *og i tråd med Dewey-teori. De skal også gjelde hele dokumentet (ikke deler)*
- 082 og 083 skal alltid inneholde delfelt \$q (organisasjonskode).
Hensikten er at vi kan kontakte biblioteket som har klassifisert om vi er uenig i klassifiseringen. De som deltar i fellesklassifiseringen må få kode fra LC og bli registrert i MARC Code List for Organizations
- 082 og 083 skal også inneholde informasjon om hvilken utgave det er klassifisert etter (dels indikatorer, dels \$2 (edition number))

Tillatte løsninger - Fellesklassifikasjon

- Det er anledning til å bruke tillatte løsninger når man klassifiserer på vegne av fellesskapet, men dette må markeres tydelig (\$mb)
- Alternativt kan en velge å legge tillatt løsning i lokal felt (092)

007	ta
008	190912s2019 no e 00 0 nob
020	__ a 9788283721461 q heftet
035	__ a (NO-OsBAS)150330190
040	__ a NO-OsBAS b nob e rda d NO-OsNB
043	__ a e-no--
082	04 a 344.48107 q NO-OsBAS 2 23/nor
083	0_ a 344.07 2 23/nor

Hva gjør vi med importposter?

- **Prinsippet om proveniens:** I BIBSYS-konsortiet gjør vi visse tilpasninger til postene (skreller bort lokale felt ved import, retter til norsk språk m.v.). Vi fjerner ikke spor etter katalogkilden eller tilpasser ut over det nødvendige

Emnedatakomiteen sier følgende om gjenbruk av klassifikasjonsnummer

Ved kopikatalogisering beholdes klassifikasjonsnummeret dersom det kommer fra en autorisert kilde ...

Autoriserte kilder er:

LC (Library of Congress)

BL (British Library) [StDuBDS]

NB (Nasjonalbiblioteket)

Andre kilder kan også vurderes som autoriserte, for eksempel bidragsytere til nasjonalbibliografien og de enkelte lands nasjonalbibliotek

Hva gjør vi med importposter?

- Dersom førsteindikator er 7 - Other edition specified in subfield \$2 og vi ikke finner spor av utgave, vil dette nummeret antageligvis slettes
- Det samme kan gjelde om vi finner dublerende Dewey (fra samme utgave) og dersom vi ikke finner kilden for klassifikasjon (\$q - institusjon)

Hvem klassifiserer de nasjonalbibliografiske postene?

- To ulike koder på de nasjonalbibliografiske postene:
 - Nasjonalbiblioteket - \$qNoNB
 - Bokbasen - \$qNoBas
- Nasjonalbiblioteket kontaktes hvis biblioteket er uenig i klassifikasjon

Hva gjør Nasjonalbiblioteket når boka er klassifisert fra før av andre?

- Hvis vi er enige i klassifikasjonen som ligger i 082 fra før, endrer vi til NBs kode
- Andre biblioteks kode beholdes i 083
- Hvis vi er uenige i klassifikasjonen, tar vi kontakt med bibliotek som har kode

082	04 a 839.82374 q NO-OsBAS 2 23/nor
083	0_ a 839.828 2 23
100	1_ a Skagen, Kaj d 1949- 0 (NO-TrBIB)90053117 4 aut
245	14 a Den forseglede ordre : b opplevd og tenkt 1949-2019

Nummeret er OK, men er ikke dypt nok. Mangler indikator og signatur. **Slettes?**

082 0 4	\$\$a 859.335 \$\$q NO-OsBAS \$\$2 23/nor
083 0	\$\$a 859 \$\$2 23

Forkortet nummer. Mangler indikator og signatur. **Slettes**

Feil nummer. Mangler indikator og signatur. **Slettes**

082	04 a 839.82174 q NO-OsBAS 2 23/nor
083	0_ a 839.828 2 23
245	00 a "Djupsnøen er en kvit forbannelse" : b om Hans Børlis lyrikk

Nasjonalbiblioteket ønsker

- At bibliotekene tar kontakt hvis de ikke er enige i hvordan NB eller Bokbasen har klassifisert
- At bibliotekene bruker koder
- At bibliotekene svarer på henvendelser
- At kontaktlisten for Dewey-spørsmål oppdateres (med koder) og at det tydeliggjøres at denne lista er for de som deltar i fellesklassifikasjonen

Spørsmål til debatt

- Hvordan kan vi best sikre et godt samarbeid rundt klassifikasjon (Dewey eller andre system) i BIBSYS-konsortiet?
- Hvordan få ut informasjon til alle?

Emneord

- Ikke MARC-felt for særegne emnesystem
- I stedet brukes indikatorer

Second Indicator

Thesaurus

- 0 - Library of Congress Subject Headings
- 1 - LC subject headings for children's literature
- 2 - Medical Subject Headings
- 3 - National Agricultural Library subject authority file
- 4 - Source not specified
- 5 - Canadian Subject Headings
- 6 - Répertoire de vedettes-matière
- 7 - Source specified in subfield \$2

Emneord

- ... og \$2 (\$2 - Source of heading or term (NR))
- Koden hentes fra [Subject Heading and Term Source Codes](#)
- Noen av vokabularene er internasjonale som MeSH, noen deles av flere, noen brukes av kun én institusjon
- Emneord fra lokale system (ikke identifiserbare gjennom \$2), legges i 690 (local subject access fields)
- Nåværende praksis: Der institusjoner samarbeider om emneordssystem, legger første bibliotek inn emneord, andre bibliotek godtar disse, eventuelt supplerer

Importposter

- Hovedregel: Emnedata som er relevante for konsortiet beholdes
- Lokale emneord (69X) slettes, likeså frie nøkkelord (653) og alle emneord som ikke har navngitt kilde (andre indikator = 4)
- Emneord med verdien 7 i andreindikator, og kilde angitt i delfelt \$2 beholdes som hovedregel
- Antagelse: I UH-sektoren med studenter fra mange ulike land og med mange ulike språkbakgrunner er det å importere emneord på ulike språk en berikelse

653 frie nøkkelord (index term uncontrolled)

- Fremdeles er det en del bibliotek som bruker frie nøkkelord

245	00 a Lærerutdanning i nordiske land c Eyvind Elstad (red.)
653	_0 a lærerutdanning a lærarutdanning a pedagogutdanning
653	_5 a Norden a Norge a Noreg a Finland a Sverige a Danmark

- Bør vi stimulere til mest mulig bruk av etablerte system?

Spørsmål til diskusjon

- Er det behov for større grad av samarbeid rundt emneord i BIBSYS-konsortiet?

Metadatakvalitet, dubletter

Dublettsjekk mot Network

- Husk å søke i Network!
- ISBN/ISSN. Ikke godta 0 treff. Søk videre med for eks. ord i tittel
- Start med vide søkekriterier.
 - Snevre inn søket når du ser at du får for mange treff. (Det kan være lurt å skumme gjennom en litt lang liste også.)
- Bruk CZ til elektroniske ressurser. Ikke registrer dem eller importer dem til NZ.
- Er dublettkontrollen mindre viktig i Alma enn den var i "BIBSYS Blå"?

Merge Records & Combine Inventory

- Fungerer dessverre ikke for tiden. Ex Libris jobber med saken og det har de gjort lenge.
- Veldig uheldig fordi dette er vårt redskap til å fjerne dubletter når de først har kommet inn.

Sletting av "tomme" poster

- Med "tomme" menes poster som ingen har Inventory på. Jeg har fått spørsmål om Unit regelmessig sletter slike poster. Vi begynte med det for et par år siden, men måtte legge det på is, bl.a. fordi NB ville la slike poster ligge i tilfelle de fikk dem tilsendt til deponering. Vi bør absolutt ta opp denne ballen igjen og slette "tomme" poster.

Import av nye bibliotek

- "Merge" med eksisterende post eller importere en ny?
- Alma har innebygde match-metoder for å sjekke om posten finnes fra før eller ikke. Slike match-metoder er aldri perfekte. Vår erfaring med Alma er at metodene enten er for strenge eller for "snille". Hvis den er for streng, slippes det inn nye poster selv om den egentlig finnes fra før. Dermed har vi laget dubletter. Hvis den er for "snill" risikerer vi å bruke (dvs merger og legger holdings på) feil post.

Importfilter for eksterne ressurser

- I profilene for de eksterne ressursene har vi filter som fjerner felter som ikke har noe i vår base å gjøre. Dette har ingen betydning for spørsmålet om dubletter, men det betyr noe for hvor mye redigeringsarbeid man får når man henter inn en post til Alma. Her er det sikkert rom for diskusjon om endringer.

Leader og 008

- LDR/17 indikerer hvilken status posten har. # (dvs. blank) betyr fullkatalogisert. "5" betyr foreløpig post. Alle verdier i MARC 21 er gyldige, og dere vil kunne finne flere statuser i Alma, men når dere selv oppretter poster, er det disse to verdiene som er aktuelle.
- LDR/6 og 7 er blitt kjempeviktige i forbindelse med importer fordi match-metoden vi bruker er basert på riktige verdier her. Og de er også viktige i forbindelse med sletting av tomme poster. Feil i disse dataene **kan** for eksempel medføre feil-matching eller sletting av poster som ikke skulle vært slettet.
- 008 bør man også være oppmerksom på. Sammen med LDR spiller den en viktig rolle i Alma.

Metadataakvalitet og overgangen til RDA

Hvor mange bruker RDA?



Metadatakvalitet i lys av RDA

- Et regelverk i forkant av format (MARC21) og visning (Oria)
- En overgangsperiode jf. Nasjonal bibliotekstrategi og lenkede data
- Hvordan vurderer vi metadatakvalitet når regelverket åpner for mer presise beskrivelser?
- Hva er kvalitet?
 - “the **standard** of something as measured against *other things of a similar kind*; the degree of excellence of something.”
 - Metadatakvalitet i lys av relasjoner mellom entiteter, ikke kun deskriptive katalogdata
- Se fremtidige muligheter og læringspotensialet

Opplæring og erfaringsutveksling

- Nasjonalbiblioteket?
- Bibliotekutdanningene?
- De fire store (BOTT)?
- UNIT (Arbeidsgruppe for resource management)?
 - Verktøykasse for katalogisering (RDA)
- Lokal opplæring
- Opplæring i mindre formelle og uformelle nettverk (basert på størrelse eller felles faginteresser).

Nasjonal bibliotekstrategi 2020-2023

Katalogiseringsregler

Katalogiseringsregler er regelverk for en formell beskrivelse av dokumenter, blant annet navn på de ansvarlige for utgivelsen, titler, tidspunkt for utgivelse, emne og lignende. Gjeldende katalogiseringsregler i Norge er Resource Description and Access (RDA) i norsk utgave.

Alle metadata som Nasjonalbiblioteket produserer, er åpent og gratis tilgjengelig. I strategiperioden skal dataene gjøres tilgjengelig som lenkede data.

Nasjonalt verksregister (NB)

- Det sentrale verksbegrepet i RDA
- Nasjonalt verksregister er under arbeid

Katalogspråk

- I følge Katalogiseringsreglene (1.0E) skal opplysninger som er hentet fra selve dokumentet, gis på det språk og det skriftsystem (om mulig) som er benyttet i dokumentet i følgende felt:
 - Tittel
 - Ansvarsangivelse
 - Utgave
 - Utgivelse, distribusjon etc.
 - Serie
- Resten av opplysningene føres opp på det språk og det skriftsystem som katalogavdelingen bruker (BIBSYS-konsortiet = norsk)

RDA – 1.4. Språk og skriftsystem

- Registrer de følgende elementene på det språket og i det skriftsystemet som er brukt i de kildene de hentes fra:
 - Tittel
 - Ansvarsangivelse
 - Utgave
 - Nummerering for løpende periodika
 - Produksjon/utgivelse/distribusjon/framstilling
 - Serie
- Legg merke til bortfallet av «om mulig»!

RDA - Transkribere?

- Å transkribere er ikke hovedanbefalingen!

Alternativ LC-PCC PS NLA PS BL PS D-A-CH LAC/BAC-BAnQ KBSP

Hvis et element fra denne listen ikke kan registreres i samme skriftsystem som er brukt i kilden, registreres elementet i transkribert form.

Valgfri tilføyelse LC-PCC PS NLA PS BL PS D-A-CH LAC/BAC-BAnQ KBSP

Registrer et element oppført i **1.4 RDA** i transkribert form i tillegg til den formen som er registrert i det skriftsystemet som brukes i kilden.

- Nasjonal anbefaling: NB/Katalogkomiteen: alternative regler kan brukes

Gjeldende BIBSYS-praksis

- Dokumenter med andre skriftsystemer registreres som tidligere i transkribert form
- Ved behov kan det også registreres originalskrift i autoritetsposter. Navneformen registreres da som se-henvisning i et 4xx-felt

Behov for avklaring i konsortiet

- Å gjengi det skriftsystemet som finnes i originalkilden er det enkleste og billigste alternativet
- Å transkribere øker kostnadene
- Å transkribere samt i tillegg å bruke originaltittel er det mest kostnadskrevenende alternativet
- Gir translitterering en merverdi eller større forståelse av dokumentet vi beskriver (for hvem?)?
- Ikke all transkribering gir entydig gjengivelse av originalskrift
- Det finnes flere transkriberingsstandarder

Spørsmål, kommentarer, diskusjon

- Hvordan sikrer vi god metadatakvalitet i fellesbasen?
- Språk og skriftsystem:
 - Originalskrift eller transkripsjon: Hva er mest hensiktsmessig for brukerne våre?
 - Er tiden inne for å revurdere om katalogspråket skal være norsk?