

National strategi for autoritetsdata - Bilag

Bilagsfortegnelse

Bilag 1: Et nationalt repository for autoritetsdata

Bilag 2: Autoritetsdata – Datatyper og RDA-sammenhænge

Bilag 3: Brugsscenarier

Bilag 4: Autoritetsdata i dag – status

Bilag 5: Oversigt over væsentlige attributter til forskellige autoritetsdatatyper

Bilag 6: Ordliste og definitioner

Bilag 1: Et nationalt repositorium for autoritetsdata

Rapport fra arbejdsgruppe 1

Indholdsoversigt

| | |
|---|----|
| Indledning | 1 |
| Rammer og funktionelle krav til det nationale autoritetsdatarepositorium | 3 |
| Konceptuel datamodel | 6 |
| Use cases for afprøvning af datamodel | 7 |
| Udformning af URI'erne i repositoret, herunder persistente og unikke ID-numre | 12 |
| Relationer til andre identifikatorer for personer (ISNI, ORCID mv.) | 14 |
| Bilag 1 | 16 |

Indledning

Bibliografisk Råd afleverede i april 2015 en foreløbig indstilling vedr. den fremtidige nationale autoritetsdatastrategi. Inden den endelige indstilling kan afgives til Slots- og Kulturstyrelsen ønsker Bibliografisk Råd, at der skal ske yderligere afklaring på en række områder. Det drejer sig om følgende:

- a. Typer af data, der kan behandles som autoritetsdata
- b. Datamodel for et nationalt autoritetsdata-repositorium
- c. Analysearbejde af identifikatorer til autoritetsdata
- d. Sammenhænge mellem RDA og autoritetsdata

Med dette formål nedsattes 2 arbejdsgrupper:

- Arbejdsgruppe 1: Til pkt. b og c
- Arbejdsgruppe 2: Til pkt. a og d

Nærværende notat udgør rapporteringen fra gruppe 1.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Gruppens opgave

Gruppens opgave, jf. dens kommissorium (Bilag 1.1), har været at

- opstille en model for det nationale autoritetsdata-repositorium. Modellen skal adressere såvel sammenhængen til eksisterende data som det fremadrettede datasamarbejde i et linked data univers.
- udarbejde modeller for autoritetsentiteter i bibliotekernes infrastruktur, på en måde, der sikrer optimal udnyttelse og genbrug af autoritetsdata i danske biblioteker.
- opstille en model for, hvorledes de danske nationalbibliografiske autoritetsdata sammen med bibliotekernes autoritetsdata kan interagere med VIAF.
- på baggrund af ovenstående modeller, komme med input til et egentligt udviklingsprojekt med henblik på etablering af et nationalt autoritetsdata-repositorium

Her udover har gruppen skullet give en vurdering af hvilke veletablerede og udbredte systemer for persistente og unikke ID'er, det vil være relevant at implementere i danske bibliotekers autoritetsdata.

Gruppens sammensætning

Anders Cato, Slots- og Kulturstyrelsen

Leif Andresen, Det Kongelige Bibliotek

Pernille Saul, Guldborgsund-bibliotekerne

Susanne Thorborg, DBC

René Høtbjerg Øhlenschläger, Københavns biblioteker

Desuden har Knut Anton Bøckman, Det Kongelige Bibliotek, deltaget i drøftelserne vedr. URI'er, datamodel og sammenhængen til DEFF-projektet "Linked, Open and Social".

Susanne Thorborg har fungeret som tovholder for gruppen og Anders Cato som gruppens sekretær

Rammer og funktionelle krav til det nationale autoritetsdatarepositorium

I dette afsnit opridses en række væsentlige forudsætninger som arbejdsgruppen mener, bør være grundlaget for bibliotekernes samarbejde om et nationalt autoritetsdatarepositorium. Endvidere listes et antal funktionelle krav, der - som de væsentligste – bør opfyldes af et nationalt autoritetsdatarepositorium.

Forudsætninger

Fælleskatalogen, dvs. Danbib, er rammen for det nationale autoritetsdatarepositorium, når der er tale om samarbejde om dataskabelse og genbrug af data. For brug af data i forbindelse med søgning og præsentation mv. er interessentkredsen meget bredere.

Det nationale autoritetsdatarepositorium etableres som et samarbejde, hvor alle deltagere i fællesskabet forpligter sig til at overholde de fælles retningslinjer (policy) der er fastlagt. (I det følgende, for så vidt det angår dataskabelse, skal "alle" forstås som "alle der har forpligtet sig til samarbejdet")

Repositoriet skal kunne rumme og tilgængeliggøre såvel nationalbibliografiske autoritetsdata som autoritetsdata fra andre deltagere i samarbejdet.

Datamodellen skal kunne understøtte en konsekvent implementering af autoritetsdata i fælleskatalogen. Således skal den bl.a. kunne håndtere ufuldstændige eller usikkert identificerede autoritetsentiteter.

Repositoriet skal kunne fungere som indgang til andre datasæt, fx til VIAF. Autoritetsentiteter skal derudover kunne linkes til andre datasæt, som er beskrivende for den pågældende entitet. Dette skal muliggøre, at man henviser til eksisterende autoritetsentiteter i stedet for selv at tilføje mange specielle oplysninger i entiteten. Et eksempel kunne være et link til en autoritetspost i GeoNames (<http://www.geonames.org/>) fra en post, der beskriver et sted.

Repositoriet skal kunne anvendes i sammenhæng med andre domæner, der arbejder med identifikation af ophav og rettighedshavere til publikationer, fx ORCID.

Modellen skal give mulighed for at berige eksterne systemer med data fra det nationale repositorium.

Mulighed for interaktion med EUROPEANA bør undersøges nærmere.

Modellen tager i sin første version udgangspunkt i personentiteter, men skal også kunne anvendes for andre entitetstyper af relevans for biblioteker.

Implementering af repositoriet skal ske på en sådan måde, at det vil være muligt at interagere i forhold til BIBFRAME, når dette format foreligger i en endelig version.

Funktionelle krav

De funktionelle krav er for overblikkets skyld delt op i de, der primært vedrører dataskabelse, de der vedrører anvendelse af data i forbindelse med søgning og præsentation af data (resource discovery) og endelig dem, der især vedrører den nationale infrastruktur og internationalt samarbejde.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

I virkelighedens verden er de forskellige områder ikke så skarpt opdelt, og nogle af kravene kan med god ret høre hjemme i flere af grupperne – ligesom et vist overlap ikke kan undgås.

Herunder følger en ikke-prioriteret liste over de funktionelle krav, der på nuværende tidspunkt er identificeret. I forbindelse med implementering af repositoret skal der ske en yderligere kvalificering og prioritering af kravene.

Dataskabelse

Der skal formuleres regler for samarbejdet, og den implementerede datamodel skal give mulighed for at kontrollere at reglerne overholdes.

Det skal være muligt for alle deltagere i samarbejdet at oprette nye autoritetsentiteter. Oprettelse skal kunne ske både som eksport fra lokale systemer og som interaktiv oprettelse i repositoret.

Den oprettende institution skal kunne tildele en URI i henhold til det fastlagte URI-skema til nye autoritetsentiteter.

URI'en er identifikator for den pågældende autoritetsentitet.

Som en del af URI'en indgår et identifikationsnummer, der skal trækkes fra en central nummerrulle. Denne nummerrulle skal kunne anvendes udenfor repositoret via en enkel web-service.

Identifikationsnummeret skal være unikt for hele repositoret og skal indeholde et checkciffer.

Der skal defineres relevante autoritetsniveauer og det skal kunne markeres hvilket niveau et udsagn om en entitet har. Der skal fastlægges roller med regler for, hvem der må hvad, på de forskellige niveauer.

Det skal være muligt at supplere eksisterende autoritetsentiteter med yderligere egenskaber efter nærmere retningslinjer, og det skal kunne identificeres, hvem der bidrager med hvad til den enkelte del af entiteten

Uanset om en beslutning om RDA som danske katalogiseringsregler kommer før eller efter implementering af det nationale autoritetsdatarepositorium bør det, af hensyn til international udveksling af data kunne angives relationer svarende til alle niveauer i RDA. (Se udybende herom i bilag 2, s. ?? vedr. sammenhænge til RDA)

Modellen skal herudover også understøtte andre nødvendige, standardiserede relationer mellem entiteter. Som standarder for angivelse af relationer kan især peges på SKOS (Simple Knowledge Organization System).

Datamodellen skal sikre, at der kan skelnes mellem metadata der beskriver selve ressourcen og metadata vedr. beskrivelsen af ressourcen, såkaldte administrative metadata. Som eksempler på administrative metadata kan nævnes status for beskrivelsen, fx i hvilken grad den er kontrolleret, beskrivelsens proveniens, beskrivelsens oprettelses- og evt. ajourføringsdato mv.

Det skal være muligt, fx med en attribut at angive en nærmere bestemmelse af en egenskab for en entitet, fx angivelse af hvilket sprog der er anvendt ved beskrivelsen af egenskaben.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Søgning og præsentation

Søgning og præsentation af data er ikke en del af repositoriets datamodel, men modellen skal sikre, at søgninger i repositoret skal kunne ske i både repositoriets data og i VIAF.

Datamodellen skal også understøtte, at det skal være muligt at præsentere/få udleveret udvalgte data, dvs. data i forskellige views. Som eksempler på relevante afgrænsninger kan nævnes nationale tilhørsforhold eller leverandøren af autoritetsdata.

Datamodellen skal desuden være åben for, at andre autoritetsdatasæt vil kunne blive inkluderet i søgning

Bibliotekernes fælles infrastruktur og det internationale samarbejde

Det skal være muligt at koble autoriteter sammen med bibliografiske data i fælleskatalogen og/eller lokale kataloger. En autoritetsentitet skal dog også kunne eksistere uden at der findes en tilknyttet bibliografisk entitet.

Datamodellen skal sikre, at entiteterne skal kunne anvendes til validering i både søgesituation og dataskabelse

Det nationale repositorium (NAR) skal kunne fungere i sammenhæng med lokale autoritetsbaser (LAR)

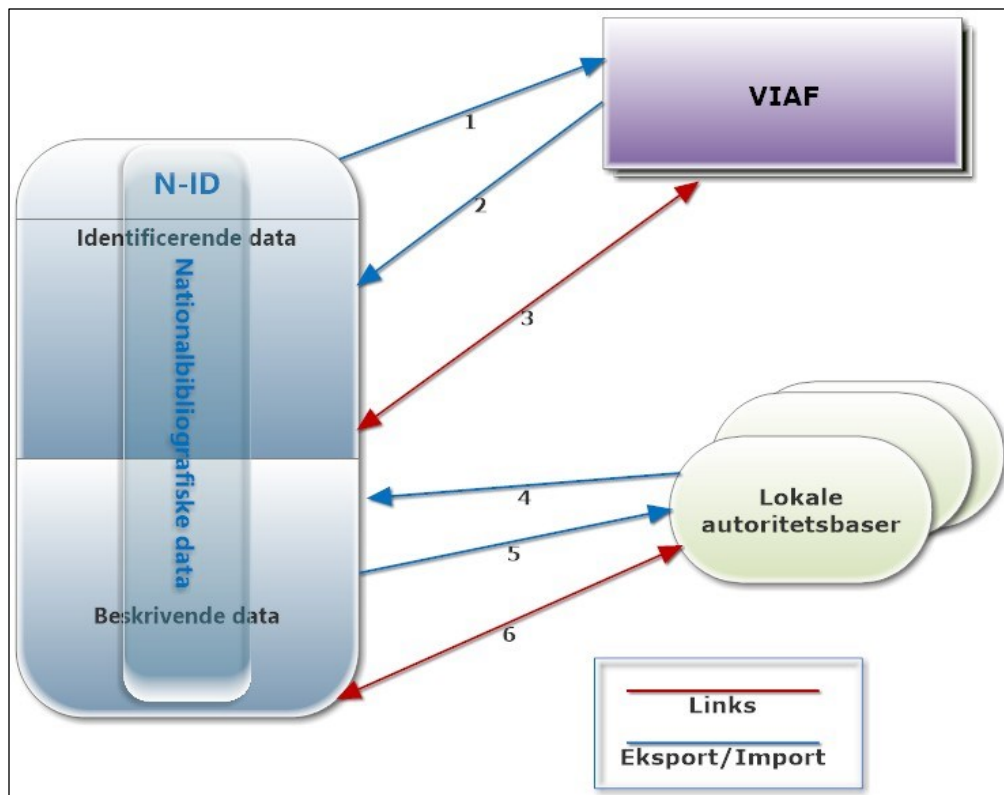
Repositoriet skal kunne tilgås af forskellige systemer (standardiserede og åbne grænseflader), således at det bl.a. skal være muligt at lave et API, der kan hente og aflevere/redigere metadata til repositoret.

Der skal kunne udleveres data fra repositoret til VIAF

Det skal være muligt at håndtere "fejl" under dataeksport og dataimport til og fra VIAF og andre systemer VIAF, fx når der forekommer 2 permalinks, der i realiteten dækker over én og samme entitet.

Konceptuel datamodel

Den konceptuelle model illustrerer helt overordnet sammenhængen mellem det nationale



autoritetsdatarepositorium, VIAF og lokale autoritetsbaser.

Fig.1: Konceptuel model

Forklaringer til fig.1:

En entitet i det nationale autoritetsdatarepositorium (NAR) indeholder 3 elementer: Et unikt ID for entiteten (N-ID), et eller flere identificerende dataelementer, fx VIAF-ID, ISNI-nummer eller ORCID. Herudover et eller flere beskrivende dataelementer, fx entitetens navn og variantformer heraf eller oplysninger om nationale tilhørsforhold.

En delmængde af entiteterne i det nationale repositorium udgøres af autoritetsdata fra Nationalbibliografien.

Flow'et mellem repositorier er angivet med pile:

- Entiteter fra det nationale repositorium uploades til VIAF (1)
- Der vil kunne importeres visse dataelementer fra VIAF til det nationale repositorium (2)

- Der oprettes links mellem det nationale repositorium og VIAF (3)
- Lokale systemer uploader entiteter til det nationale repositorium (4)
- Lokale systemer downloader entiteter fra det nationale repositorium (5)
- Hvis match, beriges den nationale entitet med link til den lokale entitet og URI returneres til den lokale base (6)

Use cases for afprøvning af datamodel

Til brug for en vurdering af den foreslåede datamodel har gruppen udpeget et mindre antal use cases, nemlig dem, som vi anser for afgørende for modellens gyldighed.

I det følgende beskrives hvordan de valgte use cases fremtræder i datamodellen.

- Oprettelse af ny fællesressource
- Oprettelse af fællesressource for entitet, der allerede findes som lokal autoritetsentitet.
- Flere biblioteker ønsker at benytte samme data om en entitet
- Tilføjelse af supplerende data til en fællesressource
- Præsentation af udvalgte dele af en ressource
- Tilføjelse relation mellem en fællesressource og en lokalressource
- Indberetning af bibliografisk post, der indeholder VIAF-URI som eneste identifikation af personnavn

Use case 1: "Oprettelse af ny fællesressource"

Der vil i denne model være tale om oprettelse af en ressourcerejistrering i autoritetsrepositoriet [På figur: Nat.Aut.Repos.]

Rettigheder:

- En brugerpolitik og identificeringsprocedure styrer, hvem der har rettigheder til at oprette ressourcerejistreringer.
- Oplysning om hvem der foretager registreringer gemmes (dataproveniensen) [administrative metadata]

Type:

- Der skal angives/vælges hvilken type ressource, der er tale om (person, organisation, sted, andet...)
- Til hver type af ressource er knyttet en profil af obligatoriske og/eller supplerende datatyper. [beskrivende autoritetsdata]
- Til hver type af ressourcer er knyttet en eller flere obligatoriske match med eksterne kilder. [aut-links]

Identifikator:

- En unik identifikator trækkes fra central nummerrulle og tildeles den nye ressource.
- En URI på ressourcen publiceres. [N-ID]

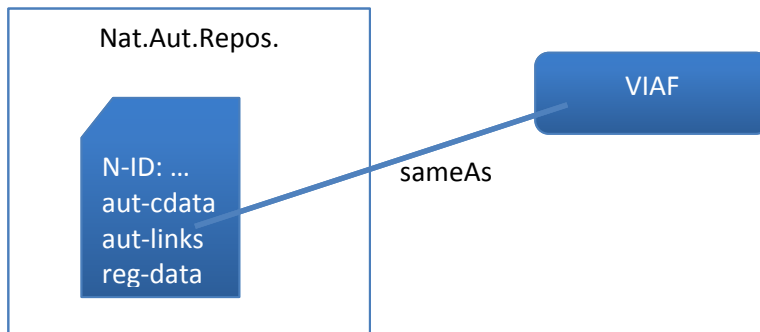
Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Kvalitetssikring:

- Obligatoriske data-typer (fx navn) for ressourcen kontrolleres
- Supplerende oplysninger kan registreres efter nærmere profil for ressourceypen
- På baggrund af indtastede/importerede oplysninger foretages match i basen for at finde potentielle dubletter, der foreslås korrektion (for at undgå dubletter)

Ekstern linkning:

- Ressourcen forsøges matchet mod eksisterende eksterne kilder iflg. profilen.
- Ved match kan ressourceregistreringen beriges med data fra ekstern kilde efter evt. forhåndsdefineret skabelon
- Ved manglende match mod relevante kilder (i nedenstående eksempel VIAF) markeres ressourcen som kandidat til eksport dertil.



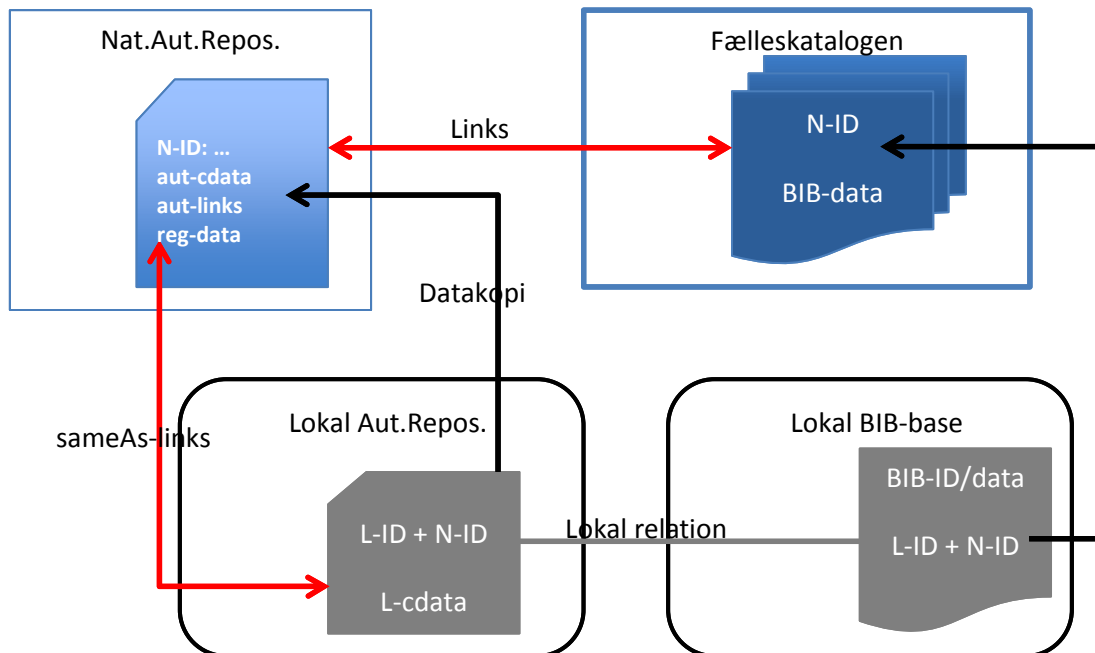
Use case 2: "Oprettelse af ny fællesressource for entitet der allerede findes som lokal autoritetsentitet"

Som i Use case 1, blot kan data fra de lokale registreringer genbruges [link til VIAF udeladt fra figur]. To overordnede genbrugsmuligheder kan skitseres – linkning eller kopiering:

Mulighed 1 (Linked Data model): Obligatoriske data-oplysninger kopieres [sort pil] til Autoritetsrepositoriets registrering, men supplerende oplysninger kan beholdes lokalt og der oprettes en sameAs-relation til denne [rød pil]. Institutioner med rettigheder til at oprette ressourceregistreringer har også ret til at registrere sameAs-relationer til åbne lokale ressourcer.

Hvis den lokale ressource ikke er åbent tilgængelig, skal der oprettes en ny repository-registrering som i Use case 1.

Mulighed 2 (uden linked data): Eksisterende registreringer fra lokale ressourcer kopieres til ny fællesressource, som i Use case 1 [sort pil].



I begge tilfælde: Lokalressourcer beriges med (link til) fællesressource [N-ID], så fællesressorens identifikator tilknyttes eksporten af bibliografiske ressourcer fra det lokale katalog til fælleskatalogen.

Det er en lokal beslutning, hvordan forholdet mellem Lokal autoritets-ID [L-ID] og den nationale autoritets-ID [N-ID] forvaltes: ved at erstatte L-ID med N-ID, ved sameAs-linkning, eller ved at kopiere N-ID til den lokale registrering.

Use case 3: "Flere biblioteker ønsker at benytte samme data om en entitet"

Da autoritetsdatarepositoriet skal indgå i en linked-data sammenhæng, hviler det grundlæggende på en "open world" forståelse, dvs. at hvem som helst kan sige, hvad som helst om hvad som helst. Repositoriet bidrager her med at levere identificerbare, offentlige ressource-registreringer, som der kan siges noget om. Der siges noget om disse ressourcer ved, at de benyttes i nationalbibliografien og fælleskatalogen, og ved at de hævdes at være *sameAs* en ressource fx i VIAF. Dette er nationalbibliografiske data i repositoriet (jf. use case 1). Andre deltagere med oprettelsesrettigheder i repositoriet kan etablere *sameAs* links *ud til* lokale autoriteter for supplerende oplysninger (jf. use case 2). Alle andre kan sige noget om repositoriets ressourcer ved at linke *ind til* dem. En resources backlinks peger dermed på de data som andre (fx biblioteker) "hænger på" autoritetsressourcen. Det vil være hensigtsmæssigt, hvis der etableres en service, så man kan filtrere disse backlinks på proveniens, fx kun biblioteker, eller bestemte biblioteker, hvis udsagn, man kunne ønske at knytte sig til.

Use case 4: "Tilføjelse af supplerende data til en fællesressource"

Deltagere, der har rettigheder til det, kan tilføje supplerende data (iflg. skabelon for pågældende ressourcestype) i registreringen i repositoret, eller som *sameAs*-relation *ud til* lokale registreringer (jf. Use case 2).

Enhver kan supplere data ved selv at publicere data om en lokalt defineret ressource og hævde, at denne er *sameAs* en ressource i repositoret. Disse supplerende datas troværdighed afhænger bl.a. af tillid til bidragsyderen.

Use case 5: " Præsentation af udvalgte dele af en ressource"

"Sammensat ressource" forstås her som en ressource hvis beskrivelse er sammensat af

- 1) nationalbibliografiske data i repositoret (obligatoriske og supplerende)
- 2) supplerende oplysninger i lokale data som der linkes *ud til* fra repositoret
- 3) udsagn i lokale data som linker *ind til* repositoret

Et bibliotek, der ønsker at præsentere data om denne ressource kan udvælge disse fra de forskellige grupper ud fra disse kriterier:

Fra 1)

- a) bestemte datatyper (fx navn, fødselsår,..) ud fra bestemte ressource typer
- b) bestemte dataleverandører (via proveniensoplysninger i repositoret)

Fra 2)

- a) bestemte datatyper ud fra bestemte ressource typer
- b) bestemte dataleverandører (via proveniensoplysninger i repositoret og/eller linkdestination)

Fra 3)

- a) bestemte dataleverandører (via backlink-holder)
- b) bestemte standardiserede datatyper (identificerbare gennem fx schema.org)

Data i gruppe 1 vil være direkte tilgængelige i repositoret; data i gruppe 2 og 3 vil være tilgængelige via links. Måske kan det overvejes at tilbyde en tjeneste med caching af data fra gruppe 2 og for udvalgte, kendte kilder i gruppe 3 for at opnå en hurtigere præsentation.

Use case 6: "Tilføj relation mellem en fællesressource og en lokalressource"

Dette foregår enten som en *sameAs*-linkning fra repositorets beskrivelse til den lokale ressource (for berettigede deltagere), eller ved at der linkes ind til repositorets ressource i og med at denne benyttes i en lokal beskrivelse.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Use case 7: "Et bibliotek tilknytter et VIAF-URI til et personnavn i en bibliografisk post, og indberetter posten til fælleskatalogen"

Systemet tjekker om den pågældende VIAF-URI er repræsenteret i NAR.

- 1) Findes VIAF-URI'en som et "sameAs" link i en eksisterende entitet knyttes biblioteket til pågældende entitet. Hvis bibliotekets navneform ikke allerede findes i NAR-entiteten tilføjes den hertil. NAR-ID tilføjes til den indberettede post i fælleskatalogen.

Systemet returnerer link mellem VIAF-URI og NAR-ID til biblioteket. Det er bibliotekets eget valg, om man vil gøre brug af denne opdatering.

- 2) Findes VIAF-URI'en ikke i NAR, opretter systemet en entitet med et nyt NAR-ID, og med bibliotekets navneform samt VIAF-URI'en som eneste indhold.

Systemet returnerer link mellem VIAF-URI og NAR-ID til biblioteket. Det er bibliotekets eget valg, om man vil gøre brug af denne opdatering.

Udformning af URI'erne i repositoret, herunder persistente og unikke ID-numre

NB! Dette afsnit omhandler kun identifikatorer af relevans for personentiteter

Grundlæggende følges anbefalinger fra [ISA: D7.1.3 - Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC](#).

Kriterier gennemgås her fra venstre mod højre, ud fra denne skabelon:

skema://domæne:port/type/kategori/reference

Skema: i denne sammenhæng vil det være protokollen http eller https – krypterede forbindelser vinder hurtig udbredelse også uden for datatrafik, der indeholder brugeres personfølsomme data. Sikre tjenester der skal benytte repositorets datatjeneste vil få en "blandet indhold" advarsel hvis denne service kører som http. Tilbud om https er dog et *implementeringsspørgsmål*

Domæne (og port): der anbefales at benytte en dedikeret service, uafhængig af publiceringsorganisation. Da denne kan skifte bør også servicen have et dedikeret domæne. For eksemplets skyld her binar.dk – præfikset med data for at indikere dataservice (til forskel fra sider om servicen); port er typisk 80, eller 443 for https, men *dette er et implementeringsspørgsmål*

http(s)://data.binar.dk

Type: her [skelnes](#) mellem en URL, der som URI identificerer et virkelig objekt, og en URL der fører til information om objektet. Der vil være en [implementeringsbeslutning](#) om der for dataset i binar.dk's størrelsesorden bør benyttes *hash-URIs* for virkelige objekter (det gør VIAF) eller *303 redirects*. Her et eksempel med *303 redirects*:

http(s)://data.binar.dk/id/{reference}

http(s)://data.binar.dk/dok/{reference}

Eller med hash:

http(s)://data.binar.dk/{reference}#{evt.klartekst}

Implementeringsbeslutning

Kategori: Under forudsætning af, at den sidste reference er unik på tværs af begrebslig kontekst, synes der ikke at være afgørende tekniske grunde til at benytte en underinddeling med kategorier, dvs. /person/, /sted/, /organisation/, etc. Kilder der kun har én slags entiteter (ISNI, ORCID) gør det naturligt nok ikke, men heller ikke fx VIAF, [GND](#), [BNF](#) eller [KB i Sverige](#). LOC gør det for deres autoritets-filer. Fordelen ved at benytte begrebsunderdelingen er primært, at det letter den menneskelige forståelse af referencen. For en SPARQL-forespørgsel vil man kunne bruge kontekst-sensitive præfikser som fx

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

@autpers <http://data.binar.dk/id/person>

@autsite <http://data.binar.dk/id/sted>

Det er dog ikke disse tekststrengene, men brug af type-beskrivelse af entiteter i RDF og i SPARQL, der vil kunne reducere load og give en bedre performance. Den menneskelige læsbarhed er imidlertid stadig et argument for brug af kategorikontekstualisering, og der vil givetvis være nyttigt med en konkordans mellem disse kategorier og de definerede RDF-typer.

Værd at overveje, hvis man ønsker angivet kategorikontekst, er hvordan man ønsker at trække grænser mellem kategorier – vil man operere med fx Agent som samlebetegnelse, eller skelne mellem Person og Organisation? Vil man have kategorier som Emner og Hændelser – hvad vil fx 'den kolde krig' være i dette tilfælde? Er 'det tredje rige' et Sted eller et Emne? The Library of Congress som også producerer klassifikation, synes i det mindste at have deres bud på det. Spørgsmålet er, om det er formålstjenligt at have det med i en URI.

Et yderligere aspekt er sprogligt valg af kategorier. I et nationalt autoritetsrepositorium vil det være naturligt at begreberne er på det lokale sprog, fx :

`http(s)://data.binar.dk/id/person`

`http(s)://data.binar.dk/id/sted`

Der vil evt. skulle tages hensyn til hvordan disse forklares for internationale brugere af servicen, se evt. [Style Guidelines for Naming and Labeling Ontologies in the Multilingual Web](#)

Samlet anbefaler gruppen, at der ikke opereres med kategori-kontekstualisering i URI'er.

Reference: Referencen vil være den tidligere beskrevne *identifikator*, og skal være unik for hele repositoret. (Også uafhængigt af kategori-kontekstualisering, hvis man alligevel ønsker at benytte dette.)

Relationer til andre identifikatorer for personer (ISNI, ORCID mv.)

Gruppen har i første omgang udelukkende set på identifikatorer for personer. Analyser vedr. mulige identifikatorer for andre autoritetstyper som fx korporationer, værker, lokaliteter mv. bør indgå i forarbejdet til implementering af disse, i den takt, det besluttes – jf. indstilling fra Gruppe 2, bilag ?, side ??.

Internationale person identifiers

De to mest oplagte identifier systemer for personer er:

- ISNI - International Standard Name Identifier

isni.org

https://en.m.wikipedia.org/wiki/International_Standard_Name_Identifier

ISNI skal ses som en generel identifier for forfattere af værker registreret af biblioteker / bibliotekssamarbejder.

- ORCID - Open Researcher and Contributor ID

<http://orcid.org/>

<https://en.m.wikipedia.org/wiki/ORCID>

ORCID dækker som udgangspunkt nulevende forfattere og forskere. Den enkelte person registrerer sig selv.

ISNI og ORCID samarbejder. ISNI er i sagens natur mere dækkende end ORCID.

I forhold til den praktiske brug af identifiers, er VIAF nok så relevant, da den tager udgangspunkt i ret mange landes nationalbibliografier m.v.

VIAF tildeler i praksis også en identifier i form af en VIAF URI. Denne bruges bl.a. i forhold til WorldCat identities. I praksis må VIAF URI være mere dækkende end ISNI

Nationale forhold

Udover DBC's autoritetsbase med personer er der en relevant ressource i form af Biblioteksafgiftens base.

Forfatternummer

Alle modtagere - og potentielle modtagere - af biblioteksafgift har et forfatternummer, som er linket til de faustnumre, der repræsenterer de bøger, som den pågældende er registreret som modtager af biblioteksafgift

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

for. Den må endvidere være indeholde modtagere af forfatterafgift, som er døde indenfor de sidste 10-20 år alt efter oprydningsspraksis.

Der er derfor her et potentiale for sammenkædning af forfatternavne og udgivelser for nyere forfattere. Alle nulevende er med.

Adgang til at bruge data fra Biblioteksafgiftens base er nok ikke helt lige til, men en model, som ikke udstiller personfølsomme data burde være mulig.

Bilag 1.1

Kommissorium for BIR-arbejdsgruppe vedr. model for nationalt autoritetsdata-repositorium mv. (arbejdsgruppe 1)

Opgaver

Gruppen har til opgave, at

- opstille en model for det nationale autoritetsdata-repositorium. Modellen skal adressere såvel sammenhængen til eksisterende data som det fremadrettede datasamarbejde i et linked data univers.
- udarbejde modeller for autoritetsentiteter i bibliotekernes infrastruktur, på en måde, der sikrer optimal udnyttelse og genbrug af autoritetsdata i danske biblioteker.
- opstille en model for, hvorledes de danske nationalbibliografiske autoritetsdata sammen med bibliotekernes autoritetsdata kan interagere med VIAF.
- på baggrund af ovenstående modeller, komme med input til et egentligt udviklingsprojekt med henblik på etablering af et nationalt autoritetsdata-repositorium

Her udover skal gruppen give en vurdering af hvilke veletablerede og udbredte systemer for persistente og unikke ID'er, det vil være relevant at implementere i danske bibliotekers autoritetsdata.

Tidsramme

Gruppen skal aflevere sin endelige indstilling til Bibliografisk Råd inden d. 30. november 2015. Milepæle undervejs fremgår af vedlagte bilag.

Sammensætning

Anders Cato, Kulturstyrelsen

Leif Andresen, Det Kongelige Bibliotek

Pernille Saul, Guldborgsund-bibliotekerne

Susanne Thorborg, DBC

René Høtbjerg Øhlenschläger, Københavns biblioteker

Desuden har Knut Anton Bøckman deltaget i drøftelserne vedr. URI'er og sammenhængen til DEFF-projektet "Linked, Open and Social".

Susanne Thorborg er formand for gruppen og Anders Cato fungerer som gruppens sekretær

Bilag 2: Autoritetsdata – Datatyper og RDA-sammenhænge

Analyse og anbefalinger fra BIR Autoritetsdataarbejdsgruppe 2

| | |
|--|----|
| Indledning..... | 2 |
| Kapitel 1 : Hvilke data har vi pt. nationalt? | 4 |
| Kapitel 2 : Overordnede RDA –regler | 8 |
| Kapitel 3 : Hvad er iflg. RDA core på områderne? | 10 |
| Kapitel 4 : Hvad vil vi udover core-elementerne anbefale som elementer i et dansk niveau på områderne og i overensstemmelse med mulighederne i RDA og hvorfor (hvilke brugerbehov er der tale om)? | 18 |
| Kapitel 5: Variant access points..... | 22 |
| Kapitel 6: Relationships between persons, families, and corporate bodies | 27 |
| Kapitel 7: Værkdata – skal vi gentage eller referere når værkdata sammensættes af elementer fra forskellige entiteter?..... | 29 |
| Kapitel 8: Vil det være muligt at genbruge internationalt tilgængelige data? | 32 |
| Kapitel 9: Samlet anbefaling..... | 35 |

Indledning

Arbejdsgruppens kommissorium og sammensætning

Arbejdsgruppen blev nedsat efteråret 2015 med følgende kommissorium og sammensætning

Kommissorium for BIR-autoritetsdata arbejdsgruppe 2 – datatyper og RDA-sammenhænge

1) Indledning

Bibliografisk Råd (BIR) har nedsat to arbejdsgrupper, som skal arbejde videre med de opgaver i den nationale strategi for autoritetsdata, som skal på plads, inden der skal laves en endelig indstilling omkring autoritetsdata

Følgende 4 opgaver skal der arbejdes med:

- e. En anbefaling til typer af data
- f. Opstille en datamodel for et nationalt autoritetsdata-repositorium
- g. Analysearbejde af identifikatorer
- h. Opstille sammenhænge til RDA

[Arbejdsgruppe 1](#) vedr. model for nationalt autoritetsdata-repositorium mv. tager sig af [pkt. b og c](#)
[Arbejdsgruppe 2](#) vedr. datatyper og RDA-sammenhænge tager sig af [pkt. a og d](#)

2) Opgaver

Gruppen har til opgave, at

- Indstille til BIR hvilke typer af autoritetsdata man anbefaler skal være de danske også med udgangspunkt i de internationalt tilgængelige data eksempelvis VIAF
- Opstille begrundelser for anbefalingerne
- Med udgangspunkt i de anbefalede autoritetsdata angive de sammenhænge der er til RDA
- Endvidere opstille formatmæssige krav og udfordringer

3) Tidsramme

Gruppen skal aflevere sin endelige indstilling til Bibliografisk Råd senest d. 1. dec. 2015.

4) Arbejdsgruppens sammensætning:

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Hanne Hørl Hansen (DBC) Gruppens Tovholder

Anders Cato (Slots- og Kulturstyrelsen) Gruppens sekretær

Henriette Fog (Det Kongelige Bibliotek)

Dorete Bøving Larsen (Statsbiblioteket)

Arbejdsgruppens arbejde

Arbejdsgruppen har taget udgangspunkt i de typer af autoritetsdata, der omtales i den nationale strategi for autoritetsdata (afsnittet: Typer af autoritetsdata).

Vi har dog i nærværende analyse og anbefaling valgt ikke at inddrage typer af autoritative data, der har karakter af indhold fra lukkede lister/vocabularer (sprog og materialetyper).

Udelukkelsen skyldes primært, at denne type data mere hensigtsmæssigt kan håndteres ved at koble sig på internationale initiativer som bl.a. RDA registry for materialetyper <http://www.rdaregistry.info/> og eventuelt tilsvarende dokumentation for sprog- og landekoder.

Ligeledes er emneområdet (emneord og klassifikation) udeladt, dog dækkes korporationer, stednavne, personer og titler som emneord af de generelle afsnit herom. Grundarbejdet på området blev ikke udført som aftalt, men da RDA pt. ikke indeholder anvisninger på emneområdet, mener vi, at dette område ikke pt. har en central betydning for samarbejdet om autoritetsdata.

Gruppens anbefalinger bygger på et miks af elementer bestående af:

Hvad skal der til i forhold til RDA – herunder

hvad er nødvendigt (core)?

hvad er derudover ønskeligt set med danske øjne ud fra bl.a. brede brugerhensyn og lokale hensyn?

Hvad kan lade sig gøre – både i forhold til

- kan noget opbygges ud fra det vi allerede har eller kan skaffe os adgang til?
- realistiske arbejdsbyrder indenfor de nuværende økonomiske rammer
- lovgivningsmæssigt (fx persondataloven)

Gruppen tager i sine anbefalinger overordnet ikke stilling til, hvordan det skal håndteres, hvis der lokalt vælges en anden vedtagen form end den nationale form. Dette har vi forudsat løst via datamodellen.

Format

Gruppen har forholdt sig til summen af den information, det er muligt at angive indenfor danMARC2's nuværende hjælpepostformat.

Da den fremtidige formattype for danske autoritetsdata endnu ikke er fastlagt og ikke nødvendigvis er MARC, nævnes kun hvilke relevante oplysninger, det ikke er muligt at anføre i det nuværende format og ikke præcis, hvordan de fremadrettet bør fremgå.

Overordnet set indeholder RDA-rapporten fra 2012 (ajourført 2015) et notat om RDA's autoritetsdata (afsnit 4.2.), hvor der allerede er peget på en række elementer, der ikke kan angives indenfor det nuværende hjælpepostformat. Kapitlet fra RDA-rapporten er medtaget som bilag 1 i denne rapport.

Kapitel 1 : Hvilke data har vi pt. nationalt?

Resume: Pt. har DBC en intern autoritetsbase over personnavne i en stor del af nationalbibliografien. Den er omfattende og vil være et glimrende udgangspunkt for etableringen af et dansk autoritetsrepositorium på person- og familienavne. Navne fra bibliotekernes mange udenlandske materialer fremgår ikke af basen. Der er ingen dansk autoritetsbase på korporationsnavne, men kontrollerede navne fremgår af nationalbibliografien. Det er yderst begrænset, hvad vi har af autoritative data på titler. Musik- og filmområdet er pt. de eneste dataleverandører og her er udformningen for størstedelens vedkommende meget forskellig fra RDA's anvisninger, fordi der har været anvendt danske standardtitler. Herudover nævnes en række problemstillinger (forskelle mellem nuværende regler/praksis og RDA), der skal adresseres ved fastlæggelsen af den danske profil.

Personnavne

Nationalbibliografien publicerer det, der svarer til indholdet af RDA's *authorized accesspoint* samt *variant accesspoint*, integreret i de bibliografiske poster. Der sker ingen koordinering mellem DBC's del af nationalbibliografien og KB's Dansk musikfortegnelse og Dansk kortfortegnelse, hvorfor fx *Tjajkovskij* ses translittereret forskelligt i forskellige dele af nationalbibliografien.

Folkebibliotekerne skal anvende samme kontrol af navneformer som DBC på det, der indberettes til Danbib og har derfor principielt navneformskontrol på linje med nationalbibliografien.

Forskningsbibliotekerne har ingen fælles praksis på området; Nogle har navneformskontrol på udvalgte områder, andre igen opererer slet ikke med navneformskontrol.

Vedtagne former og deres henvisninger fremfindes og genbruges typisk via skanning og efterfølgende fremfindning af posten i den lokale katalog eller Danbib, hvorefter felterne overføres maskinelt eller manuelt til den aktuelle katalogisering.

Internt på DBC er der etableret en autoritetsbase over personnavne brugt i Dansk bogfortegnelse, Dansk billedfortegnelse og Dansk lydfortegnelse fra 1999 ff. Pt. indgår navne i musikoptagelser således ikke. Der arbejdes pt. på en udvidelse med poster, der vil inkludere personnavne fra bibliografiske poster fra før 1999, hvor autoritetspost allerede er oprettet. Navnet i de bibliografiske poster er erstattet af et link til

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

autoritetsposten. Før ibrugtagning i 2013 er der dannet poster maskinelt og fra 2013 sker oprettelsen og ajourføring af autoritetsposter løbende i samspil med registreringen af materialerne. Når der ses bort fra formelle søgekoder mv. er indholdet i autoritetsposterne begrænset til det, der har fremgået af felt 100/700 og felt 900 i de bibliografiske poster. I de bibliografiske poster forekommer kun et link til en autoritetspost og vedtagen form og henvisninger indflettes fra autoritetsposten ved udlevering. Der arbejdes på at lægge forklarende noter om pseudonymer, der hidtil har indgået i den bibliografiske posts felt 512, ind i autoritetsposterne og samtidig slette noten i de bibliografiske poster for så at genbruge noten fra autoritetsposten ved udlevering af den bibliografiske post. Det betyder, at der skal kunne skelnes mellem noter der også bør fremgå af den bibliografiske post og noter der udelukkende skal fremgå af autoritetsposten i autoritetsposten.

Indholdet af DBC's autoritetsbase over personnavne indberettes til VIAF som danske nationalbibliografiske data og vurderes som en god basis for autoritetsdata på personnavne. En berigelse af posterne med data fra VIAF er på trapperne. Berigelsen vil bestå i, at autoritetsposterne løbende suppleres med VIAFs oplysninger om fødselsår/dødsår, nationalitet, køn, ISNI mv.

I relation til RDA skal man være opmærksom på, at navneskift for personer medfører, at samtlige poster rettes til ny form, mens de nuværende danske regler blot anviser navneskift med henvisninger, således at ældre poster ikke rettes.

RDA vil også betyde en mindre udvidelse af omfanget af personnavne, fordi der iflg. RDA også skal oprettes autoritetsposter på fiktive karakterer (fx Miss Piggy).

Personnavne, der ikke indgår i den danske nationalbibliografi, er ikke omfattet af autoritetsbasen. Mange af disse navne vil fremgå af udenlandske institutioners indberetning til VIAF. Der fastlægges ikke i VIAF én navneform pr person, men i princippet én form for hver af de institutioner, der indberetter navnet. Hvis et værk oversættes til dansk vil DBC dog oprette en autoritetspost på ophavet til værket.

Familier

Familier har pt. stort set kun været brugt som entitet i forbindelse med emnebehandling – nærmere betegnet som underdeling af klassemærker, hvor de ses formateret i felt 652 *a, der både dækker person- og korporationsnavne samt som emneord formateret som personnavn (felt 600). I danMARC's feltoversigt, felt 600 (personnavne som emne) ses 2 eksempler på slægter.

Familier indgår pt. i DBC's emneordsbase på linje med personnavne som emneord. Det er hensigten, at personnavne som emneord vil blive flyttet til autoritetsbasen for personnavne og de vil efter dette kunne vælges anvendt på linje med personnavne. Hvorvidt de nogensinde i praksis vil blive anvendt som opslag på ophav i felt 100 eller 700 er dog tvivlsomt, - de vil højst sandsynligt fortsat kun anvendes som emne.

Korporationsnavne

Nationalbibliografien publicerer det, der svarer til indholdet af RDA's *authorized accesspoint* samt *variant accesspoint* og *relationships between accesspoints*, integreret i de bibliografiske poster.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Folkebibliotekerne skal anvende samme kontrol af navneformer som DBC på det, der indberettes til Danbib, og har derfor principielt navneformskontrol på linje med nationalbibliografien.

Vedtagne former og deres henvisninger fremfindes og genbruges typisk via skanning og efterfølgende fremfindning af posten i den lokale katalog eller Danbib, hvorefter felterne overføres maskinelt eller manuelt til den aktuelle katalogisering.

Udover opslagsformen og eventuelle henvisninger er der ikke registreret yderligere data.

Stednavne kan optræde som en del af et korporationsnavn, enten som det foretrukne navn for en administrativ enhed, eller som tilføjelse til et korporationsnavn. Om stednavne, se særskilt afsnit.

Der er pt. forskel på praksis i hhv. folke- og forskningsbiblioteker på valg af navneform på internationale korporationer, der ikke har dansk som officielt sprog. Her foretrækker folkebibliotekerne og DBC en dansk form, mens forskningsbibliotekerne foretrækker en navneform fra et af de officielle sprog i følgende prioriterede rækkefølge: norsk, svensk, engelsk, tysk, fransk, spansk, italiensk, russisk, øvrige sprog. RDA lægger op til en praksis, der ligner forskningsbibliotekernes (RDA 11.2.2.5.2 og 11.2.2.5.3)

Stednavne

Stednavne kan indgå som dataelement i alle øvrige entiteter. Her er det behandlet separat.

Nationalbibliografien benytter De danske katalogiseringsregler, som forenklet siger, at for landenavne benyttes landets navn på dansk, mens der for lokale stednavne benyttes den lokale form. Hvis to områder har samme navn, adskilles de ved tilføjelse af navnet på en større geografisk enhed.

Titler (work/expression)

RDA forudsætter, at man laver Authorized Access Points. For titler med ophav består Authorized Access Point af hovedophav efterfulgt af den vedtagne form af titlen på værket/expression. Det, der pt. bedst modsvarer disse Access points, er uniforme titler og de standardtitler for musik, som folkebibliotekerne anvender. Uniforme titler bruges generelt ikke pt. i Danmark, selvom der findes regler for en del typer af værker, specificerede i kap. 4 i katalogreglerne. Kun uniforme titler på Bibelen og musik- og filmtitler vides anvendt. For musik findes der to varianter: mere AACR2-lignende uniforme titler til hvilke man bruger felt 240, og de allerede nævnte standardtitler, som følger en ældre regeltradition og som placeres i et eget felt, 239. På filmområdet anvendes udelukkende standardtitler (=danske lanceringstitler).

Nationalbibliografien publicerer således kun i meget begrænset omfang og kun på film- og musikområdet, en form for vedtagne titler, der dog ikke svarer til indholdet af RDA's *authorized accesspoints* samt *variant access points* og *relationships between access points*. I positivt fald er oplysningerne integreret i de bibliografiske poster.

Elementerne af autoritative data på titler er begrænset til det, der fremgår af indholdet i felt 239, 240 og til dels felterne 241 (originaltitel) samt felt 245*a (danske originalværker i 1. udgave). Det feltindhold, der bedst modsvarer Authorized Access Points, er helt klart 240, som bruges enten alene (anonyme værker) eller sammen med det ophav i 100/110 som værket har, - men i meget stærkt begrænset omfang.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Felterne 239 og de få danske 240-felter fungerer som vedtagne former. De og deres henvisninger fremfindes og genbruges typisk via skanning og efterfølgende fremfindning af posten i den lokale katalog eller Danbib, hvorefter felterne overføres maskinelt eller manuelt til den aktuelle katalogisering.

Udover opslagsformen og eventuelle henvisninger er der ikke registreret yderligere data.

Med RDA er kravene til at lave Authorized Access Points langt større. Alle manifestationer skal kobles til minimum et authorized accespoint for det indeholdte værk eller den indeholdte expression.

På en række værker vil kravet om værktitler kunne opfyldes ved at maskingenere autoritetsdata ud fra eksisterende bibliografiske data:

Det drejer sig om originale danske 1. udgaver, hvor værktitel i langt de fleste tilfælde svarer til manifestationens titel. (Ifølge Statistikbanken var der i 2014 knap 5700 originalt danske kommercielle førsteudgaver).

Derudover kan en række originaltitler evt. ligeledes anvendes som udgangspunkt for dannelsen af værktitler. Dette skal dog analyseres nærmere, da originaltitler typisk er gengivet fra det oversatte materiale og fx derfor ikke kan forventes translittereret i henhold til vedtagne standarder. (Ifølge Statistikbanken var der i 2014 ca. 3100 til dansk oversatte kommercielle førsteudgaver).

Generelt skal vi dog beslutte om oprettelse af autoritetsposter på originalt danske titler først skal ske, såfremt der kommer flere udgaver af samme værk. Dette er en praksis der pt. følges i USA. Fordelen herved vil være rationalet i ikke at oprette poster, der ikke bliver brug for – omvendt kan en sådan beslutning gøre danske data mindre brugbare internationalt, fx hvis en dansk bog, der kun er udkommet i en enkelt udgave i Danmark, bliver oversat. Generelt kan en sådan løsning accepteres her og nu, så længe vi befinder os inden for MARC-miljøet, men ved datamodel- og formatskifte må det sikres, at de fremtidige klienter vil håndtere oprettelse af værkposter smidigt i en integreret proces, så RDA til fulde efterleves.

Kapitel 2 : Overordnede RDA –regler

Resume: RDA mangler helt anvisninger for translitteration. Dette skal adresseres: Hvad skal man konkret gøre i danske biblioteker (både før og efter Unicode)? De manglende translitterationsregler er i en eller anden grad ligeledes en hæmsko for rationelt internationalt genbrug.

Ifølge RDA's hovedregel medtages artikler i navne og titler. En alternativ regel, der følges af mange biblioteker, tillader udeladelse af artikler i navne og titler. Dette er problematisk i en genbrugssammenhæng.

Derudover peges der på en række oplysninger, der pt. ikke er plads til indenfor det danske hjælpepostformat

Kapitel 8 angiver *General Guidelines on Recording Attributes of Persons, Families and Corporate Bodies*. Heraf fremgår bl.a. følgende:

8.4 Language and Script

Record names in the language and script in which they appear on the sources from which they are taken.

Alternative LC-PCC PS NLA PS BL PS D-A-CH

Record a transliterated form of the name either as a substitute for, or in addition to, the form that appears on the source.

Record other identifying attributes of a person, family, or corporate body in the language and script specified in the applicable instructions in chapters [9 RDA](#) – [11 RDA](#).

Kapitel 5 angiver *General guidelines on recording attributes of works and expressions*. Heraf fremgår bl.a. følgende:

5. 4 Language and Script

Record titles for works in the language and script in which they appear on the sources from which they are taken.

Alternative LC-PCC PS NLA PS BL PS D-A-CH

Record a transliterated form of the title either as a substitute for, or in addition to, the form that appears on the source.

Record other identifying attributes of a work or expression in the language and script specified in the applicable instructions in chapter [6 RDA](#).

Record the descriptive attributes of a work or expression covered in chapter [7 RDA](#) in a language and script preferred by the agency creating the data.

Dette punkt er en general central problemstilling: Hvordan forholder vi os til translitteration eller ej og i positivt fald: hvilke translitterationsstandarder skal præcis anvendes i en verden med genbrug på kryds og tværs? Der

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

er behov for translitteration bl.a. aht. de biblioteker, der anvender elementer af alfabetisk opstilling. RDA anfører ingen foretrukne standarder.

Dette kan ikke løses alene indenfor Danmark. Vi kan anlægge den pragmatiske holdning, at så længe danMARC2 eksisterer og opererer med udveksling ved hjælp af latin1-tegnsettet (på nær i felt 880), skal danske biblioteker principielt altid genbruge data translittereret til latinsk alfabet og efter de standarder, der pt. anvendes. Dog accepteres, at genbrugte data kan være translittereret efter anden standard. Hvis translitteration bevidst ønskes efter anden standard end angivet i nugældende regler, skal dette markeres enten som lokale data eller svarende til det, der indenfor det nugældende hjælpepostformats felt 008 *p er markeret med koder. Det langsigtede mål er naturligvis overgang til Unicode, men Unicode fjerner ikke behovet for translitteration aht. opstilling mv.

Indledende artikler: Dansk praksis har været at angive en indledende artikel, men markere hvis sortering skal ske på det efterfølgende led.

RDA har som hovedregel, at den indledende artikel angives, men LC-PCC, BL og NLA (alle engelsksprogede) benytter den alternative regel, der siger, at en indledende artikel springes helt over. D-A-CH benytter hovedreglen. Der er altså ikke nogen, der følger dansk praksis med en visbar form og en form til sortering.

Det tyske nationalbibliotek har været spurgt om deres overvejelser i forhold til genbrug ved at vælge hovedreglen men svar er desværre ikke modtaget endnu.

Hovedreglen er klart den mest anvendelige online, idet man ikke ved, om søgeren inkluderer en indledende artikel eller ej. Hvis hovedreglen vælges risikerer vi dog at skulle redigere mange værkposter på udenlandske værker fremfor at genbruge dem direkte, fordi en indledende artikel skal påføres.

Den kommende danske profil må endeligt fastlægge den fremtidige praksis for Danmark. Uanset hvilken praksis, der vælges, vil det få betydning i forhold til udveksling og genbrug af data.

Følgende data knytter sig til RDA's generelle kapitel 8 og korresponderer med det nuværende danske hjælpepostformat på følgende vis:

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|--|---|--------------------------------|
| Scope of usage (of the name chosen) (RDA 8.8) | Mangler | Minus |
| Date of usage (RDA 8.9) | Mangler | Minus |
| Status of identification (RDA 8.10) (enten: <i>fully established, provisional</i> eller <i>preliminary</i>) | 008 v og delvist også felt 042 – men matcher ikke fuldstændigt. | Medtages |
| Undifferentiated Name Indicator (RDA 8.11) (= der er flere med samme | Mangler | Minus |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | | |
|------------------------------|--|------------------------------|
| vedtagne form) | | |
| Sources consulted (RDA 8.12) | 670/671 | Medtages om muligt |
| Cataloguers note (RDA 8.13) | Mangler (667 er defineret som intern – men det er meningen at denne Cataloguers note skal udveksles) | Medtages i nødvendigt omfang |

Kapitel 3 : Hvad er iflg. RDA core på områderne?

Resume: Der peges på en stribe af core-oplysninger, der pt. ikke er plads til indenfor det danske hjælpepostformat.

Ligeledes adresseres væsentlige problemstillinger:

Anførsel af datoer. Såvel udformning (ikke sproguafhængigt iflg. RDA) som omfanget i indhentning af oplysninger kan være problematisk.

Anførsel af stednavne: RDA anviser, at man primært skal vælge den sproglige form, der tilsvarende den katalogiserende institutions sprog – alternativt stedets egen form (Rom eller Roma).

Nationale/lokale valg på begge områder kan i en eller anden grad modvirke internationalt samarbejde, men omvendt kan internationale valg være i modstrid med brugernes behov. Kan datamodellen løse problemet?

På titelområdet skal der tages stilling til, hvordan ophav og titler, der tilsammen udgør authorized accesspoint skal formateres i autoritetsposter, hvis et MARC-format vælges. Pt. er det kun muligt at anføre titeldata i felt 130. Da felt 100/110 og felt 130 ikke kan forekomme samtidigt i en post, kan felt 130 derfor kun bruges til anonyme værker.

Personnavne

Følgende elementer defineres som core iflg. RDA 0.6.7 og korresponderer med det nuværende danske hjælpepostformat på følgende vis:

| Element | Hjælpepostformat |
|--|---|
| Preferred name for the person | 100 a,h,e |
| Title of the person (adelstitler, religiøse titler mv.) | 100 f |
| Date of birth | Mangler (dog 100 c, hvis nødvendig del af authorized access point) |
| Date of death | Mangler (dog 100 d, hvis nødvendig del af authorized access point) |
| Other designation associated with the person(helgen, ånd mv.) | 100 f |
| Profession or occupation (kun hvis navnet ikke er et egentligt personnavn, men en frase el. lign.) | Mangler (dog 100 f, hvis nødvendig del af authorized access point) |
| Fuller form of name (RDA 9.5) (opløsning af initialer hvis | Mangler |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|---|--------------------------------|
| der ikke er anden måde at adskille to navne på) | |
| Identificer (ISNI, andre autoritetsposters id mv.) | Mangler delvist – men også 001 |

Konsekvent indhentning af fødseldato- og år samt opfølgning omkring dødsdato og –år problematisk, idet det vil kræve ekstra opsøgende arbejde og vedligeholdelse og dermed bl.a. fordyre den nationalbibliografiske registrering. Nationalbibliografien har tradition for at søge at indhente fødselsår på alle –sen-navne og en række udvalgte meget brugte navne (Holm, Smidt, Møller mv.).

Vi kan af ressourcemæssige hensyn blive nødt til at gradbøje forpligtelsen til: oplysningerne anføres som minimum, såfremt de indgår i forbindelse med authorized accesspoint eller hvis de fremgår af materialet.

Ved navneskift på personer forudsætter RDA, at der rettes bagud, således at alt omkring navnet fremgår af én autoritetspost. Dansk praksis har siden 1999 været at bibeholde gammel form i gamle poster og oprette ny form i nye poster. Konsekvensen heraf er flere autoritetsposter på samme person med henvisninger imellem.

Hvis vi vælger at følge RDA skal der findes en løsning på opstilling på folkebiblioteker mv. De fleste biblioteker retter nemlig ikke materialernes rygge, men lader dem stå under det navn de oprindeligt blev udgivet under.

Bemærkninger om enkelte dataelementer

Generelt er der et problem med datoer. Disse er i eksemplet i RDA 9.19.1.3 angivet som:

EXAMPLE

Smith, John, 1936 May 5–

Smith, John, 1936 December 17–

Det har fra flere sider været anført, at der burde vælges en i maskinsammenhæng mere hensigtsmæssig, sproguafhængig og international standard for angivelse heraf – om det så skal være 19361217 eller 19361712 ...

Danske poster har altid angivet det som *f. 1936-12-17*. Denne metode fremgår også af DBC's autoritetsbase på personnavne.

Vi opfordrer til at en dansk profil anviser, at der anvendes en ISO-standard for angivelse heraf

Familier

Følgende elementer defineres som core iflg. RDA 0.6.7 og korresponderer med det nuværende danske hjælpepostformat på følgende vis:

| Element | Hjælpepostformat |
|-------------------------------|------------------|
| Preferred name for the family | 100 a |
| Type of family | 100 f |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|--|---|
| Date associated with the family | Mangler (dog 100 f, hvis nødvendig del af authorized access point) |
| Place associated with the family (kun hvis nødvendigt for at skelne, RDA 10.5) | Mangler (dog 100 f, hvis nødvendig del af authorized access point) |
| Prominent member of the family (kun hvis nødvendigt for at skelne, RDA 10.5) | Mangler (dog 100 f, hvis nødvendig del af authorized access point) |
| Identifier | Mangler delvist – men også 001 |

Korporationsnavne

Følgende elementer defineres som core iflg. RDA 0.6.7 og korresponderer med det nuværende danske hjælpepostformat på følgende vis:

| Element | Hjælpepostformat |
|--|--|
| Preferred name for the corporate body (RDA 11.2.2) | 110 s,a,c |
| Location of conference, etc. (<i>Place associated with the corporate body is a core element for conferences, etc. For other corporate bodies, place associated with the corporate body is a core element when needed to distinguish a corporate body from another corporate body with the same name.</i> (RDA 11.3)) | 110 j, hvis der er tale om en konference Ellers: Mangler (dog 110 e, hvis nødvendig del af authorized access point for andet end en konference) |
| Number of a conference (RDA 11.6) | 110 i |
| Date of conference, etc. (<i>Date associated with the corporate body is a core element for a conference, etc. For other corporate bodies, date associated with the corporate body is a core element when needed to distinguish a corporate body from another corporate body with the same name</i> (RDA 11.4)) | 110 j, hvis der er tale om en konference Ellers: Mangler (dog 110 e, hvis nødvendig del af authorized access point for andet end en konference) |
| Associated institution (for conferences, etc., if the institution's name provides better identification than the local place name or if the local place name is unknown or cannot be readily determined) (RDA 11.5) | Mangler (dog 110 e, hvis nødvendig del af authorized access point) |
| Other designation associated with the corporate body (<i>Other designation associated with the corporate body is a core element for a body with a name that does not convey the idea of a corporate body. For other corporate bodies, other designation associated with the corporate body is a core element when needed to distinguish a corporate body from another corporate body with the same name.</i>) (RDA 11.7) - Type of corporate body (RDA 11.4.1.4) - Type of jurisdiction (RDA 11.7.1.5) | Mangler (dog 110 e, hvis nødvendig del af authorized access point) |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|---|--------------------------------|
| - Other designation (RDA 11.7.1.6) | |
| Identifier for the corporate body (RDA 11.12) | Mangler delvist – men også 001 |

I forhold til dansk praksis er der ikke stor forskel på, hvad der er core elements. Dog er der forskelle, når man kigger på udformningen af elementerne; her skal specielt nævnes nogle få forskelle.

Bemærkninger om enkelte dataelementer

Preferred name:

Om *datoer*, se kommentarer under personnavne.

Om *numre på konferencer*: RDA foreskriver, at disse skal angives som ordenstal (dansk praksis har været at bruge arabertal). RDA foreskriver endvidere, at ordenstal angives efter det pågældende lands praksis, altså fx 2nd (engelsk), 3^e (fransk), 4^o eller 4^a (italiensk), 5. (dansk og tysk).

Other designation associated with the corporate body:

Type of corporate body: De danske katalogiseringsregler har her blot: Hvis en korporations navn i sig selv ikke angiver, at det betegner en korporation, tilføjes en passende betegnelse af generel karakter. RDA angiver: Record the type of corporate body in a language preferred by the agency creating the data. Select terms from a standard list of names of types of corporate body, if available. If there is no equivalent term for the type of corporate body in a language preferred by the agency, or in case of doubt, record the type of corporate body in the official language of the corporate body.

Også her vil der altså være tale om, at tilføjelser oftest vil være udformet efter den katalogiserende institutions sprog.

Type of jurisdiction: Her angiver både de danske katalogiseringsregler og RDA at typen af jurisdiktion normalt angives på den katalogiserende institutions sprog

Anden identificerende tilføjelse: Ikke store forskelle mellem dansk praksis og RDA. Igen udformes tilføjelser på den katalogiserende institutions sprog

Stednavne

Det er pt ikke udarbejdet retningslinier for hvad der er core .

Bemærkninger om enkelte dataelementer

Navne på steder gengives iflg RDA altid i den katalogiserende institutions form. Dette gør det med nuværende datamodeller svært at genbruge data med stednavne som en del af indførslen. Det betyder, at ikke blot landenavne men også stednavne skal 'oversættes' til en dansk form af navnet, når en sådan er almindeligt anvendt, eller føres tilbage til originalsproget, hvis en dansk form ikke er kendt.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

For områder inden for et land, stat osv medtages landets/statens navn altid som identificerende tilføjelse. Eksempler fra RDA: *Kiev (Ukraine)*, *Region Sjælland (Denmark)*, *Lucca (Italy)*. Her har man i de danske katalogiseringsregler kun medtaget landenavne, hvis de var nødvendige for at skelne.

For byen Rom har de danske katalogiseringsregler formen *Roma*, mens RDA tilsvarende vil have *Rom (Italien)* som dansk form (og *Rome (Italy)*, hvis posten er udarbejdet i et engelsktalende land).

Alt dette gør det med nuværende datamodeller svært at genbruge data med stednavne som en del af indførslen. Det betyder, at ikke blot landenavne men også stednavne skal 'oversættes' til en dansk form af navnet, når en sådan er almindeligt anvendt, eller føres tilbage til originalsproget, hvis en dansk form ikke er kendt, samt at lande- eller statsnavne skal tilføjes til alle navne på mindre områder.

Kommende datamodeller kan forhåbentlig håndtere forskellige sprogs foretrukne navn for samme område og dermed lette udveksling.

I de danske katalogiseringsregler har man normalt undladt at anføre Danmark som første led i statslige korporationsnavne, og blot markeret udeladelsen indirekte i formatet. Denne praksis bør genovervejes, da den altid vil besværliggøre dataudveksling. I stedet kunne man måske overveje at undertrykke 'Danmark' som første led i indeksering og fremvisning i lokale kataloger.

RDA ser dog aktuelt på udformningen af stednavne, da det netop er et af de områder, der halter i et internationalt perspektiv. En repræsentant fra EURIG deltager i dette arbejde.

Titler

Et Authorized Access Point) for work eller expression er core og består af Authorized Access Point for ophav (person/korporation) + Preferred title for værket + evt. tillæg.

Som hovedregel er det værkets titel på originalsproget (for værker efter 1500) der udgør preferred title. For værker før 1500, diverse samlinger, musik mv. gælder dog en række særlige regler. En preferred title skal, hvis den ikke er unik, have diverse tilføjelser, afhængigt af hvad for slags materiale det er. Musiktitler får typisk flest tilføjelser, men også ved religiøse og juridiske værker, samt en del anonyme klassikere skal en tilføjelse anføres. Se videre 6.2.2.8-11.

Følgende data knytter sig til RDA's generelle kapitel 6 og korresponderer med det nuværende danske hjælpepostformat på følgende vis (regler for core skrevet med kursiv efter ethvert element). Det skal bemærkes, at det danske hjælpepostformat kun opererer med felt 130, hvilket modsvarer reglerne for anonyme værker. Der findes ingen dansk modsvar til formatering af værker med ophav.

| Element | Hjælpepostformat |
|--|--|
| Preferred title for a work (RDA 6.2.2.8) | Mangler (dog 130a hvis anonymt værk) |
| Preferred title for part of a work (RDA 6.2.2.9) | Mangler (dog 130s hvis |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|---|---|
| | anonymt værk) |
| Preferred title for a compilation of works (RDA 6.2.2.10) | Mangler (dog 130a hvis anonymt værk) |
| Form of work (RDA 6.3) <i>Core hvis RDA 6.27.19.1 opfyldes</i> | Mangler (dog 130 hvis anonymt værk) |
| Date of work (RDA 6.4) <i>Core for at differentiere mellem to værker med samme titel</i> | Mangler (dog 130u hvis anonymt værk) |
| Place of origin of the work (RDA 6.5) <i>Core for at differentiere mellem to værker med samme titel</i> | Mangler (dog muligvis 130 j hvis anonymt værk) |
| Other distinguishing characteristics of the work (RDA 6.6) <i>Core for at differentiere mellem to værker med samme titel</i> | Mangler (dog 130j hvis anonymt værk) |

Særregler for musik:

| Element | Hjælpepostform |
|--|---|
| Preferred title for a musical work (RDA 6.14.2.1 – 6.14.2.5) | Mangler (dog 130a hvis anonymt værk) |
| Preferred title for a part or parts of a musical work (RDA 6.14.2.7) | Mangler (dog 130s hvis anonymt værk) |
| Preferred title for a compilation of works (RDA 6.14.2.8) | Mangler (dog 130a hvis anonymt værk) |
| Medium of performance (RDA 6..15.1) | |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|--|---|
| <p><i>Medium of performance is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i></p> | <p>Mangler (dog 130d hvis anonymt værk)</p> |
| <p>Numeric designation of a musical works (RDA 6.16.1.3) <i>Numeric designation is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i></p> <ul style="list-style-type: none"> • Serial number (RDA 6.13.1.3.1) • Opus number (RDA 6.13.1.3.2) • Thematic index number (RDA 6.13.1.3.3) | <p>Mangler (dog 130 e,f,g hvis anonymt værk) <i>Et lille formatmæssig problem her da danMARC2 skiller på numeriske betegnelser og opusnummerb etegnelse/te matiske nummer. Det gør man ikke i MARC21. DanMARC adskiller også delnummer til opus i g. Ikke hellere det findes i MARC21.</i></p> |
| <p>Key (RDA 6.17.1.3) <i>Key is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i></p> | <p>Mangler (dog 130h hvis anonymt værk)</p> |
| <p>Arrangements, transcriptions etc. (RDA 6.18.1.4) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i></p> | <p>Mangler (dog 130k hvis anonymt værk)</p> |
| <p>Sketches (RDA 6.18.1.5) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i></p> | <p>Mangler (Man kunne bruge 130k hvis anonymt)</p> |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|--|---|
| | værk) |
| Vocal and chorus scores (RDA 6.18.1.6) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i> | Mangler (Man kunne bruge 130 k hvis anonymt værk) |

Kapitel 4 : Hvad vil vi udover core-elementerne anbefale som elementer i et dansk niveau på områderne og i overensstemmelse med mulighederne i RDA og hvorfor (hvilke brugerbehov er der tale om)?

Resume: Der gives en række anbefalinger for hvilke oplysninger det vil være hensigtsmæssigt at medtage.

Den danske profil er dog ikke hermed lagt, idet der er tale om minimumsoplysninger. Juridiske aspekter omkring typen af oplysninger på personer bør vurderes af en juridisk fagperson. Det vises, hvilke elementer der pt. ikke er plads til i hjælpepostformatet.

Personnavne

Udover core-elementerne er det muligt at anføre nedenstående oplysninger iflg. RDA.

Foreslået dansk niveau dækker her over, hvad der hvis muligt minimum foreslås medtaget ved dansk oprettelse af et navn.

I tilfælde af genbrug (fysisk eller on the fly) af udenlandske data foreslås alle elementer medtaget, hvis de forekommer:

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|---|------------------|--------------------------------|
| Gender (RDA 9.7) | Mangler | Medtages, hvis det fremgår |
| Country associated with the person (RDA 9.10) | 678 n | Medtages, hvis det fremgår |
| Place of residence etc. (RDA 9.11) | Mangler | Minus |
| Address of the person (RDA 9.12) | Mangler | Minus |
| Affiliation (RDA 9.13) | 678 f | Minus |
| Language of the person (RDA 9.14) | 678 l | Medtages, hvis det fremgår |
| Field of activity (RDA 9.15) | Mangler | Minus |
| Profession or occupation der ikke optræder som nødvendig del af authorized accesspoint (RDA 9.16) | Mangler | Medtages, hvis det fremgår |
| Biographical information (RDA 9.17) | 678 a | Minus |

De i dansk sammenhæng nye elementer foreslås medtaget af hensyn til mere specifikke søgninger (køn, nationalitet, originalsprog), der alle har været efterlyst i det nuværende bibliografiske miljø.

Uddannelse/erhverv foreslås medtaget, fordi oplysningen ofte vil være mere velegnet til identifikation end fødselsår. Man kunne forestille sig oplysningen som en ekstra tilføjelse til authorized accesspoints ved opslag i registre eller i forbindelse med navigation på nettet ved autoritetsposter som linked data.

En lang række data er ikke foreslået at indgå i et dansk minimumsniveau. Dels fordi oplysningerne i mange tilfælde ikke vil være tilgængelige, dels fordi flere af oplysningerne skønnes at ville bringe os på kant med

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

persondataloven. Vi kan formentlig ikke registrere tilknytninger til foreninger, biografisk information der fx indeholder oplysninger om religiøsitet mv. Generelt bør indholdet af en fuld RDA-registrering af en personentitet gennemgås juridisk, så vi får klart svar på, hvilke oplysninger vi præcis må registrere, hvilke vi må gengive fra andre mv.

Familier

Udover core-elementerne er det muligt at anføre nedenstående oplysninger iflg. RDA.

Foreslået dansk niveau dækker her over, hvad der hvis muligt minimum foreslås medtaget ved dansk oprettelse af et navn.

I tilfælde af genbrug (fysisk eller on the fly) af udenlandske data foreslås alle elementer medtaget, hvis de forekommer:

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|-----------------------------------|------------------|--------------------------------|
| Hereditary title (RDA 10.7) | Mangler | Minus |
| Language of the family (RDA 10.8) | 678 l | Minus |
| Family history (RDA 10.9) | 678 a | Minus |

Som det ses anbefalet kun core som dansk minimum – dette set i lyset af behovet.

Korporationsnavne

Udover core-elementerne er det muligt at anføre nedenstående oplysninger iflg. RDA.

Foreslået dansk niveau dækker her over, hvad der hvis muligt minimum foreslås medtaget ved dansk oprettelse af et navn.

I tilfælde af genbrug (fysisk eller on the fly) af udenlandske data foreslås alle elementer medtaget, hvis de forekommer:

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|------------------------------------|------------------|--------------------------------|
| Date of establishment (RDA 11.4.4) | Mangler | Medtages, hvis det fremgår |
| Date of termination (RDA 11.4.4.1) | Mangler | Medtages, hvis det fremgår |
| Period of activity (RDA 11.4.5) | Mangler | Minus |
| Address of the corporate body | Mangler | Minus |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | | |
|---|--|---|
| (RDA 11.9) | | |
| Language of the corporate body (RDA 11.8) | Mangler (i hjælpepostformatet er sprog i 678 l knyttet til personer, ikke korporationer) | Medtages, hvis det fremgår, når der er tale om internationale korporationer |
| Field of activity of the corporate body (RDA 11.10) | Mangler | Minus |
| Corporate history (RDA 11.11) | 678 a ? Det fremgår ikke helt af hjælpepostformatet, om hele 678 er beregnet på personer | Medtages, hvis det fremgår, og er nyttigt for at forstå valg af navneform, eller for at fastholde en kompliceret sammenhæng med andre entiteter |

For navneskift vil det være meget nyttigt at have oplysninger om den periode et givet navn dækker. Eksempel: Kulturministeriet, Kultur- og Kommunikationsministeriet, Ministeriet for kulturelle anliggender osv., som alle er skiftende navne for det samme ministerium. Derfor anbefales det, at lægge date of establishment og date of termination ind når de let kan findes, især for korporationer med tidligere eller senere navneskift.

Language of the corporate body er specielt nyttigt til at forstå, hvorfor en given sprogform er valgt som autoriseret navn.

Corporate history kan bidrage til i mere tekstlig form at redegøre for forhold, der tjener til at identificere et ophav. Eksempel fra RDA: The Salem Female Academy was originally founded in 1772 and chartered as Salem Female Academy and College in 1866. In 1907 name was changed to Salem Academy and College, and in 1912, the institution was separated into Salem College and Salem Academy

Stednavne

For stednavne anføres nedenstående oplysninger iflg. RDA

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|--|---|---|
| Preferred name for the place (RDA 16.2.2) | 154 a, hvis der er tale om emne, men indgår også f.eks. i 110 s, e, j | Medtages |
| Terms indicating type of jurisdiction (RDA 16.2.2.8) | Mangler til generel identifikation af stednavne men kan angives i parentes i samme delfelt som stednavnet | Medtages |
| Identifier for the place (RDA 16.3) [To be added in a later release] | Mangler | Der træffes først afgørelse, når RDA er blevet udformet |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Titler

Udover core-elementerne er det muligt at anføre nedenstående oplysninger iflg. RDA.

Foreslået dansk niveau dækker her over, hvad der hvis muligt minimum foreslås medtaget ved dansk oprettelse af en titel

Der skal noteres at det danske hjælpepostformat kun har felt 130, hvilket modsvarer reglerne for anonyme værker. Der findes intet dansk modsvar til værker med ophav.

I tilfælde af genbrug (fysisk eller on the fly) af udenlandske data foreslås alle elementer medtaget, hvis de forekommer:

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|--|------------------|---|
| Coded dates | Mangler (048) | Medtages, hvis det fremgår |
| Content type | Mangler (336) | Medtages, hvis det fremgår |
| Format of notated music | Mangler (348) | Medtages, hvis det fremgår |
| Form of work | Mangler (380) | Medtages, hvis det fremgår |
| Other Distinguishing Characteristics of Work or Expression | Mangler (381) | Medtages, hvis det fremgår (men måske mindre vigtigt) |
| Medium of Performance | Mangler (382) | Medtages, hvis det fremgår |
| Numeric Designation of Musical Work | Mangler (383) | Medtages, hvis det fremgår |
| Key | Mangler (384) | Medtages, hvis det fremgår |
| Audience Characteristics | Mangler (385) | Medtages, hvis det fremgår |
| Time Period of Creation | Mangler (388) | Minus |

Kapitel 5: Variant access points

Resume: Variant Access Point svarer til danske henvisninger og er ikke core ifølge RDA. En praksis som stort set svarer til den nuværende danske praksis, vil derfor kunne videreføres. I Danmark vil vi derudover gerne anvende udvalgte variant access points som dansk hyldeopstilling, ligesom vi gerne vil bygge bro mellem opstilling efter nuværende danske standardtitler og authorized accesspoint for work fra RDA. Markering af fælles dansk opstilling bør kunne anføres i autoritetsposten.

Endvidere dokumenteres mangler i det nuværende hjælpepostformat.

Variant access points svarer overordnet set til det, der i de danske katalogiseringsregler kaldes 'se-henvisninger'. Hos DBC (og dermed også folkebibliotekerne) optræder de i form af 9xx-felter i de enkelte bibliografiske poster. Traditionelt skelner danMARC2s bibliografiske format ikke mellem se- og se også-henvisninger, sådan som RDA skelner. Et manglende delfelt x antyder dog, at der er tale om en se-henvisning.

Det danske hjælpepostformat skelner: at se-henvisninger er formateret i et felt 4xx og se også-henvisninger er formateret i et felt 5xx

Katalogiseringsreglerne anfører: Når navnet på en person eller korporation eller en uniform titel er kendt, eller kan tænkes at blive søgt, under en anden form end den, der er anvendt som den vedtagne form, udformes en henvisning fra den ikke-anvendte form til den anvendte form. [...] Henvisninger udformes dog ikke fra former, der ligner den vedtagne form eller hinanden så meget, at de er overflødige.

Der er ikke nogen krav om at lave indførsler på variantnavne

RDA angiver ikke nogen relationship designators for forholdet mellem variant access point og authorized access point (man kan altså ikke se typen af variation som f.eks. fuldstændighed, anden sprogform, anden translitteration osv.)

Personnavne

RDA angiver en række eksempler, hvor man kan lave variant access points (RDA 9.2.3):

Variantnavn, anden translitteration, virkeligt navn, hvis kun et pseudonym har været anvendt, verdsligt navn, hvis det gejstlige navn er valgt (eller omvendt), tidligere navn eller senere navn, når disse ikke er valgt, anden sproglig variant af navnet, anden del af navnet, mere eller mindre fuldstændigt navn osv. Dette ligger helt på linje med de danske katalogiseringsregler.

Det anbefales derfor at den danske profil anviser at fortsætte med at lave indførsler svarende til hidtidigt niveau for se-henvisninger.

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|--|------------------|---|
| Variant name for the person (RDA 9.19.2) | 400 a, h, e | Medtages hvis det med rimelighed kan anses for at |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | | |
|---|-------|-------------------------------------|
| | | blive søgt på (katalogisators skøn) |
| Title of the person (adelstitler, religiøse titler mv.) (if considered important for identification) | 400 f | Medtages, hvis vurderet relevant |
| Date of birth (if considered important for identification) | 400 c | Medtages, hvis vurderet relevant |
| Date of death (if considered important for identification) | 400 d | Medtages, hvis vurderet relevant |
| Other designation associated with the person(helgen, ånd mv.) (if considered important for identification) | 400 f | Medtages, hvis vurderet relevant |
| Profession or occupation (kun hvis navnet ikke er et egentligt personnavn, men en frase el. lign.) (if considered important for identification) | 400 f | Medtages, hvis vurderet relevant |
| Fuller form of name (RDA 9.5) (opløsning af initialer hvis der ikke er anden måde at adskille to navne på) | 400 k | Medtages, hvis vurderet relevant |

Hvis henvisningen går fra virkeligt navn til pseudonym, skal resultatet af den juridiske vurdering, der pt. finder sted om dette forhold i de danske regler, naturligvis respekteres også under RDA.

For variant access points ligger der allerede i dag mange muligheder for at hente data, dels fra eksisterende danske se-henvisninger, dels fra VIAF, LC name headings mv.

Familier

RDA angiver en række eksempler, hvor man kan lave variant access points (RDA 10.2.3):

Anden translitteration, anden del af hierarkiet ved underordnede korporationer, fuldt udskrevne former, akronymer, osv.

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|--|------------------|---|
| Variant name for the family (RDA 9.19.2) | 400 a | Medtages hvis det med rimelighed kan anses for at blive søgt på (katalogisators skøn) |
| Type of family | 400 f | Medtages |
| Date associated with the family (if considered important for | 400 f | Medtages, hvis vurderet relevant |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | | |
|--|-------|----------------------------------|
| identification) | | |
| Place associated with the family (if considered important for identification)) | 400 f | Medtages, hvis vurderet relevant |
| Prominent member of the family (if considered important for identification) | 400 f | Medtages, hvis vurderet relevant |

For variant access points ligger der allerede i dag mange muligheder for at hente data, dels fra eksisterende danske se-henvisninger, dels fra VIAF, LC name headings mv.

Korporationsnavne

RDA angiver en række eksempler, hvor man kan lave variant access points (RDA 11.2.3):

Anden translitteration, anden lingvistisk form, anden del af navnet, anden bogstavering osv.

Dette svarer til de nuværende danske katalogiseringsregler.

| Element | Hjælpepostformat | Foreslået dansk minimumsniveau |
|---|-------------------|---|
| Variant name or form of name | 410 s, a, c, h, g | Medtages hvis det med rimelighed kan anses for at blive søgt på (katalogisators skøn) |
| Type of corporate body if considered important for identification | 410 e | Medtages, hvis vurderet relevant |
| Place associated with the body if considered important for identification | 410 e | Medtages, hvis vurderet relevant for identifikation |
| Associated institution if considered important for identification | 410 e | Medtages, hvis vurderet relevant |
| Date associated with the body if considered important for identification | 410 e | Medtages, hvis vurderet relevant |
| Type if jurisdiction if considered important for identification | 410 e | Medtages, hvis vurderet relevant |
| Other designation if | 410 e | Medtages, hvis vurderet |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | | |
|--|-------------|----------|
| considered important for identification | | relevant |
| Number, date and location of a conference, etc | 410 i, k, j | Medtages |

Med de bemærkninger der er om authorized access point vedrørende sprog vil der være mulighed for genbrug af data.

Stednavne

RDA har endnu ikke udformet regler for dette

Titler

RDA angiver en række eksempler, hvor man kan lave variant access points (RDA 6.2.3):

Varianttitel, anden translitteration, tidligere titel eller senere titel, anden sproglig variant af titlen, mere eller mindre fuldstændig titel osv. Dette ligger helt på linje med de danske katalogiseringsregler.

Til en ikke unik varianttitel kan anføres tilføjelser. Type af tilføjelse afhænger af typen af materiale - præcis på samme måde som Authorized Accesspoint. Musik har flest tilføjelser, men også ved religiøse og juridiske værker samt en del anonyme klassikere er tilføjelser relevante. Se videre 6.2.2.8-11. Tilføjelse til varianttitler må formodes anvendt i begrænset omfang.

Dansk hyldeopstilling bør kunne fremgå af expressionoplysningerne (fx den danske titel på en til dansk oversat roman). Oplysningen bør fremgå som en markering på relevant varianttitel. En sådan markering mangler ligeledes formatmæssigt.

Det samme gør sig gældende med musikvarianttitler, der er opstillingselement i en dansk kontekst.

Ligeledes vil det være relevant at markere evt. gamle standardtitler, der bør anføres som varianttitler, som værende gamle standardtitler, således at der kan bygges bro mellem gammel og ny opstilling.

| Element | Hjælpepostformat |
|--|------------------|
| Alternative Linguistic Form of Title for the Work (RDA 6.2.3.4) | Mangler |
| Other Variant Title for the Work (RDA 6.2.3.5) | Mangler |
| Variant title of a musical work (RDA 6.14.3) | Mangler |
| Form of work (RDA 6.3) <i>Core hvis RDA 6.27.19.1 opfyldes</i> | Mangler |
| Date of work (RDA 6.4) <i>Core for treaty og for at differentiere mellem to værker med samme titel</i> | Mangler |
| Place of origin of the work (RDA 6.5) <i>Core for at differentiere mellem to værker med samme titel</i> | Mangler |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|---|---------|
| Other distinguishing characteristics of the work (RDA 6.6) <i>Core for at differentiere mellem to værker med samme titel</i> | Mangler |
|---|---------|

Særregler for musik

Der skal noteres at det danske hjælpepostformat ikke har felt 430 (men 530), for se-henvisninger for anonyme værker. Der findes intet dansk modsvar til værker med ophav.

| Element | Hjælpepostformat |
|--|---|
| Variant title for a musical work (RDA 6.14.3) | Mangler |
| Medium of performance (RDA 6.15.1) <i>Medium of performance is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i> | Mangler |
| Numeric designation of a musical works (RDA 6.16.1.3) <i>Numeric designation is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i> <ul style="list-style-type: none"> • Serial number (RDA 6.13.1.3.1) • Opus number (RDA 6.13.1.3.2) • Thematic index number (RDA 6.13.1.3.3) | Mangler Et lille formatmæssig problem her da danMARC2 skiller på numeriske betegnelser og opusnummerbetegnelser/tematiske numrer. Det gør man ikke i MARC21. DanMARC adskiller også delnummer til opus i g. Ikke heller det findes i MARC21. |
| Key (RDA 6.17.1.3) <i>Key is a core element when needed to differentiate a musical work from another work with the same title. It may also be a core element when identifying a musical work with a title that is not distinctive</i> | Mangler |
| Arrangements, transcriptions etc. (RDA 6.18.1.4) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i> | Mangler |
| Sketches (RDA 6.18.1.5) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i> | Mangler |
| Vocal and chorus scores (RDA 6.18.1.6) <i>Other distinguishing characteristic of the expression is a core element when needed to differentiate an expression of a work from another expression of the same work</i> | Mangler |

Kapitel 6: Relationships between persons, families, and corporate bodies

Resume: I Danmark har vi tradition for "se også" henvisninger mellem tidligere og senere navne. Det er stort set den eneste relation vi pt. har registreret. RDA har ingen core-elementer omkring beskrivelse af relationer og relationer kan udtrykkes på forskellig vis (som tekststreng, i noteform mv.). Området er i det hele taget lidt uopdyrket og synes at udvikle sig ved det øgede fokus på entitetsregistrering. En række relationer er oplagte også at belyse (fx forholdet mellem en musikgruppe og dens medlemmer, personer og embeder). Det vil være på linje med dansk praksis for funktionskoder at udtrykke en relation i kodeform (med en underliggende identifikator for det indhold, koden dækker).

Relationships between persons, families, and corporate bodies svarer overordnet set til det, der i de danske katalogiseringsregler kaldes 'se også-henvisninger'. Hos DBC (og dermed også folkebibliotekerne) optræder de i form af 9xx-felter i de enkelte bibliografiske poster. Traditionelt skelner danMarc2s bibliografiske format ikke mellem se- og se også-henvisninger, men hjælpepostformatet skelner, lige som RDA skelner. Et delfelt x, som (bl.a.) indeholder ordene 'se også' antyder dog, at der er tale om en se også-henvisning.

Om disse anfører de danske katalogiseringsregler:

Se også-henvisninger - faste henvisninger i forbindelse med bevidst brug af flere navne: Hvis en person skriver under flere navne (f.eks. virkeligt navn og pseudonym), vil der i katalogen være flere samtidige vedtagne former af navne for samme person. I sådanne tilfælde udformes faste se også-henvisninger mellem de forskellige vedtagne former for de anvendte navne. Til denne type henvisninger kan der i beskrivelsen udformes en forklarende note.

Se også-henvisninger - postspecifikke henvisninger i forbindelse med ændringer af vedtagen form: Hvis en person eller en korporation forandrer navn, vil der i katalogen være flere vedtagne former af navne for personen eller korporationen. I tilfælde af navneskifte udformes derfor postspecifikke se også-henvisninger i de første af de bibliografiske poster, hvori det nye navn anvendes, således at brugeren ledes frem og tilbage mellem det tidligere og senere navn. Til denne type henvisninger kan der ud over teksten se også være tilføjet en forklarende tekst. [...]Flere samtidige former vil endvidere kunne eksistere i katalogen, når der er udformet et opslag på et statsoverhoved mv., der skal dække det embede, vedkommende besidder, samtidig med at der er oprettet en navneform for vedkommende som person. I sådanne tilfælde henvises der mellem de forskellige samtidigt anvendte former.

Det er ikke et krav at lave relationer mellem forskellige entiteter

Der er ikke i RDA noget minimumsniveau angivet. Om relationer siges der:

The data recorded to reflect relationships between persons, families, and corporate bodies should enable the user to:

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

a)

find persons, families, or corporate bodies that are related to the person, family, or corporate body represented by the data retrieved in response to the user's search

b)

understand the relationship between two or more persons, families, or corporate bodies.

To ensure that the data created using RDA meet those functional objectives, the data should reflect all significant bibliographic relationships between related persons, families, and corporate bodies.

RDA foreskriver ikke en bestemt metode for at angive relationer. Disse kan enten etableres som relationer mellem identifiere, som relationer mellem authorized access points eller som en tekstlig beskrivelse ('note').

Relationship designators kan anvendes. I de danske katalogiseringsregler har det hidtil kun været praksis at benytte 'se også under det tidligere navn' og 'se også under det senere navn', evt udbygget med en note i den bibliografiske post.

RDA har en hel række betegnelser for relationer mellem korporationer og andre entiteter, herunder andre korporationer, se RDA afsnit 29-32 og K. Disse betegnelser kan benyttes om muligt. Dette vil kunne ske i form af en tekststreng. Senere må man antage, at denne tekststreng kan erstattes af en identifier.

Som minimum bør der laves relationer mellem tidligere og senere navne for samme entiteter, og for virkeligt navn og pseudonym, når begge har været anvendt som opslag.

Det vil være relevant at anvende Relationship designators fra RDA som fx *hierarchical subordinate* og *absorbed corporate body*, *alternate identity*, *member* (fx for en person, der er medlem af en musikgruppe)

Format:

Alle de dataelementer, der indgår i et authorized access point skal kunne registreres, derudover skal man kunne angive en relationship designator.

De store institutioner i de angelsaksiske lande og Tyskland er godt i gang med at ændre deres autoritetsposter til RDA. I den sammenhæng lægger de også links til relaterede entiteter ind.

Kapitel 7: Værkdata – skal vi gentage eller referere når værkdata sammensættes af elementer fra forskellige entiteter?

Resume: Eksemplerne illustrerer centrale problemstillinger: Når et ophav indgår i beskrivelsen af et værk, skal ophavet så optræde i udskrevet form eller som link til autoritetsposten på ophavet?

Skal der oprettes selvstændige autoritetsposter på værket og på dets expressions og skal værkets titel på en expression optræde som link?

Grundlæggende anbefaler vi linkning fremfor fuldt udskreven form.

I MARC21 i VIAF ses ophav knyttet til værket som anvist nedenfor ligesom alle expressions af et værk ses behandlet fuldt udskrevet enkeltvis som nedenfor. Der gøres brug af gentagelser men også i en vis grad links til andre entiteter.

Litterært værk

Autoritetspost-personnavn

035 00 #a (OCoLC)oca03756841

035 00 #a (ISNI) 0000000120966265

035 00 #a (SELIBR)355636

100 10 #a Blixen, Karen, #c 1885-1962

500 00 #a Dinesen, Isak, #c 1885-1962

500 00 #a Andrézel, Pierre, #c 1885-1962

998 00 #a Dinesen, Isak, #d 1885-1962 #2 (ISNI) 0000000118910457

998 00 #a Andrézel, Pierre, #c 1885-1962#2 (ISNI)0000000107972409

Autoritetspost-værk:

035 00 #a(VIAF)184814361

100 10 #a Blixen, Karen, #c 1885-1962. #t Den afrikanske farm

993 00 #a Blixen, Karen, #c 1885-1962 #2 (ISNI) 0000000120966265#3 worktoname

(I VIAF ser posten ikke sådan her ud. LC bruger Dinesen som ophav til værket og ser den engelske titel "Out of Africa" som preferred title. Men konstruktionen bliver den samme)

Autoritetspost-expression 1:

035 00 #a(VIAF)184814361

100 10 #a Blixen, Karen, #c 1885-1962. #t Den afrikanske farm. #l Engelsk

400 10 #a Blixen, Karen, #c 1885-1962. #t Out of Africa

993 00 #a Blixen, Karen, #c 1885-1962 #2 (ISNI) 0000000120966265#3 exprtoname

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Autoritetspost-expression 2:

035 00 #a(VIAF)251331799
035 00 #a (SELIBR)355636
100 10 #a Blixen, Karen, #c 1885-1962. #t Den afrikanske farm. #l Svensk
400 10 #a Blixen, Karen, #c 1885-1962. #t Den afrikanska farmen
400 10 #a Blixen, Karen, #c 1885-1962. #t Afrikansk pastoral
678 00 #a Ursprungligen utgiven 1937 med titeln: Afrikansk pastoral
993 00 #a Blixen, Karen, #c 1885-1962 #2 (ISNI) 0000000120966265#3 exprtoname
993 00 #a Blixen, Karen, #d 1885-1962 #2 (SELIBR)41509 #3 exprtoname

Måske skulle der også være links fra expressions-posterne til værkposten, men det synes man ikke at have pt.

Musikværk

Autoritetspost-personnavn

035 00 #a (ISNI) 0000000108617388
035 00 #a (VIAF) 197250
100 10 #a Nielsen, Carl, #c 1865-1931
400 00 #a Nielsen, Carl August, #c 1865-1931

Autoritetspost-værk

035 00 #a (VIAF)175195654
046 00 #k 1922
100 00 #a Nielsen, Carl, #d 1865-1931. #t Kvintetter, #m fløjte, obo, klarinet, horn, fagot, #n op. 43
400 00 #a Nielsen, Carl, #d 1865-1931. #t Kvintetter, #m strygere, #n op. 43
400 00 #a Nielsen, Carl, #d 1865-1931. #t Kvintet for fløjte, obo, klarinet, horn og fagot, opus 43
382 00 #a fløjte #n 1 #a obo #n 1 #a klarinet #n 1 #a horn #n 1 #a fagot #n 1 #s 5 #2 lcmp
383 00 #b op. 43
700 00 #a Nielsen, Carl, #d 1865-1931. #t Quintets, #m flute, oboe, clarinet, horn, bassoon, #n op. 43
993 00 #a Nielsen, Carl, #d 1865-1931 #2 (VIAF) 197250 #3 worktoname

Modelmæssigt kunne man sagtens forestille sig, at ophavets navn i værkautoritetsposten udelukkende var et link til autoritetsposten på ophavet og altså ikke udskrevet.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Hvor sammenkoblede strukturer er det hensigtsmæssigt at etablere når man samtidig vil bevare overblik?

Omvendt vil en flad struktur uden maskinelt link mellem ophav og værkpost hindre maskinelt dannet overblik over fx relationer mellem ophav og flere værker ligesom manglende link mellem værkpost og poster på expressions vil hindre maskinel dannelse af fuldt overblik over omfanget af alle et værks expressions.

I forbindelse med diverse workshops har værktøjet RIMMF været anvendt for at illustrere entitetsregistrering i et ikke-MARC-miljø. I dette værktøj er der anvendt link mellem ophav og værktitler fremfor udskreven form af ophavet. Grundlæggende vurderer vi, at det ikke er ideelt med udskrevne former. Linkning vil fungere bedre i forbindelse med rettelser ligesom linkning må formodes at ville fungere bedre i en linked data sammenhæng. Omvendt er linkning en udfordring i forbindelse med visning og i lokalkatalogen.

I nogle tilfælde vil en dansk ønsket opstilling i folkebibliotekerne adskille sig fra RDA's preferred title. (fx *Tryllefløjten* fremfor *Die Zauberflöte* eller den danske titel på en roman fremfor den udenlandske originaltitel). I arbejdsgruppen har vi ikke været helt enige om, hvordan det bør løses.

En mulighed er, at det konsekvent altid fremgår på registreringen af manifestationer.

En anden mulighed er at man ud fra en autoritetspost på en expression skal kunne identificere en evt. særlig dansk opstilling på hylden af den aktuelle expression, såfremt et værk er selvstændigt udgivet. Når dette foreslås, skyldes det at opstilling i en folkebibliotekssammenhæng bliver betragtet som en naturlig del af autoritetsdata.

Kapitel 8: Vil det være muligt at genbruge internationalt tilgængelige data?

Resume: Der peges på de enkelte genbrugsmuligheder internationalt, men det påpeges også at problemer med sprogformer og translitteration fortsat vil være en hindring for et ikke "håndholdt" genbrug.

Personnavne

DBC's autoritetsbase med personnavne forventes allerede nu at kunne hente dataleverancer med oplysninger som ISNI, køn og nationalitet fra VIAF for at flette dem ind i de danske poster.

Særligt forskningsbiblioteker vil grundet megen udenlandsk litteratur have behov for langt flere autoritetsdata end dem, der fremgår af DBC's autoritetsbase. Her er VIAF også den naturlige mulighed, men VIAF er en foreningsmængde af mange sæt af autoritetsdata, så der vil skulle lægges en prioritering ved genbrug. På latinske alfabeter kunne dette være, at autoritetsdata primært genbruges fra nationalbibliografier, der omfatter personens værker på originalsproget. Dette svarer til hidtidig tradition på området og er også om end ikke direkte udtrykt så i tråd med RDA 8.2:

8.2 Functional Objectives and Principles

...

Representation. The name or form of name chosen as the preferred name for a person, family, or corporate body should be:

a)

the name or form of name most commonly found in resources associated with that person, family, or corporate body

or

b)

a well-accepted name or form of name in a language and script preferred by the agency creating the data.

...

Language preference. The preferred name for a person, family, or corporate body should be the name or form of name found in resources associated with that person, family, or corporate body in the original language and script of the content. In some cases, the original language and script is not a language and script preferred by the agency creating the data. When this occurs, the preferred name or form of name should be one found in resources associated with that person, family, or corporate body, or in reference sources, in a language and script preferred by the agency.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Common usage or practice. When there is more than one part in the name of a person or family, the part chosen as the first element of the preferred name should reflect the usage or practice in the country and language most closely associated with that person or family.

Ved ikke-latinske alfabeter kunne man som hovedregel foretrække autoritetsdata fra Library of Congress. Under alle omstændigheder bør der tages stilling til translitteration.

Om genbruget så skal ske rent fysisk eller on the fly er en del af afgrænsningen for det nationale repositorium , datamodel og forventet funktionalitet.

Familier

Når DBC's personnavne som emneord (herunder familier) er lagt ind i DBC's autoritetsbase for personnavne vil de kunne beriges i forhold til VIAF på linje hermed. Om yderligere er nødvendigt i en dansk kontekst er tvivlsomt, omfanget taget i betragtning.

Korporationsnavne

En række oplysninger vedrørende korporationsnavne vil være at finde i VIAF og hos LC. For valg af sprogformer, når et korporationsnavn optræder med flere former, og for translitterationer samt for stednavne og tilføjelser til navne (stednavne, typer af korporationer, typer af jurisdiktioner) vil der dog være en del tilretning, da RDA så entydigt fokuserer på den katalogiserende institutions sprog. Dertil kommer den udbredte praksis med udeladelse af indledende artikler, der såfremt artikler vælges medtaget i en dansk kontekst, også vil øge behovet for tilretninger.

Stednavne

Det vil være vanskeligt at genbruge stednavne, når RDA så entydigt benytter den katalogiserende institutions sprog som udgangspunkt for valg af navn. Dette gør det dog ikke mindre nødvendigt at koble sig på andres arbejde med identifikatorer for samme navn, f.eks. VIAF, GeoNames, LC authority files.

Titler

Titler kan ligeledes være vanskelige at genbruge. Jo tættere man følger RDA desto bedre muligheder for genbrug. Reglen siger, at preferred title er det samme som værkets originaltitel, hvilket er enkelt genbrugsmæssigt. Men der findes alligevel mange muligheder for variationer. - især når det drejer sig om titler fra sprog, som man ikke forstår eller titler fra andre skriftssprog og alfabeter.. Nøddeknækkeren er et eksempel herpå. Originaltitlen er egentlig *Щелкунчик*, eller med dansk translitterering *Sjtjelkuntjik*, alternativt

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Šelkuntčik (translittereret med ISO 9). Selvom vi skulle vælge originaltitlen, så kan vi ikke genbruge de engelsksprogede data. I en engelsksproget katalogpost er preferred title nemlig *Shchelkunchik* baseret på Library of Congress' translitterering.

Musiktitler er generelt ikke velegnede til genbrug. Man skal vælge form/genre i henhold til det sprog som katalogiseringen er på. En dansk tilføjelse vil f.eks. være Symfonier, medens den engelske er Symphonies, og den tyske Sinfonien. Også instrumentnavne mv. skal være på det lokale sprog.

Preferred titles er formentlig mest velegnet til sammenligning, men ikke til genbrug. Man kan dog lave varaintformer på andre sprog indgå i autoritetsposter, hvis man ønsker. Alternativt kan man forlade sig på at autoritetsposter fra forskellige lande matches i VIAF og dermed linkes sammen.

Kapitel 9: Samlet anbefaling

Gruppen havde følgende opgave

- Indstille til BIR hvilke typer af autoritetsdata man anbefaler skal være de danske også med udgangspunkt i de internationalt tilgængelige data eksempelvis VIAF
- Opstille begrundelser for anbefalingerne
- Med udgangspunkt i de anbefalede autoritetsdata angive de sammenhænge der er til RDA
- Endvidere opstille formatmæssige krav og udfordringer

Genbrug

Analysen viser, at direkte genbrug af udenlandske autoritetsdata som danske autoritative data i mange tilfælde ikke vil være muligt:

Årsagerne hertil er flere:

- Som genbrugskilde er VIAF ikke en entydig kilde men svarer til en international fælleskatalog. I nogle tilfælde kan man kompensere herfor ved at prioritere genbrugskilder, men nedenstående punkter komplicerer alligevel genbruget.
- Skriftsprog og translitteration
- Udbredt udenlandsk praksis omkring udeladelse af artikler. Om dette skal være dansk praksis skal endeligt fastlægges i den danske profil, men da det ikke er specielt brugervenligt skulle argumentet herfor udelukkende være af genbrugshensyn.
- RDA's preference mht. at foretrække tilføjelser mv. på den katalogiserende institutions sprog

MEN: udenlandske autoritetsdata vil i mange tilfælde kunne bruges som udgangspunkt for og inspiration til berigelse af danske data.

Regler

Arbejdet med autoritetsdata i et nationalt repository er ikke bundet til et valg af RDA men kan isoleret set være ønskeligt for tydeligere at illustrere relationer. Man kan godt implementere et dansk autoritetsrepository uden RDA, men man skal i givet fald så selv fastlægge regler for, hvad der skal indgå i beskrivelsen af de enkelte entiteter. Her vil det være naturligt at se på, hvilke elementer RDA arbejder med bl.a. for at kunne genbruge oplysninger.

Dansk hyldeopstilling og standardtitler

Dansk hyldeopstilling bør betragtes som en del af danske autoritetsdata via en særlig markering. Den titel der vælges til opstilling vil i langt de fleste tilfælde alligevel indgå som varianttitel i autoritetsposten.

De hidtidige danske standardtitler på musik bør ligeledes fremgå af autoritetsdata og derved bygge bro imellem hidtidig og fremtidig praksis.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Formatet

Hvad enten implementering vælges inden for et MARC-miljø eller ej, skal der tages højde for at alle relevante elementer, der hidtidig er fremgået af hjælpepostformatet eller som noget nyt fremgår af RDA, kan angives. I analysen er anført samtlige konstaterede mangler indenfor det nuværende format.

Hvis implementeringen sker inden for MARC, skal den hidtidige logik omkring navne helt for sig og titler helt for sig (enten felt 100 eller felt 130 i en post – aldrig begge) løses, så der kan ske den nødvendige kobling af ophav og titler som authorized access points.

Praktisk brug af autoritetsdata

Mange værker udkommer kun en enkelt gang og oversættes formentlig heller aldrig. Der bør tages stilling til, om vi i dansk sammenhæng vil (og kan) oprette værkposter på alt fra starten, hvor vi i hvert fald for de bibliografiske posters vedkommende, befinder os i et MARC-miljø.

Dertil kommer at afgøre præcis, hvordan værk- og expression posters sammenhæng skal illustreres i en dansk kontekst. Skal vi have en værkpost, der linker til sit ophav, en- eller flere expressionposter, der linker til værkposten og ikke gentager oplysninger herfra eller skal etableres flade strukturer med gentagelser af oplysninger.

Svaret på ovenstående hænger meget sammen med de muligheder for rationelle arbejdsgange og overblik, genbrug mv., som de fremtidige værktøjer vil kunne stille til rådighed.

Brugsscenerier omkring samspil mellem bibliografiske poster og entitetsposter og imellem entitetsposter bør beskrives nærmere både helt generelt og også specifikt. (Et eksempel på sidstnævnte: Hvad skal der ske fx når man i en bibliografisk musikpost har en henvisning til en autoritetspost og ønsker standardtitlen fremvist, fordi materialet er af ældre dato og er opstillet herunder?)

Der skal endvidere tages stilling til, hvad der vil være hensigtsmæssigt at udføre retrospektive rettelser på når datamodel og format kendes.

Prioritering af typerne – rækkefølge for implementering mv.

Arbejdsgruppen anbefaler følgende prioritering af typerne:

- 1) Personnavne

Man vil med fordel kunne starte med dette område, da DBC allerede har etableret en pt. intern autoritetsbase dækkende en væsentlig del af det nationalbibliografiske område. En opgave vil bl.a. være at finde en løsning på håndtering af de ikke-nationalbibliografiske navne samt opgradering af et i første omgang udenlandsk navn til dansk nationalbibliografi. Disse spørgsmål forventes besvaret via datamodellen.

- 2) Værktitler på musik.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Så snart vi går i gang med at bruge RDA, vil dette være et presserende problem, så opbygning af repositorium inkl. løsning på dansk opstilling og relation til gamle standardtitler på dette område bør prioriteres højt.

Det bør undersøges, hvor mange danske standardtitler vi egentligt har haft anvendt på musik. Det bør efterfølgende vurderes, om vi skal lave en gennemgang af disse eller som en anden mulighed: starte med at lægge standardtitlerne op som autoriteten og rette til RDA efterhånden, som titlerne genbruges, når der kommer nye materialer. I vurderingen bør også indgå om noget i større omfang kan klares maskinelt fx via match med VIAF og oversættelse til dansk af udvalgte termer.

3) Korporationsnavne

Gerne sideløbende med punkt 2 vil det være relevant at opbygge autoritetsdata på nationalbibliografiske korporationsnavne på linje med personnavne.

4) Øvrige værk- og expressionstitler

BILAG 2.1 : Uddrag fra RDA-rapporten 2015, Kapitel 4 vedr. Autoritetspostformat

4.2 Autoritetspostformat

Først og fremmest vil der være behov for, at der udarbejdes et udvidet gennemarbejdet format til autoritetsposter, som er et essentielt moment i forhold til RDA. Vi har arbejdet med nogle få autoritetsposter i MARC21 format og har overvejet om der i forbindelse med autoritetsposter på personer kan opstå problemer i forhold til den danske persondatalovgivning. Der er tale om meget detaljerede personoplysninger, evt. baseret på internetsøgninger med angivelse af kilde.

Vi har uformelt spurgt en repræsentant fra Datatilsynet ved et-gå-hjem-møde i FABITa herom.

Den umiddelbare udlægning var at hvis en person selv har offentliggjort oplysninger om sig selv fx på en bog må vi gerne gengive det.

4.2.1. Hvor mangler felter/delfelter/koder i hjælpepostformatet, for at det vil kunne benyttes til RDA?

Recording the Status of Identification (RDA 5.7 og 8.10.1.3)

RDA foreskriver, at man angiver status for en autoritetspost.

I Marc21 sker dette i:

008 position 33 (**33 - Level of establishment**)

a - Fully established

b - Memorandum

c - Provisional

d - Preliminary

n - Not applicable

| - No attempt to code)

I danMarc2 Hjælpepostformatet er der dels mulighed for at angive et katalogiseringsniveau i 008 *v, dels at angive, om der er tale om en autoriseret form i 042 *a. Hvis man skal understøtte RDA bør disse to felter nok revideres.

Angivelse af, om en autoritetspost gælder for et work, expression, person, family, corporate body, concept, object, event eller place (de entities, der er definerede i RDA) – i danMarc2 er der følgende muligheder i 004

*x:

Delfelt *x Kode for hjælpeposttype

- Der anvendes følgende koder:
- m Emneordspost
- n Navneformspost (person- eller korporationsnavn)
- t Titelpost

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Special coded dates (*RDA 9.3, 6.10*)

I Marc21 er det muligt at benytte specifikke delfelter til start- og slutdatoer – dette vil det måske også være hensigtsmæssigt at indføre i danMarc2 hjælpepostformatet.

Preferred title og variant title of work og expressions

Der mangler delfelter i felt 100/110/400/410/500/510 til preferred title og variant title of work og expressions, part of work, date of work, language of work, anden identification af expression/work works (*RDA 6.x*)

Der mangler delfelter til **fuller form of name** i 100

Der mangler felt/delfelter til at angive **Content type** til at skelne expressions (*RDA 6.9*)

I Marc21 er der bl.a. defineret følgende specifikke felter til **karakteristika for entiteter**:

- **Felt 370 Associated Place**
- **Felt 371 Address**
- **Felt 372 Field of Activity**
- **Felt 373 Associated Group**
- **Felt 374 Occupation**
- **Felt 375 Gender**
- **Felt 376 Family Information**
- **Felt 377 Associated Language**
- **Felt 380 Form of work**
- **Felt 381 Other Distinguishing Characteristics of Work or Expression**
- **Felt 382 Medium of Performance**
- **Felt 383 Numeric Designation of Musical Work**
- **Felt 384 Key**
- **Felt 667 Nonpublic General Note**
- **Felt 688 Application History Note**

Kommentar 2015

Det bør nævnes, at alle disse problemer er overgangsproblemer. Den dag, der er defineret en afløser for et MARC-format, hvad enten det så hedder BIBFRAME eller noget andet, vil datastrukturer se helt anderledes ud

Bilag 3: Brugsscenarier

Som baggrund for autoritetsdatastrategien er udarbejdet en række brugsscenarier.

Disse illustrerer nogle typiske brugsscenarier for autoritetsdata i biblioteker, men der er ikke tale om en udtømmende liste.

AUTORITETSDATA – BRUGSSCENARIER – OPHAV

Brugsscenarie 1

Eksempel:

En studerende søger overblik over udgivelser af *Søren Kierkegaards* værker på dansk, engelsk og tysk. Han foretager en ordsøgning med tegnsæt i almindelig dansk tastatur. Lænker / relationer skal desuden kunne finde *S. A. Kierkegaard*, *Søren Aabye Kierkegaard* med diverse varianter af ø og aa samt pseudonymerne Johannes Climacus, Victor Eremita etc. (I VIAF ses 10 varianter, i LCs autoritetspost ses 27 varianter, hvoraf dog nok ikke alle er relevante på engelsk og tysk)

Kommentar:

Selv om brugeren kender *Søren Kierkegaards* fulde navn, har han ikke forudsætninger for at vide, at databasen ikke samler alle relevante værker under ét, og slet ikke for at gennemskue, at navnet kan forekomme med forskelligt tegnsæt. Et traditionelt henvisningsnetværk ville sende brugeren videre fra én variant til den næste, men ambitionen må være, at søgesystemet udfører denne funktion bag om ryggen på brugeren.

Krav til funktionalitet:

Autoritetsdata skal gøre det muligt at samle adgangen til værker med forskelligt formulerede variationer af samme navn.

Brugsscenarie 2

Eksempel:

En låner søger i en databrønd på *shostakovich symphony 7*. Han har en forventning om at få både musik og anmeldelser. Bibliotekets katalogiserede materialer er registrerede med navneformen *D. D. Šostakovič*. Men andre biblioteker i databrønden har formen *Dmitrij Sjostakovitj*, og en stor mængde artikelstof er registreret med formen *Shostakovich*.

Kommentar:

Brugeren bør kunne finde både bøger, artikler og musikaler uanset hvilken form, der forekommer i de enkelte materialer.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

I eksemplet ligger også den problemstilling, at katalogiseringsreglerne giver valgfrihed m.h.t. standard for transskribering af kyrillisk tegnsæt, et forhold, som de færreste brugere har forudsætninger for at gennemskue. En fagperson fra musikområdet vil sandsynligvis ikke søge på samme navneform, som en fagperson fra et slavisk institut.

Krav til funktionalitet:

Autoritetsdata skal gøre det muligt at søge, finde og genkende information i den for brugeren kendte form, uanset hvilken form, der optræder i materialet.

Problemstillingerne, der er beskrevet i brugsscenarie 1 og brugsscenarie 2, har principielle funktionalitetskrav fælles, men der er tale om forskelle både i søgesituationen og i datastrukturen. I scenarie 1 kunne der godt være tale om søgning i én database, som ganske vist indeholder data importeret fra forskellige kilder – f. eks. DBF og LC, men lagrede i samme format og system, som den katalogiserende institution må antages at have en vis redaktionel indflydelse på. I scenarie 2 søges i en brønd med sammenbragt dataindhold i flere forskellige formater. Her har katalogisator ingen indflydelse på de forskellige ”klumper” af data, men man kunne godt tænke sig indflydelse på de transaktioner, der forgår bag om ryggen på brugeren. Det er netop her, autoritetsdata kunne være til hjælp.

Brugsscenarie 3

Eksempel:

Bibliotekaren skal katalogisere cd'en *Ave Maria : sange til Jomfru Maria gennem 900 år*, og skal blandt andet lave opslag på Hildegard af Bingen. I Danbib findes 46 hits på *Hildegard af Bingen 1098 1179*, der findes 111 hits på *Hildegard, Saint 1098 1179*, der findes 161 hits på *Hildegard von Bingen*, og der findes 27 på *Hildegardis Bingensis*. Bibliotekaren importerer en post med formen *Hildegard von Bingen*. Det er vigtigt at kunne etablere relationer mellem disse varianter.

Kommentar:

Eksemplet indeholder varianter, der stammer fra poster fra forskellige kilder - DBF, DMO og LC, og poster, hvor navneformen skyldes valgfriheden i katalogiseringsreglerne. I en brugergrænseflade, hvor brugeren navigerer ved hjælp af hyperlinks, vil hver af de ovennævnte former fungere som adskilte enheder, og det vil derfor være nødvendigt at kunne etablere lænker mellem varianterne. Problematikken har meget tilfælles med den, der er beskrevet i brugsscenarie 4.

Krav til funktionalitet:

Det er nødvendigt, at der kan etableres lænker mellem variantformer. Autoritetsdata skal levere materiale til lænkerne.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Brugsscenarie 4

Eksempel:

Katalogisator registrerer både noder og fonogrammer af *Niels W. Gade* samt bøger om denne. I DMF er navneformen: *Gade, Niels W. 1817-90*. I DMO og DBF er navneformen: *Gade, Niels W.*, og i LC er formen *Gade, Niels W – Niels Wilhelm, 1817-1890*. Derudover findes i Danbib også varianterne *Gade, Niels Wilhelm 1817-1890*, og *Gade, Niels Wilhelm f. 1817*.

Kommentar:

Varianterne vil ikke give problemer i ordsøgning, men det er vigtigt at sørge for lænke mellem relationerne. Med traditionel katalogisering ville katalogisator have valgt én navneform fra én af kilderne og så have tilrettet navnet ved import fra de andre kilder. Variationerne er endda så små, at man nok ikke ville overveje henvisninger. Men vi har i efterhånden mange år importeret poster fra nationalbibliografiske kilder uden at redigere yderligere i dem, fordi de hver især repræsenterer en autoritet. Her er det nødvendigt at operere med et entydigt ID og/eller lænker mellem variantformerne. Problematikken har en del tilfælles med brugsscenarie 3, men her er det vigtigere at få udpeget ét ID

Krav til funktionalitet:

Autoritetsdata skal gøre det muligt for katalogisator at benytte data fra forskellige kilder og knytte dem til en entydig identifikation.

Brugsscenarie 5

Eksempel:

Der findes i Danbib 6 indførsler med *Ole Hansen* med forskellige fødselsår – og der findes 809 poster uden fødselsår, hvor der ikke kan skelnes. Der er brug for, at hver forfatter forsynes med entydigt ID, evt. med lænke til værkfortegnelse.

Kommentar:

Det forekommer udsigtsløst at skulle koble ID på alle de indførsler, der deles af flere ophav. Der er dog flere initiativer i gang, når det drejer sig om forskeridentifikation, men de vil i alle tilfælde kun kunne dække en brøkdel af de entiteter, der er registrerede. Men kan forstille sig, at ID på ophav bliver mere udbredt, men der vil altid være et stort "gråt felt". Dog vil valg af en foretrukken form, når flere varianter foreligger, være muligt og i visse tilfælde nødvendigt. For biblioteker, der opstiller efter forfatter/ophav, er det vigtigt at kunne afmærke en foretrukken form.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Krav til funktionalitet:

Der skal være mulighed for tydelig angivelse af, hvilken variant, der foretrækkes. For folkebiblioteker vil den nationalbibliografiske form være gældende, for andre biblioteker skal man kunne ejermærke egen foretrukken form.

Se også navneformerne i brugsscenario 2 og 3. Afmærkning er uomgængelig i tilfælde, hvor navnet benyttes som opstillingselement.

Ved etablering af brugergrænseflade forekommer der problemer i forbindelse med varianter på navne. Hvis posterne ligger i uredigeret form, som de er importerede – og det vil de jo bl.a. gøre for alle de poster, der er modtaget som dataleverancer, vil der følgelig være lige så mange varianter som i eksemplerne ovenfor. Hvis hyperlinks blot etableres i den enkelte post, vil den kun fange poster med samme form. Derfor bør lænken gå til en autoritetspost med alle de former, der er tilgængelige – evt. ved at supplere med varianter fra VIAF.

Brugsscenario 6

Eksempel:

I forbindelse med digitalisering af et værk, ønskes der klarhed over hvorvidt værket er beskyttet af ophavsret. Værkets MARC post har ingen information omkring forfatterens dødsår, som kunne give en indikation om hvorvidt værket er i public domain eller ej.

Kommentar:

ARROW.eu er et EU-projekt som netop har gjort adgangen til søgning efter rettighedshavere lettere. ARROW (Accessible Registries of Rights Information and Orphan Works) benytter sig bl.a. af VIAF og andre autoritetsbaser.

Krav til funktionalitet:

Ved at tilgå autoritetsdata vil det være muligt at trække information omkring rettighedshavers dødsår og dermed få klarhed over status på copyright i forhold til et evt. public domain spørgsmål.

Brugsscenario 7

Eksempel:

Låner er interesseret i Afghanistan og vil gerne have romaner og noveller, som foregår i Afghanistan eller det geografiske område, der i dag er Afghanistan. Hun ønsker sig også faglitteratur om forskellige perioder i landets historie, gerne perioden omkring år 1300. Generelt er det Djengis Khan og mongolernes tid i Afghanistan, der har lånerens interesse.

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Kommentar:

Ved hjælp af de nuværende nationale metadata vil dette informationsbehov kræve megen mental bearbejdning hos personen selv, hvis informationsbehovet skal opfyldes. Lad os fx antage, at Afghanistan i dag anvendes som isoleret term til at repræsentere et objekts geografiske emne, så kan den informationssøgende ikke se, at Afghanistan er et geografisk område etableret i 1700-tallet – og ledes ej heller til andre termer der beskriver samme geografiske område historisk set.

Krav til funktionalitet:

Autoritetsdataene skal her kunne beskrive geografis historie gennem relationer til andre begreber (værdier), der kan beskrive - og derved verificere - hvad det geografiske område Afghanistan skal forstås som. Det kunne fx være:

- Etableringstidspunkt (foundedIn)
- staveformer (inkl. sprogvariationer)
- geografiske koordinater
- nabolande (fx hasGeographicNeighbour "Kina")
- region (fx isPartOfGeographicalRegion "Centralasien")
- geografisk historie (fx HistoricallyAGeographicPartOf Det mongolske kejserherdømme)

Brugsscenario 8

Eksempel:

En låner vil gerne læse romaner der på en eller anden måde handler om migration. Lånerens forældre flyttede fra Marokko til Danmark i starten af 1970'erne, så migrantfortællinger med udgangspunkt i Marokko foretrækkes. Migration, frivillig såvel som ufrivillig - i alle dets aspekter - interesserer brugeren, så hun ønsker også faglitterære materialer, som forklarer og/eller diskuterer aspekter af migration.

Kommentar:

Det er i dag en utopi, at en bruger kan få sit informationsbehov opfyldt gennem en entydig verifikation af relevante materialer om fx emnet migration pga. mange forskellige emnesystemer i den nationale infrastruktur. Og det er af flere årsager urealistisk at forestille sig, at autoritetsdata for emner skal kunne løse diskrepanserne mellem de forskellige emnesystemer. Men brugsscenarierne viser den kompleksitet, som autoritetsdata for emner skal arbejde med, hvad en infrastruktur af national interesse skal imødekomme, hvis emnenavigation skal muliggøres.

Krav til funktionalitet:

Radikalt set er brugsscenario-kravet, at brugeren har behov for hjælp til at afklare informationsbehovet. Det kan antages, at det er det sociologiske fænomen migration, der er den primære interesse, men hvilken type

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

migration? "migration" skal altså beskrives semantisk. Det deraf afledte krav er, at de forskellige emnesystemers betegnelser for samme skal sammenkobles.

Et realistisk kompromis er, at autoritetsdata til emnenavigation kan opnås ved at specificere semantiske relationer i udvalgte emnesystemer af national interesse (fx DK5 og DBC's emneord) og derigennem (i en løbende proces) knyttet relationer til dele af andre emnesystemer.

Bilag 4: Autoritetsdata i dag – status

Situationen angående autoritetsdata i DK kan betegnes som broget. Der har ikke tidligere været nogen fælles politik eller indsats med henblik på at skabe fælles kontrollerede autoritetsdata, primært på grund af manglende ressourcer. Dette har især præget fag- og forskningsbibliotekerne, mens folkebibliotekerne har genbrugt de nationalbibliografiske data, producerede af DBC.

På det nationalbibliografiske område har der været en fælles indsats mellem DBC og Det Kongelige Bibliotek, som dog afspejler de variationer, som er indlejret i katalogiseringsreglerne.

I formatarbejdet under Bibliografisk Råd er udarbejdet et fælles dansk autoritetsformat, som bygger på MARC 21, således at det burde være muligt med dataudveksling på det område, hvis man ønsker det.

I det følgende opridses nogle af de vigtigste problemstillinger i relation til de enkelte datatyper.

Navne på personer og korporationer

Katalogiseringsreglerne i DK anviser i en række tilfælde både regler og alternative regler, hvilket gør det umuligt at tale om én og kun én autoritativ form.

Ved udformning af navne på personer/korporationer gælder f. eks, at der i den nationalbibliografiske registrering på DBC og det Kongelige Bibliotek er foretaget navneformskontrol således at der i den enkelte nationalbibliografi kun er én foretrukken form, men der kan være forskellige valg i Dansk Musik Fortegnelse, Dansk Kort Fortegnelse og Dansk Periodica Fortegnelse.

Det er derfor nødvendigt at opbygge et system, hvor flere fortrukne former kan sameksistere enten i én autoritetspost eller ved hjælp af relationer mellem flere varianter, der kobles til samme ID.

I forbindelse med kontrol af navne bør den øverste autoritet principielt være den nationalbibliografi, der omfatter de originale udgivelser. Således er den danske nationalbibliografi autoritet for danske publikationer.

DBC anvender samme regler/praksis for udformning af navne til nationalbibliografi og til bibliotekskatalogisering, og her drejer det sig ofte om de regler, der i katalogiseringsreglerne er angivet som "alternative".

Folkebibliotekerne følger i alt overvejende grad DBC's praksis.

Fag- og forskningsbibliotekerne følger – især ved udformning af korporationsnavne – ofte en anden praksis (se i denne sammenhæng bort fra, at der i mange tilfælde slet ikke foretages navneformskontrol).

Varianter

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

I tilknytning til hver af de foretrukne former kan der være etableret en eller flere henvisninger fra ikke anvendte former.

Disse henvisninger kan være materialespecifikke, dvs. udarbejdet på baggrund af varianter af navnet, der findes i et givet materiale (også kaldet postspecifikke henvisninger), eller de kan være generelle, dvs. udarbejdet ud fra en generel opfattelse af, hvilke variantformer, brugerne kan tænkes at søge på.

Der kan fra institution til institution være forskel på, hvilket "sæt" af henvisninger, man har knyttet til en foretrukken navneform.

Værk-/titeldata

Katalogiseringsreglernes definition på uniform titel/standardtitel giver tre versioner:

- 1) Standardiseret titel til identificering i kataloger og bibliografier af et værk, der er udkommet under forskellige titler.
- 2) Standardiseret titel, der anvendes til at skelne værker med enslydende bibliografiske titler fra hinanden.
- 3) Standardiseret fællestitel, der anvendes til at samle materiale med flere værker eller uddrag el.lign. af flere værker (f.eks. samlede værker eller værker i en bestemt litterær eller musikalsk genre) af et bestemt ophav

Men da beslutning om anvendelse af uniforme titler og/eller "standardtitler" i henhold til reglerne er overladt til den enkelte katalogiserende institution, er praksis tilsvarende varierende. Uniforme titler på andet end musik anvendes kun i begrænset omfang

Definition 1) anvendes i musikkatalogiseringen til at samle varierende værktitler og værktitler på forskellige sprog, men der er forskellig praksis i hhv. Dansk Musik Fortegnelse og i Danske Musik Optagelser, idet man i DMF benytter uniforme titler, mens man i DMO benytter "standardtitler" som er folkebibliotekernes version af uniformerede titler.

Dansk Periodica Fortegnelse benytter definition 2), hvor den standardiserede titel anvendes til at sondre mellem værker med enslydende bibliografiske titler. Hvis der er flere enslydende titler, tilføjes data (udgave, by, år, medium, udgiver) for at kunne skelne titlerne fra hinanden. Dette gøres både i Key title til ISSN (Qualifying term) og i DPF (Tilføjelser). Disse tilføjelser bliver således brugt til at skelne mellem enslydende titler og ikke til at samle titler.

Selvom praksis for uniforme titler mv. altså varierer, er der ikke som ved navnedata tale om, at der anvendes forskellige valgfrie regler for udformning.

På den baggrund vil det være rimeligt at overveje etablering af fælles autoritetsposter indenfor hver datatype – selvom alle institutioner muligvis ikke vil gøre brug af alle typer.

Emnedata

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

Emnedata er det mest uensartede område, og et fælles nationalt emne-repository vil ikke være et realistisk scenarie.

Til gengæld kan man godt forestille sig, at der opbygges autoritets-repositorier for enkelte emnesystemer af national interesse, som fx DK5 og DBC's emneord .

I forhold til emneområdet vil det største fremadrettede perspektiv dog ligge i mulighederne for at berige autoritetsposter med links, der kan forbinde de forskellige systemer med hinanden, fx i form af relationsangivelser mellem DK5 og Dewey.

Geografiske data/stednavne

Vi har pt. ingen praksis for dette område - udover emnedata

Bilag 5: Oversigt over væsentlige attributter til forskellige autoritetsdatatyper

Nedenfor angives de *væsentligste* attributter, der bør knyttes til hver enkelt type. Udarbejdelsen af en detaljeret beskrivelse inklusiv angivelse af et anbefalet minimumsniveau skal indgå som en del af implementeringen af autoritetsdatastrategien. I samme forbindelse skal også krav i forlængelse af en evt. overgang til RDA indtænkes.

Personer

Væsentlige attributter:

- Vedtaget navn
- Variantformer – herunder varianter i forskellige transskriptioner og alfabeter
- Identifikatorer knyttet til personen (ISNI, ORCID)
- Relationer til biografisk og historisk information mv., fx oplysning om navneskift, fødsels- og dødsår, nationalitet, køn, arbejdsmæssig tilknytning

Korporationer

Væsentlige attributter:

- Vedtaget navn
- Variantformer – herunder varianter i forskellige transskriptioner og alfabeter
- Tidspunkt for konference
- Over- og underordnede korporationer
- Identifikatorer
- Relationer til historisk information mv., fx oplysning om navneskift, start- og slutår

Titler (uniforme titler og "standardtitler")

Væsentlige attributter:

- Vedtaget titel
- Variantformer – herunder varianter i forskellige transskriptioner og alfabeter
- Identifikatorer
- Relaterede titler, fx titler på andre dele af samme værk

Bemærk! Efter en evt. overgang til RDA vil uniform titel og standardtitel ikke længere være relevant, idet disse titler vil indgå som "authorized access point" repræsenterende et "work" eller en "expression".

Work og expression

Work- og expressionentiteter er knyttet til RDA, som dermed vil være bestemmende for hvilke attributter, det skal udpeges her.

Emne ('topic')

Væsentlige attributter:

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

- Vedtaget term/klassifikationskode
- Variantformer – herunder også tekstlig "oversættelse" af klassifikationskode
- Identifikatorer
- Semantiske relationer, fx over- og underordnede begreber
- Mapning til andre vokabularer eller taxonomier

Lokaliteter

Væsentlige attributter:

- Dansk navn for lokalitet
- Variantformer
- Identifikatorer, fx ID fra Geonames
- Landekode(r)

Sprog

Væsentlige attributter:

- Dansk navn for sprog
- Variantformer – herunder varianter i forskellige transskriptioner og alfabeter
- Identifikatorer
- Sprogkode(r)

Materialetyper

Væsentlige attributter:

- Dansk navn for materialetype
- Variantformer – herunder varianter i forskellige transskriptioner og alfabeter
- Identifikator
- Kode for materialebetegnelse(r)

Bilag 6: Ordliste og definitioner ⁱ

| | |
|-------------------------------|---|
| Attributter | <p>Noget som karakteriserer en genstand og som danner basis for beskrivelse og klassifikation af genstanden.</p> <p>I autoritetsdatastrategien anvendt om de egenskaber, der kan knyttes til en entitet, fx navnet på en institution eller en persons fødselsdato.</p> |
| BIBFRAME | <p>BIBFRAME (Bibliographic Framework) er en datamodel for bibliografisk beskrivelse. BIBFRAME, der er tænkt som en afløser for MARC-formatet, bygger på principperne for linked data og har bl.a. som formål at øge nytten af bibliografiske data også udenfor biblioteksdomænet. (http://bibframe.org)</p> |
| BNF | <p>Bibliothèque nationale de France</p> <p>Det franske nationalbibliotek (http://www.bnf.fr/fr/acc/x.accueil.html)</p> |
| Entiteter | <p>Egentlig betegnelse for abstrakte eller konkrete ting, der kan eksistere på ét ganske bestemt tidspunkt (modsat en proces eller begivenhed, som tager tid). I nærværende sammenhæng anvendes begrebet fx om de elementer, hvoraf FRBR er opbygget, dvs. om de enkelte elementer der kan indgå i metadata om et materiale. Det kan fx være oplysninger om materialets ophav, emne, forlagsoplysninger etc. Beskrivelsen af det fysiske eller digitale materiale kan i sig selv være en entitet.</p> |
| Entitets/reasons-model | <p>En tænkning om informationssystemer og databaser baseret på grundbegreberne entiteter og disses egenskaber og relationer.</p> <p>FRBR er et eksempel på en objekt og entitets/reasons-model.</p> |
| FRAD | <p>Functional Requirements for Authority Data</p> <p>En udvidelse af FRBR til at også at omfatte autoritetsdata.</p> |
| FRBR | <p>Functional Requirements for Bibliographic Records</p> <p>Navnet på systemanalytisk model (en objekt og entitets-reasons-model) af funktionelle krav til bibliografiske poster, udarbejdet af en arbejdsgruppe under IFLA og udgivet i 1998. FRBR afløser de principper som blev vedtaget på Paris-konferencen i 1961.</p> |
| GND | <p>Gemeinsame Normdatei (http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html), er en international autoritetsfil skabt i et samarbejde mellem flere biblioteker indenfor det tysktalende område. Det tyske nationalbibliotek, Deutsche Nationalbibliothek, er ansvarlig for driften af GND.</p> |
| ISNI | <p>International Standard Name Identifier</p> <p>En ISO-standard til entydig identifikation af bidragydere til medieindhold såsom bøger, tv-programmer og avisartikler.</p> |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

| | |
|---------------------|---|
| KB (Sverige) | Kungliga Biblioteket (http://www.kb.se). Sveriges nationalbibliotek |
| LAR | Lokale AutoritetsdataRepositorier |
| L-ID | ID for entitet i lokalt repositorium. Skal i sammenhæng med ID for leverandøren være unikt. |
| Linked data | <p>Metode til at strukturere data, så de automatisk kobles sammen og bliver lettere at bruge, fx ved websøgning og til at dele med andre. Metoden bygger på en grundlæggende ide om at beskrive sammenhænge mellem metadata via såkaldte "triples", dvs. 3-lede udsagn der består af objekter og handlinger - også udtrykt som subjekter, objekter og prædikater.</p> <p>Eksempler på tripler: Et værk er skrevet af en forfatter En forfatter har en nationalitet En forfatter er beskrevet i et forfatterportræt</p> <p>I teknikken er indbygget brugen af maskinlæsbare URI'er (URL'er) til identifikation af udsagnets enkelte led.</p> |
| NAR | Det Nationale AutoritetsdataRepositorium (<i>foreløbig arbejdstitel</i>) |
| N-ID | Unikt ID for entitet i det nationale autoritetsdatarepositorium |
| ORCID | Open Researcher and Contributor ID (http://orcid.org) En åben, non-profit, fællesskabsbaseret bestræbelse på at etablere en registrant med unikke personnumre for forskere ("navneformer") samt en transparent måde til at forbinde forskningsprodukter og forskningsprojekter med disse entydige forskeridentifikatorer. |
| RDA | Resource Description and Access Navnet på internationale regler for katalogisering og tilgængelighed, jf. ressourcebeskrivelse. Reglerne skal fra ca. 2010 afløse AACR2 og indarbejder principperne fra FRBR (Functional Requirements for Bibliographic Records) og FRAR (Functional Requirements for Authority). |
| Ressource | Det objekt (person, ting el.lign.) der beskrives i autoritetsentiteten |
| SKOS | Simple Knowledge Organization System (https://www.w3.org/2004/02/skos/) |
| URI | Uniform Resource Identifier En streng af tegn anvendt til at identificere en ressource på internettet. En grundlæggende standard for den semantiske web. Den mest kendte form er URL. |
| URL | [Universal Resource Locator] »Adresse« på website, der er opbygget af flere dele. Et eksempel er: http://www.sa.dk/media(1209,1030)/Arkivh%C3%A5ndbogen.pdf Forrest angives den anvendte kommunikationsprotokol, jf. protokol2 http:// , efterfulgt af navnet på en server www.sa.dk - der er Statens Arkivers World Wide Web-server, jf. domæne2 - hvorefter der følger navn på mappe media(1209,1030) , og endelig filnavn og |

Bilag til National strategi for autoritetsdata - Version 2.0 (2016)

filformat Arkivh%C3%A5ndbogen.pdf for et konkret dokument (Arkivhåndbogen i pdf-format).

VIAF

Virtual International Authority Files (<http://viaf.org>)

Navn på internationalt register over autoriserede navneformer på personer, korporationer, værker mv. Det er et fællesprojekt mellem flere nationalbiblioteker, og det drives af Online Computer Library Center (OCLC). Initiativet udgik fra Deutsche Nationalbibliothek og Library of Congress.

Wikidataⁱⁱ

Wikidata er en kollektivt redigeret vidensbase drevet af Wikimedia-stiftelsen. Hensigten er at levere et fælles grundlag af visse datatyper (fx fødselsdage), som kan anvendes af bl.a. Wikipedia.

Wikidata giver mulighed for, at grundlæggende information om et emnet kan anvendes på tværs af sproglige forskelle på de nationale Wikipedia-sites. Wikidata er også tilgængelig for andre, fx inkluderer VIAF data fra Wikidata i sine autoritetsdataklynger.

ⁱ Hvor intet andet er angivet, tager definitionerne udgangspunkt i Informationsordbogen (<http://www.informationsordbogen.dk/>)

ⁱⁱ Kilde: Wikipedia (<http://da.wikipedia.org/wiki/Wikidata>)