Fremtidig nasjonalt MARC utvekslingsformat

Utredning ved Den Norske Katalogkomité

Avlevert oktober 2004 til Arbeidspakke 1 under
programmet Norsk Digitalt Bibliotek

Oslo 2004

Den norske katalogkomité

Fremtidig nasjonalt MARC utvekslingsformat

FORORD

I brev fra NDB-Rammeverk ved prosjektleder Petter Moberg av 9. juni 2004 fikk Den norske katalogkomité (DNK) i oppdrag å utrede et fremtidig nasjonalt MARC utvekslingsformat. Vi ble bedt om å vurdere NORMARC kontra MARC 21 som nasjonalt MARC utvekslingsformat. Utredningen skulle inneholde følgende elementer:

1. Sammenligning av de to formatene
2. Vurdering av formatene opp mot hverandre
3. Vurdering av faglige og ressursmessige fordeler og ulemper ved å fortsette med et nasjonalt format i forhold til å skifte til et format med bredere internasjonal anvendelse
4. Påpekning av praktiske og ressursmessige følger som et eventuelt skifte fra NORMARC til MARC 21 vil medføre (inkl. konverteringsbehov)
5. Anbefaling av framtidig nasjonalt utvekslingsformat

For å skape konsensus om konklusjonene ble vi bedt om å involvere systemleverandørgruppen. Frist for levering av utredningen var satt til 15. oktober.

Den norske katalogkomité hadde alt en tid arbeidet med en slik utredning da bestillingen kom. Vi kjente til andre land som enten hadde gått over til MARC 21 eller i svært stor grad hadde tilpasset seg til MARC 21 med et minimum av avvik. Vi ønsket selv å foreta en vurdering av en mulig overgang til MARC 21 også for Norge. Det arbeidet vi allerede hadde nedlagt, var et godt utgangspunkt for den videre utredningen.

Katalogkomitéen og systemleverandørene har hatt to møter, den 15. juni og 31. august. Et medlem av Katalogkomitéen har fungert som sekretær.

INNLEDNING

MARC 21 er resultatet av en sammenslåing av USMARC (USA) og CAN/MARC (Canada), foretatt i 1999. Hensikten var å skape et moderne format for internasjonal bruk. Det er ikke et helt nytt format. Det er heller ikke mange endringer i forhold til USMARC. Det var meningen at Storbritannia (med sitt UKMARC) også skulle være med i utviklingsarbeidet, men de trakk seg. Imidlertid har British Library sluttet seg til senere. I juni 2004 gikk dette biblioteket over til MARC 21. British Library går inn som ansvarlig for formatet på linje med Library of Congress og the National Library of Canada.

MARC-formatet er et felles grunnformat, men har i de forskjellige land foreligget i lokale varianter. Etter hvert har samarbeide om og utveksling av bibliografiske data blitt mer og mer utbredt. Data eksporteres og gjenbrukes, og det samsøkes i flere baser og ”virtuelle bibliotek”. Flere land har gått over til MARC 21. Begrunnelsen er ofte å spare inn katalogiseringskostna-

der ved gjenbruk av poster. Tidligere var UNIMARC planlagt som et internasjonalt utvekslingsformat, men det har hatt mindre suksess.

NORMARC er basert på USMARC og har kommet i tre utgaver, siste utgave i 1999. NORMARC er offisielt et utvekslingsformat til bruk ved eksport av data, men det fungerer også som registrerings- og lagringsformat i de fleste norske biblioteksystemer. Mht. eksport og import av data vil et system spare mye konverteringsarbeide ved å bruke utvekslingsformatet også internt. Ved å lage egne varianter vil en både kunne tape data og gi/få feil data ved utveksling.

NORMARC har ikke tatt i bruk hele USMARC, men har til gjengjeld utvidet med noen felt som ikke fantes i USMARC eller som betød (og fortsatt betyr) noe annet. Det dreiet seg om felt som det var behov for i norske bibliotek, f.eks. felt som angår musikk. Vi opplever nå at norske felt/delfelt som tidligere var ledige, har kommet inn i MARC 21, men med et annet innhold. Det gjelder f.eks. delfelt $j for en forfatters nasjonalitet. Skal man unngå slike problemer, må man bruke en såkalt ekte delmengde av MARC 21; dvs. man kan la være å bruke hele formatet, men ikke legge noe til eller endre innhold.

Det er anledning til å opprette lokale felt med feltkoder XX9, X9X eller 9XX, men disse kan bare brukes til eksport mellom biblioteksystemer som er enige om bruken.

Den norske katalogkomité har foretatt en vurdering og sammenligning av MARC 21, det bibliografiske formatet, med tanke på eventuell overgang til MARC 21 eller en revisjon av NORMARC. Komitéen har konsentrert seg om hva norske bibliotek vil miste i forhold til NORMARC ved å gå over til MARC 21.

Katalogkomitéen har ikke vurdert MARC 21 autoritetsformat i denne omgang. NORMARC har ikke hatt eget autoritetsformat, og komitéen regner med at valg av format vil bli vurdert i forbindelse med opprettelse av et felles norsk autoritetsregister. Andre land har valgt å følge MARC 21 på dette området.

Bestandsformatet (Holdings) er heller ikke vurdert.

INTERNE FORMAT VERSUS UTVEKSLINGSFORMAT

Som nevnt er NORMARC først og fremst et *utvekslingsformat*, utviklet for å brukes ved eksport og import av katalogdata til/fra nasjonale og internasjonale baser. I praksis er NORMARC i stor grad også blitt brukt som *registrerings*- og *lagringsformat*, dvs. som basis for det *interne formatet* i det enkelte biblioteksystem. De interne lagringsformatene er imidlertid mer rikholdige enn utvekslingsformatet ved at de f.eks. også rommer administrative data, lenkeinformasjon og bestand.

Da NORMARC ble utviklet, ble det etter ønske fra bibliotekene innført enkelte norske spesialiteter. Særlig for folkebibliotekene er det viktig at de opplysningene som ligger i de norske ”særhetene”, blir tatt vare på i et fremtidig utvekslingsformat. Dette er et forhold vi har tatt hensyn til i vår vurdering.

Selv om meget av dagens statistikk nok tas fra de interne formatene, vil vi likevel påpeke at utvekslingsformatet må ta hensyn til de krav som stilles til statistikk, først og fremst til statistikkravene som formidles av ABM-utvikling.

SAMMENLIGNING OG VURDERING AV NORMARC OG MARC 21

En fullstendig sammenligning av formatene finnes i tabellform i vedlegg 1.

GJENNOMGÅENDE FORSKJELLER

Tidligere USMARC og nå MARC 21 har med **ISBD-skilletegn** i feltene. Dette begrunnes visstnok med at MARC 21 skal være uavhengig av katalogiseringsregler. Det skal også ha sammenheng med at amerikanske katalogsystemer (biblioteksystemer) er basert på at det skrives skilletegn, og at leverandørene er skeptiske med hensyn til det programmeringsarbeidet som en endring her ville innebære. Skriving av skilletegn burde være unødvendig siden feltkoder kan styre skilletegnene. I Norge (som i andre land) har vi tradisjon for å la systemene generere skilletegn ut fra feltkodene. Skriving av tegnene medfører en fare for feilskriving. Å skrive tegnene vil oppleves som ekstraarbeide idet det blir mer manuell registrering. I tillegg blir det kostbart for biblioteksystemene å endre dette. I enkelte tilfeller vil det være en fordel å kunne bruke skilletegn, f.eks. for å skille mellom undertitler og parallelltitler i 245 $b (MARC 21).

Det er gjennomgående færre muligheter for **sortering** (alfabetisering) i MARC 21 enn i NORMARC. Noen steder mangler de helt, andre steder mangler muligheten for å sortere på annet enn å ta vekk artikkel. Enkelte steder anbefaler MARC 21 å sløyfe artikkelen i stedet. Dette stemmer ikke med norsk praksis.

MARC 21 har flere repeterbare felt og delfelt enn NORMARC har. Dette kan i enkelte tilfeller være en fordel, men det kan også være avhengig av om systemene kan holde korrekt rekkefølge ved like feltkoder.

Enkelte steder har NORMARC data oppsplittet i flere delfelt enn MARC 21. Det gjelder f.eks. felt 020 ISBN og felt 130 standardtittel.

MARC 21 har totalt flere felt enn NORMARC. Slik har det alltid vært, og det er en konsekvens av at vi i Norge har valgt ut de feltene som er nyttige for norske bibliotek.

DE ENKELTE MARC-FELT OG GRUPPER AV FELT

000 HODE /LEADER

NORMARC har henvisninger til kapitler i Katalogiseringsregler. Dette kan gi inntrykk av at det skal kodes i henhold til hvilket kapittel som er brukt som grunnlag for beskrivelsen. Noe tilsvarende finnes ikke i MARC 21. Her finnes i stedet beskrivelser ved hver kode som sier hva som omfattes av koden.

Kodene ”g” og ”k” er ikke brukt på samme måte. I NORMARC betyr de henholdsvis ”Filmer og videogrammer (kap. 7)” og ”Grafisk materiale (kap. 8)”. I MARC 21 betyr de henholdsvis ”Projected medium” (som omfatter materialer både i kap. 7 og 8) og ”Two-dimensional non-projectable graphic” (som også kan høre hjemme i flere kapitler, bl.a. kap. 8 og 9).

07 Bibliografisk kategori
Her er det noen avvik i forhold til MARC 21. For eksempel burde vår ”p” være overflødig hvis ”a” gjøres om til ”monografisk analytt”. Det vil si at typen analytt gjøres uavhengig av om vertsdokumentet er monografisk eller periodisk. Det er også kommet en ny kode her, ”i” for integrerende ressurs, som eventuelt bør innføres i NORMARC.

007 FYSISK BESKRIVELSE AV DOKUMENTET

Mange av de kodene som ikke finnes i NORMARC, er/ble utelatt fordi de ikke er/ikke var relevante for norske forhold. Som eksempel kan nevnes at vi ikke skiller mellom cartridge og cassette.

Grafisk materiale ugjennomtrengelig for lys har mer spesifikk koding i NORMARC enn hva MARC 21 kan tilby. F.eks. mangler det kode for både kunstreproduksjon, postkort og plakat. (Men 008 Visuals (som ikke finnes i NORMARC) skiller mellom kunstoriginaler og reproduksjoner.)

MARC 21 skiller ikke mellom grammofonplate (vinylplate) og kompaktplate for lydopptak. Vi mister dermed muligheten til å utveksle koder for disse platetypene.

MARC 21 inneholder ikke mye aktuell koding som ikke allerede finnes i NORMARC. Den mest relevante koden i MARC 21 som ikke finnes i NORMARC er kode for DVD-video. Ellers er 007 blitt utvidet til også å dekke taktilt materiale (pos. 00 = f), kombidokument (00 = o), noter (00 = q), fjernanalysebilder (00 = r), tekst (00 = t).

Det er en og annen divergens mellom kodene i de forskjellige posisjonene.

008 INFORMASJONSKODER

MARC 21 har generelt noe mer koding enn NORMARC. De har på enkelte områder annen koding, f.eks. for offentlige publikasjoner. Innholdskoder har flere verdier, f.eks. for flere typer juridisk materiale. På den annen side har NORMARC en annen og mer oppdelt alderskoding av målgrupper (008/22).

Monografier
008/30 Festskrifter: MARC 21 har ikke oppdeling i personer og korporasjoner
008/33 Litterære genre: flere koder i MARC 21 og annen koding, men mangler koder for antologi og pekebok som NORMARC har.

Periodika
008/20 ISSN-senter er fjernet fra MARC 21. I stedet er det innført et nytt delfelt i 022 som kalles ”Source”. Her kan en angi en kode for ISSN-senter i henhold til en liste under ”MARC Code Lists for Relators, Sources and Description Conventions”. Denne listen synes ikke å være publisert ennå.

008/21 Type periodikum: MARC 21 opererer med flere koder. NORMARC må uansett utvides i og med nye regler for løpende ressurser.
008/31 Register: MARC 21 har ikke denne koden for publiseringsmåten.

Musikk
008/18-19 Komposisjonsform: MARC 21 har her 68 tobokstavskoder. NORMARC har tilsvarende informasjon i 652, med emneord basert på en tesaurus. (18-19 er udefinert i NORMARC.)
008/30-31 Innhold i lydopptak: noen avvik i kodene.

Filmer og videogrammer
008/23-27 Medfølgende materiale, mangler i MARC 21.

01X-04X KONTROLLNUMMER OG KODER

MARC 21 har langt flere felt for numre enn NORMARC (hhv. 30 og 15).

019 er i NORMARC brukt til informasjonskoder som ikke får plass i 008. Slik bruk finnes ikke i MARC 21. Her vil vi miste muligheten for mer spesifisert koding av brukergrupper og genre. Men dette er også en overlapping med 008/33. (Som felt som ender på 9 må dette ansees som lokalt felt.)

Standardnumre
020 i MARC 21 mangler delfeltene for innbindingsinformasjon og andre tilføyelser; disse opplysningene legges i $a sammen med nummeret. Dette er en ulempe når det er et felt som indekseres. I MARC 21 anbefales det å taste inn nummeret uten bindestreker, men i LIBRIS må strekene tastes inn. Dette vil ofte være systemavhengig.
024 mangler $g Andre tilføyelser i MARC 21.
028 Utgivernummer / edisjonsnummer skal dubleres for visning i 500 i MARC 21 og i 571 i NORMARC. 571 finnes ikke i MARC 21.

Musikkoding
047 Form of musical compostion er en utvidelse av 008/18-19, som er udefinert i NORMARC, til bruk ved flere typer komposisjoner.
048 Number of musical instruments or voices code er heller ikke brukt i NORMARC, men tilsvarende forekommer i 658.
Disse feltene sammen med 008/18-19 kunne være et alternativ til noen av musikkfeltene i 65X.

05X-08X KLASSIFIKASJON

MARC 21 har 16 felt for forskjellige ”call numbers” og klassifikasjon. NORMARC har alle disse, men 11 av dem er reservert for senere bruk. De er tatt inn fordi noen biblioteksystemer importerer data fra LC og kan ha behov for å bearbeide feltene.

Felles for flere av klassifikasjonsnumrene i NORMARC er delfeltkoder som ikke finnes i MARC 21:
$e for oppslagsterm til emneregister.
$z nummerets karakter (hovednummer etc.).
$7 del av dokumentet som klassifikasjonen gjelder.

100/600/700/800 PERSONINNFØRSLER

Første indikator: MARC 21 har ikke mulighet for å skille enkelt etternavn fra sammensatt etternavn (ingen kode for sammensatt etternavn).

Andre indikator (100/800): Ingen mulighet for sortere bort artikler eller typografiske tegn som ikke skal influere på sorteringen.

Andre indikator (600): NORMARC har mulighet for å angi type emneord (hovedemne/alternativt/analytisk emne). MARC 21 har angivelse av emneordsliste brukt (se $2).

Andre indikator (700): MARC 21 har færre muligheter for å differensiere på type biinnførsler.

$i (600/700) serienummer, opusnummer finnes ikke i MARC 21.

$j nasjonalitet: MARC 21 har nå fått inn denne koden, men med et annet innhold.

$u som sorteringsdelfelt for $t eksisterer ikke i MARC 21; delfeltkoden er tatt i bruk med et annet innhold.

$w sorteringsdelfelt for annet enn bortsortering av artikler finnes ikke i MARC 21.

$x (800) ISSN finnes ikke i MARC 21 her, men er med i 490.

$0 (600) kvalifikator finnes ikke i MARC 21.

$1 (600) klassifikasjonskode (for lenking til emneord i $a) finnes ikke i MARC 21.

$2 (600) kilde for emneinnførselen, er i MARC 21 lagt i andre indikator.

$7 (600/700) del av dokumentet som innførselen gjelder, finnes ikke i MARC 21.

$8 andre karakteristika forbundet med navn, er tatt i bruk med annet innhold i MARC 21.

110/111/610/611/710/711/810/811 KORPORASJONER OG KONFERANSER

Første indikator: Ikke definert i NORMARC, type navn i MARC 21.

Andre indikator (110/111/810/811): Ikke definert i MARC 21, derfor ingen mulighet for å sortere bort artikler eller typografiske tegn som ikke skal influere på sorteringen.

Andre indikator (610/611): NORMARC angir her type emneord (hovedemne/alternativt/ analytisk emne). MARC 21 har angivelse av emneordsliste brukt (se $2).

Andre indikator (710/711): MARC 21 har færre muligheter for å differensiere på type biinnførsler.

$i (610/710) serienummer, opusnummer finnes ikke i MARC 21.

$q forklarende tilføyelse finnes ikke i MARC 21; opplysningen må legges i $a, sammen med navnet.

$u som sorteringsdelfelt for $t eksisterer ikke i MARC 21; delfeltkoden er tatt i bruk med et annet innhold.

$w sorteringsdelfelt for annet enn bortsortering av artikler finnes ikke i MARC 21.

$x (810/811) ISSN finnes ikke i MARC 21 her, men er med i 490.

$0 (610/611) kvalifikator finnes ikke i MARC 21.

$1 (610/611) klassifikasjonskode (for lenking til emneord i $a) finnes ikke i MARC 21.

$2 (610/611) kilde for emneinnførselen, er i MARC 21 lagt i andre indikator.

$7 (610/611/710/711) del av dokumentet som innførselen gjelder, finnes ikke i MARC 21.

130/630/730/830 STANDARDTITTEL

Standardtitler kan ikke sorteres på annet enn første ord uten artikkel, dvs. det finnes bare sorteringsindikator og ingen $w i MARC 21.

Delfelt $q (forklarende tilføyelse) finnes ikke i MARC 21; slik tilføyelse skal legges i $a i parentes.

2XX TITTELFELT
MARC 21 har flere tittelfelt enn NORMARC, men det er ikke noe umiddelbart behov for å ta dem i bruk.

245 i MARC 21 mangler andre muligheter for sortering enn behandling av innledende, foranstilte artikler; dvs. andre indikator finnes, men ikke $w.

For registrering av originaltittel og parallelltittel er det store forskjeller mellom NORMARC og MARC 21:

Parallelltittel skal både i NORMARC og MARC 21 i **246**, men i MARC 21 registreres også flere andre tittelvarianter i dette feltet, som har fått betegnelsen *Varying form of title*. Men i MARC 21 skal parallelltittelen også registreres i 245 $b for visning. Dette er en redundans som virker uforståelig. At MARC 21 forutsetter skriving av ISBD-tegnene i feltene, fører også til at de brukes på en annen måte. 246 mangler sorteringsindikatorer.

I MARC 21 synes også **originaltittel** å bli behandlet som en tittelvariant som registreres i **246**, mens NORMARC plasserer originaltittelen i note med egen kode 574. Ved hjelp av forskjellige verdier i indikator 1 kan det i MARC 21 styres om det skal lages, evt. ikke lages, biinnførsel på tittelvarianten og/eller om det skal lages en 500-note (generell note) med opplysningen. I tillegg kan det være aktuelt å registrere originaltittelen som standardtittel i 130 eller 240. (LIBRIS valgte i starten en annen løsning for originaltittel, idet de opprettet et eget lokalt felt, 249. Denne løsningen vil de nå gå bort fra.)

3XX FYSISK BESKRIVELSE
Det er ikke store forskjeller her mellom NORMARC og MARC 21. Forskjellene dreier seg mest om antall felt og delfelt og at de er oftere repeterbare i MARC 21.

4XX/830 SERIER
440: NORMARC har flere delfelt som ikke finnes i MARC 21:
$c (ansvarsangivelse): MARC 21 bruker antagelig heller 490 med skråstrek (/), supplert med 830.
$w: mangler i MARC 21, som andre steder; det gis ingen mulighet for å sortere på annet enn innledende, foranstilte artikler ved hjelp av en indikator.
$y (nummerering i underserien): innført i NORMARC, ser ut til å skulle legges i repetert $a og $v i 490 i MARC 21.
830 har de samme manglene vedrørende sortering. ISSN registreres under delfeltkode $x i 490 i MARC 21, mens alle opplysninger registreres i $a – med innledende, relevante ISBD-tegn - i NORMARC.

5XX NOTEFELT
MARC 21 har langt flere spesifikke notefelt enn NORMARC. Dels gir notene muligheter for å registrere flere detaljer enn vi har hatt behov for i Norge. Dels er tidligere sammenslåtte noter blitt splittet opp. Men NORMARC har også 9 felt som ikke finnes i MARC 21, og ett

som har et annet innhold: **533**. Disse vil ved en eventuell overgang til MARC 21 kunne komme i **500** (generell note), med tap av den fleksibiliteten som ligger i å ha spesifiserte noter.

I det at MARC 21 har flere felt, ligger også at det der er innebygget flere muligheter for å vise fortekster generert fra indikatorer. Disse mulighetene er ikke utnyttet i NORMARC, som ikke har indikatorer ved 5XX-kodene og bare ett delfelt. Enkelte notefelt i MARC 21 (som 505) er også utbygget med flere delfelt slik at det er muligheter for å indeksere rett fra notefeltet uten å ha registrer bi-innførsler.

Felter som finnes i NORMARC der tilsvarende felt ikke finnes i MARC 21:
503 Utgave/Historikk
512 Art, ramme eller kunstnerisk form
531 Spesielle materialdetaljer
532 Utgivelse, distribusjon, etc.
539 Serier
571 Nummerbetegnelser
572 Varianter av hovedtittel
573 Kilde for hovedtittel
574 Originaltittel

Felt **533** har ulik bruk i de to formatene:
533 Fysisk beskrivelse i NORMARC og 533 Reproduction note i MARC 21.

Varianter av hovedtittelen (572) og originaltittel (574) vil som påpekt tidligere bli registrert i 246 i MARC 21, med generert note, men denne noten vil være en generell 500-note. En eventuell indeksering vil kunne styres av indikator 1 i 246.

640-690 EMNEFELT

640 Emneinnførsel tittel og **652** Komposisjonstype finnes ikke i MARC 21. Hvor mye **640** er benyttet i norske bibliotek er usikkert, idet en for verk med hovedinnførsel på tittel, like gjerne har kunnet bruke 630 for emneinnførsler. **652** derimot brukes av de fleste bibliotek med musikksamlinger. Et alternativ til 652 kan i MARC 21 være koding i 008/18-19 og 047.

658 i NORMARC og MARC 21 har forskjellig innhold: I NORMARC brukes denne til besetning i musikk, i MARC 21 er den forbeholdt pensumrelaterte skole/undervisningsbøker. Informasjonen som legges i 658 i NORMARC, registreres i kodet form i **048** i MARC 21 (uten at vi har undersøkt om det er fullstendig sammenfall).

650, **653**, **655**, **656** i MARC 21 mangler de delfeltkodene som brukes av Biblioteksentralen og folkebibliotekene:
$0 kvalifikator
$1 klassifikasjonskode
$7 del av dokumentet som emneinnførslene gjelder
$9 språk eller språkform

LENKEFELT

Lenkefeltene **76X-78X** er så å si identiske i MARC 21 og NORMARC. Det er usikkert hvor mye de brukes i norske systemer, og anbefalingene i formatene er nokså åpne. Lenking som er basert på å hente data fra tillenket post, kompliserer utveksling av data.

84X-88X BESTAND

NORMARC har bare to felt her: **850** Eiende institusjon og beholdningsinformasjon og **856** Elektronisk lokalisering og tilgang.

MARC 21 har et eget format for bestand, *Holdings.* **850** er derfor mindre utbygget i det bibliografiske formatet, med bare to delfelt. Holdingsformatet er mer omfattende og gir plass til f.eks. registrering av tilstand for de enkelte eksemplarer.

856 er nesten sammenfallende, men NORMARCs $g URN (Uniform Resource Name) er i MARC 21 omfattet av $u Uniform Resource Identifier, som er brukt som fellesbetegnelse for URN og URL (Uniform Resource Locator). URN og URL har ulik oppbygging (syntaks) og kan derfor med fordel registreres i samme delfelt. Innholdsmessig må 856 i NORMARC og MARC 21 betraktes som likeverdige.

LOKALE FELT

MARC 21 har lokale felt tilsvarende NORMARC for oppstilling og klassifikasjon (09X), emneord (69X) og noter (59X), men ikke for lokale koder og lenker.

Da MARC 21 ble utviklet, la man som prinsipp til grunn at alle felt med 9-tall (XX9, X9X og 9XX, unntatt 490) skulle kunne brukes som lokale felt.

009 LOKALE KODER

Disse er i NORMARC reservert for lokalt bruk i de enkelte systemer. Feltet er ikke definert i MARC 21.

Se også 019 over.

9XX HENVISNINGER

NORMARCs egne felt for henvisninger (**9XX**) kan også defineres som lokale felt. Disse finnes ikke i MARC 21 da man der har et eget autoritetsformat. I Norge brukes 9XX av Nasjonalbibliografien, og av Biblioteksentralen for å få eksportert henvisninger. (LIBRIS har nå separat autoritetsregister, men kan likevel tilby eksport av poster med henvisninger lagret i de bibliografiske postene som 9XX-felt.)

OPPSUMMERING AV MANGLER OG ULEMPER I MARC 21 I FORHOLD TIL NORMARC

MANGLER

- Ikke eget (emne-)felt for komposisjonstype og besetning (musikk)
- Færre muligheter for sortering
- Ikke delfelt for personers nasjonalitet
- Ikke kobling mellom klassifikasjon og emneord

- Emneord mangler flere delfelt som brukes i Norge
- Ikke skille mellom grammofonplater og CD-plater
- Mindre spesifikke koder for grafisk materiale
- Ikke koding av sammensatt etternavn

ANDRE ULEMPER

- ISBD-tegn må genereres ved eksport av poster i MARC 21-format eller tolkes ved import i dette formatet
- Dårlig løsning for originaltittel
- Ikke eget delfelt for forklarende tilføyelse til korporasjonsnavn eller standardtittel, eller for det rene nummeret i ISBN
- Annen aldersinndeling av brukergrupper (008, 019)

HVA GJØRES I ANDRE LAND

I andre land som vi kjenner til, har man ikke primært innført MARC 21 som eksportformat. Her er det enten nasjonalbibliotek eller store bibliotek eller biblioteksystemer som tar formatet i bruk som internt format, gjerne ved overgang til et nytt system som benytter MARC 21.

Sverige: LIBRIS har gått over til MARC 21 i forbindelse med nytt system, men har laget noen lokale felt, for bl.a. originaltittel og emneord. (Lovlige lokale felt med 9-tall.) MARC 21 brukes nå av alle LIBRIS-bibliotekene, inkl. Kungliga biblioteket. Folkebibliotek som får data fra Bibliotekstjänst, bruker BTJ-MARC. Dette formatet er basert på USMARC fra 10 år tilbake. Andre bibliotek bruker de formater som systemleverandørene tilbyr. For ytterligere informasjon henvises det til nettstedet http://www.libris.kb.se/tjanster/katalogisering.jsp.

Finland: I forbindelse med konvertering til Voyager-systemet i Linnea2-prosjektet var det behov for å legge FINMARC nærmere MARC 21. Men man ville ikke gå over for fullt. Den finske versjonen fikk navnet MARC 21-Fin. De viktigste motforestillingene gjaldt felt 248 i FINMARC for deltitler (flerbindsverk?), og ISBD-tegn i postene. FINMARC brukes fremdeles i folkebibliotekene.

Nasjonalbiblioteket og universitetsbibliotek i Linnea satser sterkt på gjenbruk av poster, og bruker UseMARCON for konvertering fra WorldCat og BookWhere.

Danmark har vurdert å gå over til MARC 21, men bestemte seg for å fortsette med sitt DanMARC2. De danske katalogiseringsreglene er nylig revidert og tett knyttet til det nye DanMARC2. Disse to vurderes som mer effektive og bedre tilpasset online-kataloger enn MARC 21, slik at en overgang ville være et tilbakeskritt. Av andre argumenter de anfører, er skilletegnene, behandling av flerbindsverk og felt for musikk tungtveiende. For ytterligere informasjon henvises det til nettstedet
http://www.bs.dk/index.ihtml?side=http://www.bs.dk/nyhed2.ihtml?id=3320.

Storbritannia: British Library gikk i juni 2004 over til MARC 21 i forbindelse med at de tok i bruk sitt nye system ILS (Integrated Library System). De leverer i tillegg data i UKMARC i tre år fremover. For ytterligere informasjon henvises det til nettstedet
http://www.bl.uk/services/bibliographic/marc21move.html.

MARC OG FORHOLDET TIL ANDRE STANDARDER

En MARC-post består av tre elementer: (1) postens struktur, (2) innholdskodingen og (3) innholdsopplysningene i posten (postens innholdsmessige data).

Den fysiske kodingen i MARC-posten er basert på ISO 2709, standarden som angir det vanligste kodeformatet for internasjonal utveksling. Dette gjelder både MARC 21 og NORMARC, men for å kode MARC 21 er det også utviklet et autorisert XML-skjema. Dette har fått betegnelsen MARC XML og er strukturelt kompatibelt med ISO 2709 slik at konvertering kan skje begge veier uten tap av informasjon. Muligheter for konvertering til XML er viktig i dag, men MARC XML er et så pass allment skjema at det vil kunne brukes også på NORMARC, forutsatt at NORMARC-spesifikasjonene utvides noe. NorZIG har allerede sett på dette. Utviklingen av MARC XML vil således ikke være noen avgjørende faktor ved vurderingen av en mulig overgang til MARC 21.

Innholdskodingen (dvs. kodene og den vedtatte måten å angi og karakterisere opplysningene i en post på og senere kunne manipulere dem) er fastsatt av MARC-formatene, mens *innholdet i de dataelementene* som utgjør en MARC-post, er fastsatt av standarder som ikke inngår i formatene, men i Norges tilfelle av *Katalogiseringsregler*, den norske utgaven av *Anglo-American Cataloguing Rules*, second edition. Riktignok er innholdet i visse kodede dataelementer (f.eks. hodet, feltkodene 007 og 008) definert i MARC-formatet.

Med MARC 21 følger flere verktøy, noen av dem gratis. MARC XML er ett av verktøyene, andre er verktøy for konvertering til MODS, Dublin Core og ONIX.

Som en del av NORMARC-formatet er det definert et eget såkalt linjeformat til bruk ved utveksling av NORMARC-poster. I hvert fall BIBSYS bruker denne standarden til koding av BIBSYS-MARC-poster i søkegrensesnittet BibSøk Nett. NORMARC linjeformat brukes også ved eksport/import i Mikromarc. Noe tilsvarende format finnes ikke for MARC 21.

Z39.50 støtter både NORMARC og MARC 21.

FAGLIGE OG RESSURSMESSIGE FORDELER OG ULEMPER VED Å FORTSETTE MED NASJONALT FORMAT ELLER SKIFTE TIL INTERNASJONALT FORMAT

Frem til nå har NORMARC vært brukt som utvekslingsformat og i stor grad også som registrerings- og lagringsformat i norske bibliotek. Enkelte biblioteksystemer har benyttet ”offisielt” NORMARC-format, andre har tatt i bruk lokale varianter av NORMARC for registrering av dokumenter.

Ved en mulig overgang til MARC 21, som synes å bli et format med bred internasjonal anvendelse, kan det tenkes at biblioteksystemene velger å bruke utvekslingsformatet også som registreringsformat. Hvis systemene derimot velger å beholde sine egne registrerings- og lag-

ringsformater, eller lage lokale varianter av MARC 21, må det foretas en konvertering til og fra MARC 21 ved eksport og import.

Diskusjonen videre vil derfor i stor grad dreie seg om fordeler og ulemper ved en eventuell overgang fra NORMARC til MARC 21 som *utvekslingsformat*. En annen sak er i hvilken grad utvekslingsformatet vil påvirke registrerings- og lagringsformatet i det enkelte system.

FAGLIGE FORDELER OG ULEMPER

UTVEKSLING AV DATA (EKSPORT OG IMPORT)

Bruk av et felles internasjonalt utvekslingsformat gjør utveksling av data mellom bibliotek-systemer og samsøk mellom baser langt enklere, og kvaliteten på dataene beholdes i større grad. Det foregår mer og mer gjenbruk av bibliografiske poster mellom systemene og over landegrensene, og felles formater gir mindre ekstra opprettingsarbeid. Når det må konverteres mellom egne formater og fremmede formater, vil som regel noen data gå tapt.

NORSKE TILPASNINGER I NORMARC

Ved en overgang til MARC 21 vil vi miste en del felt og delfelt som bare finnes i NOR-MARC.

I MARC 21 finnes en rekke nye **koder** for materialtyper som ikke har vært dekket tidligere, f.eks. for taktilt materiale, DVD-video og for forskjellige typer løpende ressurser. På den annen side skilles det ikke i kodingen mellom vinylplate og kompaktplate, et skille som derimot finnes i NORMARC. Å miste denne koden vil oppleves som en ulempe. Vi vil trenge en del av de nye kodene MARC 21 byr på, men disse vil med letthet kunne innføres i NORMARC. Mangelen på disse kodene vil således ikke være avgjørende for en mulig overgang til MARC 21.

NORMARC har hatt en rekke særnorske trekk, f.eks. en spesifikk **alderskoding**, en koding av utvalgte **litterære genre**, en registrering av **nasjonalitet** på den enkelte person som har fått en hoved- eller bi-innførsel (en tilleggsopplysning man ikke finner grunnlag for i katalogiseringsreglene). Slike særtrekk vil man måtte gi slipp på ved en overgang til MARC 21. Spesielt folkebibliotekene vil se på dette som et stort tap. MARC 21 tilbyr riktignok muligheter for koding av litterære genre, men genrene er til dels noen andre enn dem vi har funnet interessante å trekke frem i Norge.

NORMARC har også innført noen **emnefelt** for musikk som er unike for den norske versjonen av MARC-formatet. Dette gjelder emneinnførsel for komposisjonstype (652) og for besetning (658). Begge feltene er tvilsomme som emnefelt siden det ikke dreier seg om en teoretisk fremstilling av komposisjonstype eller besetning, men er et uttrykk for hvilken komposisjonstype musikkverket representerer, og hvilken besetning som er brukt i musikkverket. I MARC 21 (og det tidligere USMARC) finnes det to feltkoder der det er mulig å *kode* henholdsvis komposisjonstype (008, pos. 18-19, supplert med 047) og besetning (048). Dette er en kunnskapsorganisatorisk mer korrekt måte å angi opplysningene på. Her har MARC 21 et klart fortrinn fremfor NORMARC, men det skulle ikke være noe i veien for å innføre de to feltkodene i NORMARC og avskaffe 652 og 658, hvis de faglig sett dekker det samme.

Norsk foreligger i to **språkformer**. Dette innebærer at enkelte emneord og korporasjonsnavn vil ha en bokmålsform og en nynorsk form. I NORMARC er det innført en delfeltkode ($9) der en kan angi språk eller språkform på emneordet/ordningsordet. Denne muligheten er viktig, men finnes ikke i MARC 21.

Ved overgang til MARC 21 vil en miste en del **sorteringsmuligheter** for titler og for navn når det gjelder alfabetiske utskrifter. F.eks. er det ikke mulig å sortere spesialtegn av typen þ, ð, æ. Dette må evt. løses systemteknisk. Sortering av titler som starter med artikkel i en navn-tittel-bi-innførsel er også falt ut. Selv om etablering av en database er det viktigste, vil det fortsatt være aktuelt med utskrifter til f.eks. trykte bibliografier. Dette betyr at sorteringsmuligheter ikke kan unnværes.

NORMARC har en rekke koder for spesifikke typer av **noter** i beskrivelsen. MARC 21 har fortsatt noen spesifikke koder, men til dels andre. Samtidig brukes 500 (koden for generell note) i større grad. Dette må anses som et tap. Med spesifikke koder vil en i det enkelte biblioteksystem ha mulighet for å gjøre bestemte noter søkbare, og det enkelte biblioteksystem vil oppnå større fleksibilitet ved visning. Ved bruk av 500 vil en ikke kunne styre dette. På den annen side utnytter MARC 21 mulighetene for å styre utskrifter og søkbarhet ved å bruke indikatorer og delfelt.

MARC 21 har også et format for **autoritetsregister**. Et tilsvarende format inngår ikke i NORMARC. Formatet for autoritetsregister bør kunne tas i bruk uavhengig av om vi tar i bruk MARC 21 for registrering og/eller utveksling.

TEGNSETT

I NORMARC brukes 8-bits ASCII tegnsett eller utvidet 7-bits ASCII tegnsett. MARC 21 gir muligheter for å velge mellom to tegnsettstandarder, MARC-8 (den vanligste) og UTF-8 (UNICODE). UTF-8 er vanlig tegnsett i XML, samtidig som det er et lovlig tegnsett i MARC 21. UTF-8 (UNICODE) har et stort potensial som tegnsettstandard for all tekst, og utvalget av verktøy som støtter UTF-8, er økende. UTF-8 gir større muligheter for flerspråklige kataloger idet UTF-8 kan representere direkte langt flere alfabeter enn hva som er tilfelle med andre tegnsett som er i bruk. Bruk av UTF-8 (UNICODE) er derfor en stor fordel.

Som nevnt tidligere har vi to offisielle språkformer i Norge, nynorsk og bokmål, men vi har også to offisielle språk, norsk og samisk. Samisk har en rekke spesialtegn som skal komme frem i katalogen. I landet har vi også en rekke minoritetsgrupper med behov for spesialkataloger på sine språk. Valg av standard for tegnsettet i MARC-posten er av den grunn viktig.

Skal MARC 21 tas i bruk i Norge, må det være en betingelse at UTF-8 benyttes som tegnsettstandard. Noe annet vil være lite fremtidsrettet. Beholdes NORMARC, må likeledes UTF-8 tillates. Overgang til UTF-8 vil neppe medføre de store utgiftene, men innføringen kan få store konsekvenser i systemene som skal håndtere kodene.

NYE LØSNINGER / FRBR

Implementering av FRBR (Functional Requirements for Bibliographic Records) vil med stor sannsynlighet føre til nye katalogiseringsregler og behov for nytt format. Her vet man kanskje mer om 5 år slik at meget taler for at man venter med å ta i bruk et nytt format og heller fortsetter med NORMARC.

På den annen side fungerer MARC 21 som de facto internasjonal utvekslingsstandard, og mye av utviklingen vil kanskje først komme i tilknytning til dette formatet. Ved å bruke MARC 21 vil en få en del programvare og lignende "gratis" med på kjøpet, f.eks. MARC XML.

RESSURSMESSIGE FORDELER OG ULEMPER

VEDLIKEHOLD

Velger man å innføre MARC 21, vil vi slippe vedlikehold av formatet forutsatt at man tar i bruk den engelskspråklige utgaven. Til gjengjeld blir det nødvendig med en løpende oppfølging av endringer i MARC 21. NORMARC foreligger i dag i sin tredje versjon og må vedlikeholdes i takt med utviklingen internasjonalt. Det nedlegges noen arbeidstimer pr år på dette arbeidet. Omfanget har dog vært relativt begrenset etter at første offisielle utgave forelå.

Det må imidlertid vurderes om ikke verket skal oversettes. Spesielt hvis det skal brukes som registreringsformat, vil brukerne ønske en norskspråklig utgave. Dette vil medføre utgifter til oversettelse, og i tillegg vil vi få oppgaven med å vedlikeholde den norskspråklige utgaven parallelt med at det skjer endringer i den engelske originalen.

Fordelen med oversettelse vil bl.a. være at vi da kan plukke ut en delmengde av MARC 21. I tillegg kan vi komme med norske eksempler. Med fullstendig engelsk utgave vil vi få et meget omfattende (og uoversiktlig) format med diverse koder som vi ikke trenger.

SKILLETEGN

MARC 21 (og tidligere USMARC) forutsetter bruk av ISBD-tegn med tilhørende mellomrom i tillegg til at det brukes delfeltkoder. En slik praksis gjør det mulig å repetere delfeltkoder eller å gi en delfeltkode forskjellig mening. I norske biblioteksystemer er NORMARC programmert slik at delfeltkoden automatisk gir korrekt ISBD-tegn med tilhørende mellomrom. Forutsetningen har vært en entydig bruk av f.eks. $b i 245, at $b bare ble brukt i forbindelse med annen tittelinformasjon. Årsaken til denne forskjellen er forskjellig praksis i registreringen og presentasjon av poster i MARC 21 og NORMARC. Så vidt vi vet, er det bare i 240 og 245 at MARC 21 har et faktisk behov for ISBD-tegn i MARC-posten.

Slike ISBD-tegn vil være nyttige bare når postene skal presenteres i et ISBD-lignende format. I andre tilfelle vil de utgjøre støy som må filtreres bort. Generelt mener vi at ISBD-tegn i MARC-poster burde være unødvendig. MARC-formatet burde lages slik at disse var overflødige. Vi kan ikke se stort annet enn ulemper ved en slik praksis.

PRAKTISKE OG RESSURSMESSIGE FØLGER VED ET SKIFTE TIL MARC 21

Innledningsvis har vi presisert at vi ikke kunne unngå å ta NORMARC som registrerings- og lagringsformat i betraktning, selv om dette primært er en utredning av utvekslingsformat.

En overgang fra NORMARC til MARC 21 som utvekslingsformat er mindre problematisk enn også å skulle ta MARC 21 i bruk som registrerings- og lagringsformat. Hvis systemene skal ta MARC 21 i bruk som registrerings- og lagringsformat (internt format), må atskillig mer arbeid påregnes.

Uansett må en regne med en del arbeid med å lage konverteringstabeller mellom de respektive interne formatene og utvekslingsformatet. De som har brukt utvekslingsformatet som også som internt format, har hittil vært spart for slike konverteringstabeller. Dette gjelder for de fleste systemene i Norge.

Det er grunn til å tro at en eventuell overgang til MARC 21 som internt format vil bli vurdert i de enkelte systemene, siden fordelene ved å bruke samme format både internt og til utveksling er store, men her vil nok kostnadene bli nøye vurdert. Det må påregnes en lang overgangsperiode.

Konverteringsarbeidet betegnes som komplisert fordi ikke alle MARC-koder er entydige. En må regne med at opplysninger vil gå tapt både ved overgang fra NORMARC til MARC 21 og ved MARC 21 til NORMARC da det er vanskelig å lage entydige rutiner for konvertering.

Selv om NORMARC blir brukt for registrering og utveksling også i fremtiden, vil det fortsatt være behov for konvertering ved import av utenlandske poster levert i MARC 21-format. Den norske katalogkomité vil derfor anbefale at det sentralt - uansett vedtak om fortsatt bruk av NORMARC eller om overgang til MARC 21 - utvikles konverteringstabeller og programvare for konvertering til bruk for de forskjellige biblioteksystemene. Dette vil lette tilpasningen lokalt for systemleverandørene. Det vil være økonomisk uforsvarlig å overlate et slikt arbeid til hver enkelt systemleverandør. Det foreslås at det settes av midler innenfor programmet Norsk digitalt bibliotek til dette utviklingsarbeidet.

OPPSUMMERING

- Ut fra tekniske forhold kan overgang til MARC 21 som utvekslingsformat være en fordel
- Faglig sett er det mer å tape enn å vinne ved å gå over fra NORMARC til MARC 21
- Innføring av MARC 21 vil sannsynligvis gi enklere utveksling av katalogdata
- Velges komplett engelsk utgave av MARC 21, vil vi slippe vedlikehold av formatet
- Foretrekkes en norsk oversettelse av MARC 21 (gjerne delmengde), vil det medføre utgifter til oversettelse og til vedlikehold
- Overgang til MARC 21 vil kreve bruk av ISBD-skilletegn.
- Overgang til nytt format vil gi store konverteringsutgifter; det er kostbart å skifte format
- Både NORMARC og MARC 21 er tilpasset dagens katalogiseringsregler
- Implementering av FRBR vil sannsynligvis føre til nye katalogiseringsregler og behov for nytt format, og dette taler for at man venter med å endre format
- MARC XML og UTF-8 (UNICODE) innføres også for NORMARC hvis NORMARC beholdes
- MARC 21 anbefales vurdert for bruk til nasjonalt autoritetsregister

MULIGE LØSNINGER

Ut fra sammenligningene og diskusjonene ovenfor synes det å foreligge fire mulige anbefalinger:

1. NORMARC beholdes som utvekslingsformat, oppdateres og vedlikeholdes fortløpende
2. MARC 21 innføres som utvekslingsformat, forutsatt at det utvikles et norsk tillegg slik at norske "særegenheter" (behov) kan inkorporeres i lokale felt (kfr. 9-felt reservert for lokale behov). NORMARC oppdateres og vedlikeholdes som grunnlag for registrerings- og lagringsformat
3. MARC 21 velges som utvekslingsformat og tas i bruk som det foreligger (fullstendig, engelskspråklig utgave)
4. MARC 21 velges som utvekslingsformat, og formatet – eller en delmengde av dette – oversettes

Ved å gå over til MARC 21 som utvekslingsformat får vi automatisk tilgang til en rekke verktøy (konverterings- og registreringsverktøy) og programvare som vi har bruk for. Vi vil også komme mer i harmoni med resten av Europa og det angloamerikanske området.

Å bruke MARC 21 uten bearbeidelser eller tilføyelser kan ikke anbefales. For mange sentrale opplysninger vil gå tapt ved import/eksport.

Utvikles det et norsk tillegg basert på bruk av lokale felt for norske spesialopplysninger, vil MARC 21 som utvekslingsformat ivareta norske biblioteks behov for utveksling innen Norge. Vi vil få et utvekslingsformat som fungerer både nasjonalt og internasjonalt, noe som vil være ressursbesparende. Det er et mål at det norske tillegget får et minst mulig omfang.

Komitéen regner med at man beholder dagens lagrings- og registreringsformater. En oversettelse av MARC 21 vil derfor ikke være nødvendig, idet utvekslingsformatet først og fremst vil berøre systemleverandørene og deres arbeid.

Det vil fortsatt være behov for å holde NORMARC ved like. Oppdateringene/justeringene her vil være av mindre art.

KONKLUSJON

Den norske katalogkomité vil i samråd med systemleverandørene anbefale at MARC 21 med norske tilpasninger tas i bruk som utvekslingsformat. Hvis norske tilpasninger ikke lar seg gjennomføre, anbefales det at NORMARC beholdes som utvekslingsformat. NORMARC må i så fall oppdateres.

Det anbefales at det sentralt - uansett vedtak om fortsatt bruk av NORMARC eller om overgang til MARC 21 - utvikles konverteringstabeller og programvare for konvertering til bruk for de forskjellige biblioteksystemene. Det foreslås at det settes av midler innenfor programmet Norsk digitalt bibliotek til dette utviklingsarbeidet.

Det bør tas initiativ til etablering av en europeisk MARC-gruppe som kan være med og påvirke den videre utviklingen av MARC 21.

Videre anbefales det at MARC XML med UNICODE UTF-8 som tegnsett brukes for fysisk koding av data. Dette gjelder uansett valg av MARC 21 eller NORMARC.

Det anbefales at Den norske katalogkomité får ansvar for å vedlikeholde den norske delen av utvekslingsformatet og for konverteringstabellen.