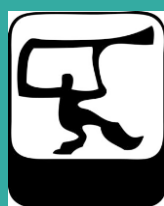




Eksport av data fra fagsystemet FIOLO - klargjøring til nytt fagsystem



NOVEMBER 2020

OPPLANDSARKIVET, NORSK SENTER FOR FOKEMUSIKK OG FOLKEDANS OG KULTURIT

Eksport av data fra fagsystemet FIOL - klargjøring til nytt fagsystem

Opplandsarkivet, Norsk senter for folkemusikk og folkedans og KulturIT

Prosjektet har en klar ramme der målet er å etablere og utvikle en metode for å hente ut, klargjøre, sikre og importere metadata fra FIOL til et nytt fagsystem som bygger på åpen kildekode. Det spesielle i dette arbeidet er at det nye fagsystemet ikke er utviklet enda. Det medfører at sikker lagring av klargjorte metadata er en viktig faktor i prosessen.

Forsidebilde:

Ola Lie Haugen sitter ute på stølsvollen og spiller (Fetstølen) ca. 1940-1945.

Fotograf ukjent / Valdresmusea AS

Mandat, rammer og deltakere i prosjektet

Ramme og rolleavklaringer i prosjektet

Prosjektet har en klar ramme der målet er å etablere og utvikle en metode for å hente ut, klargjøre, sikre og importere metadata fra FIODL til et nytt fagsystem som bygger på åpen kildekode. Det spesielle i dette arbeidet er at det nye fagsystemet ikke er utviklet enda. Det medfører at sikker lagring av klargjorte metadata er en viktig faktor i prosessen.

Ressurspersoner og roller i prosjektet

Arkivarer og fagpersoner:

Elisabeth Kværne, Valdres Folkemusikkarkiv, Valdresmusea, Opplandsarkivet
Frank Rolland, Valdres Folkemusikkarkiv, Valdresmusea, Opplandsarkivet
Stein Villa, Tradisjonsmusikkarkivet Mjøsmuseet, Mjøsmuseet, Opplandsarkivet
Kristin Akslen Hagen, Folkemusikkarkivet for Nord - Gudbrandsdalen, Gudbrandsdalsmusea, Opplandsarkivet
Tom-Willy Rustad, Folkemusikkarkivet for Sør - og - Midt Gudbrandsdalen, Gudbrandsdalsmusea, Opplandsarkivet
Ivar Møgstad, Norsk senter for folkemusikk og folkedans
Tone Erlien, Norsk senter for folkemusikk og folkedans
Silje Cathrin Fylkesnes, Opplandsarkivet (prosjektkoordinator)
Torgeir Forseth, KulturIT (prosjektleder)

Styringsgruppe:

Arne Julsrud Berg, Dir. Mjøsmuseet
Ole Aastad Bråten, Dir. Valdresmusea
Øystein Rudi/Torveig Dahl, Dir. Gudbrandsdalsmusea
Marit Stranden/Ellen Krageberg, Dir. Norsk senter for folkemusikk og folkedans
Bård Bie Larsen/Torgeir Forseth, KulturIT, Produktansvarlig forvaltningssystemer
Silje Cathrin Fylkesnes, Sekr. leder, Opplandsarkivet

Økonomisk ramme:

Støtte fra Riksarkivarens utviklingsmidler:	350 000,-
Opplandsarkivets sekretariat:	30 000,-
Norsk senter for folkemusikk og folkedans	20 000,-
Egeninnsats Opplandsarkivet og	
Norsk senter for folkemusikk og folkedans	90 000,-

Total Ramme: 490 000,-

F a s e 1 : I n n l e d e n d e k a r t l e g g i n g a

Prosjektets praktiske del ble innledet med samtaler med Per Gudmundsson ved Folkmusikens Hus i Sverige om datamodell og datastruktur i FIOl. Her ble det også innhentet noe teknisk dokumentasjon for å kunne lette prosessene med pilot for migrering av data fra FIOl. I møtet med Per Gudmundsson ble også verktøy og teknologi FIOl er utviklet med diskutert og drøftet. KulturIT innhentet tillatelse til å benytte Valdresmusea sin FIOl lisens i testøyemed i forbindelse med eksportprosessen.

Deretter foretok KulturIT utdrag av fullstendige sikkerhetskopier av FIOl databasene for de deltagende institusjonene. Det var også ansett som en viktig del av prosjektet å utbedre og undersøke hvor godt sikret FIOl databasene var i produksjonsmiljøene på de deltagende institusjonene. Denne gjennomgangen satte lyset på at folkemusikk materialet lagret i FIOl i flere tilfeller var lagret på en måte som ikke sørget for sikring av dataene. En del av den tekniske gjennomføringen av prosjektet ble derfor å sørge for at FIOl databasene ble sikkerhetskopiert ved jevnlige intervaller, samt at databaser og mediafiler ble lagret på lagringsløsninger som er vedlikeholdt. En øyeblikkskopi av hver av databasene ligger nå også lagret hos KulturIT. Dette er ikke en løsning som sikrer FIOl databasene for fremtiden, da databasene er i kontinuerlig endring.

Et vedvarende tankekors er likevel lagring av media (film, lyd og bilder) som er lenket til metadata i databasene. Disse digitale filene er ikke kopiert av KulturIT på grunn av størrelse og manglende avtaler rundt lagring. Lagringen av media ble forsøkt utbedret i prosessen, men på grunn mengden data er det her usikkert om alle deltagende institusjoner kan sørge for sikker lagring.

FIOl innehar flere funksjoner for backup av databasen. I dette prosjektet har vi benyttet en funksjon der man kan dumpe databaseinnholdet direkte til tekstfiler. Hver tekstfil representerer innholdet i en tabell i FIOl databasen. Dette har gjort det relativt enkelt, forutsatt at man har oversikt over relasjonene mellom tabellene, å kunne få et raskt overblikk over både innhold og sammenhenger uten bruk av annen programvare enn en tekstbehandler.

KulturIT installerte og konfigurerte et testmiljø for FIOl lokalt. Dette ble brukt til en grundig felt - og funksjonsgjennomgang av FIOl. Dokumentene er vedlagt rapporten som vedlegg 1 og vedlegg 2. En felt - og funksjonsgjennomgang var nødvendig først og fremst for å ha et grunnlag for å kunne uttale seg om potensielle kandidatsystemer for eksport.

Under felt - og funksjonskartleggingen ble nedlastede kopier av FIOl databasene fra de deltagende institusjonene brukt for å få en grunnleggende oversikt over institusjonenes data, samt hvilke forskjeller det var i katalogiseringen i de forskjellige databasene.

Samtaler med NRK og Nasjonalbiblioteket

Det var et mål i prosjektet å vurdere om man kunne dra nytte av NRK og Nasjonalbibliotekets pågående arbeid med å utvikle et nytt fagsystem primært for egne folkemusikksamlinger. Tidlig i prosjektet var det derfor et håp om å kunne kopiere NB's metodikk for utdrag av kunne eksistere i et nytt fagsystem utviklet av NB/NRK. Prosjektet brukte derfor tid på samtaler og utveksling av erfaringsgrunnlag og metodikk med Richard Gjems hos Nasjonalbiblioteket. I korte trekk var dette tilbakemeldingene fra Nasjonalbiblioteket fra møte 15.08.2018.

- Prosjektet er ca. et halvt år forsinket.
- Man har eksportert NB og NRKs FIOl databaser til RDF format. Andre databaser som også er brukt til katalogisering av folkemusikk er blitt eksportert.
- Ny programvare som skal ivareta NRK og NBs behov for et folkemusikkarkiv er under utvikling. «Versjon 1» er ute i test, men gamle system brukes ved siden fortsatt. Målet er at neste versjon skal være full produksjon, og at gamle system inkl. FIOl pensjoneres.
- Man har tatt utgangspunkt i NRKs datamodell for den nye løsningen.
- Det er tidligere sagt at det er ønskelig at folkemusikkarkiver fra andre institusjoner kan bruke løsningen, men infrastruktur og programvare for dette er ikke klart. Dette må utvikles, og man må ha et mandat for å kunne begynne å realisere dette.
- FIOl databasene hos NRK og NB hadde såpass forskjellig oppbygning at det ikke lot seg gjøre å lage et felles eksportskript, som senere kunne brukes på andre databaser. Det har i stor grad vært snakk om manuelt arbeid for å eksportere databasene.

KulturIT mottok mappingdokumenter som var blitt brukt for å importere FIOl materiale til det nye fagsystemet. Datamodell for NRKs fagsystem og det tilhørende mappingdokumentet fra Nasjonalbiblioteket fungerte senere i prosjektet som en referanse for KulturITs egen eksport/import pilot for institusjonene i prosjektet. Vedlagt i rapporten er mappingdokument fra Nasjonalbibliotekets eksportprosjekt (vedlegg 3).

Funn i innledende kartlegging:

- FIOl databaser og media var ikke i alle tilfeller forsvarlig lagret
- FIOl programvaren benytter seg av teknologi, programmeringsspråk og komponenter det ikke er store fagmiljøer på i Norge. Spesielt gjelder dette 4D programvaren ([https://en.wikipedia.org/wiki/4th_Dimension_\(software\)](https://en.wikipedia.org/wiki/4th_Dimension_(software))) FIOl benytter som rammeverk.
- Institusjonene har ikke alltid den tekniske kompetansen til å vedlikeholde FIOl programvaren.
- FIOl har innebygde funksjoner for enkel eksport av tabellene i relasjonsdatabasen som ligger til grunn for katalogiseringen.
- Nasjonalbibliotekets erfaring med eksport fra FIOl til NRKs fagsystem har begrenset verdi for vårt prosjekt, da tidsplan og fremtid er usikker for Nasjonalbibliotekets prosjekt med NRK.
- Skript og mappingdokument fra Nasjonalbiblioteket kan i beste fall brukes som en referanse for prosjektet, da mye av arbeidet har blitt gjort manuelt. **G r u n n l a g e t h e r v a r a t i** databaser var grunnleggende forskjellig. Noe som med stor sannsynlighet også vil være tilfellet for museenes FIOl databaser, sammenlignet med NRKs valg av løsning.

Fase 2: Demonstrasjon av uttrekk fra FIOl og import til eksternt system

I begynnelsen hadde prosjektet en forventning om å kunne dra nytte av Nasjonalbiblioteket og NRKs arbeid med eksport fra FIOl, og at man kunne benytte seg av pilotering i nettopp dette systemet.

Ut fra funnene i fase 1 i prosjektet ble det klart at man innenfor prosjektets rammer og tidshorisont vanskelig kunne forvente å sette i stand en slik eksport og import. Siden en demonstrasjon av eksport av metadata fra FIOl ikke i seg selv ville gi fagmiljøene noen stor overføringsverdi til fremtidige prosjekter, besluttet prosjektet å se etter andre alternativer for import og tilgjengeliggjøring av materialet i FIOl databasene.

Et av de viktigste kravene i prosjektet er at resultatet skal føre til overføringsverdi for folkemusikkarkiv og programvareutviklere. Det ble derfor viktig å finne en kandidat for tilgjengeliggjøring som helst baserte seg på lenkede åpne data (https://no.wikipedia.org/wiki/Lenkede_%C3%A5pne_data), samt sørget for forsvarlig forvaltning av materialet og minimal drift og vedlikeholdskostnad. Det ble på grunnlag av dette satt i gang et arbeid med utvelgelse av eksisterende forvaltningsløsninger som oppfylte kravet om overføringsverdi, forsvarlig lagring og kostnad.

I korte trekk ble utvelgelsesarbeidet fokusert mot hyllevare «DAMS» systemer, samt åpne systemer for forvaltning av autoriteter og termer. Arbeidet ble avsluttet med en beslutning om å gjennomføre en pilot for import av utvalgte FIOL data til KulturNav (<https://kulturnav.org>). Grunnlaget for beslutningen var i hovedsak basert på:

- Enkelhet for fremtidige systemutviklere å gjenbruke de eksporterte dataene vha. åpne API.
- Kostnad: KulturNav er et av få norske forvaltningssystemer som er gratis i bruk.
- Åpenhet. KulturNav er utviklet med åpen kildekode, og er åpen for bruk av alle. Se for øvrig <https://kulturnav.org/info/more>.
- KulturNav er en programvare som kan benyttes av hele ABM sektoren og data om folkemusikk og folkedans vil kunne ha stor overføringsverdi for feltet.
- Fremtidig og langsiktig forvaltning. KulturNav er støttet og utviklet i samarbeid med Kulturrådet, noe som ble ansett som en fordel i forhold til rene kommersielle aktører.
- Umiddelbar formidlingseffekt for enkelte av metadataene, som for eksempel i form av autoritetsregister for musikere, tekstforfattere og dansere innen folkemusikk.

Med bakgrunn i dette utarbeidet KulturIT et mappingdokument for flere «proof of concept» datasett i KulturNav med data importert fra Valdresmuseas database. Mappingdokumentet tok innholdet i FIOL-tabellene District, Identity, Group, Identity_district, Instrument, Role, Region, Role_type og Type og mappet dette til felter for poster i datasett i KulturNav. KulturNav er i seg selv ikke en kandidat for et fagsystem for forvaltning av opptak/media og dets metadata, men er en svært god kandidat for å forvalte deler av FIOL dataene også i fremtiden. Dataene som ble valgt ut for pilotprosjektet er spesielt godt egnet for KulturNav, da de er i form av autoritetsdata.

Selve prosessen med eksport og import fra FIOL til KulturNav er dokumentert og det er lagt opp til at man på en relativt enkel og kostnadseffektiv måte kan gjennomføre den samme prosessen på andre FIOL databaser i fremtiden.

Vedlagt med rapporten ligger mappingdokument for eksport, samt metodikk for eksportprosessen (vedlegg 4 og vedlegg 5).

Fase 3: Arbeid med FIOL Datasettene

Pilotprosjektet med import av data fra Valdres FIOL database til KulturNav resulterte i flere datasett som kunne redigeres og gjennomgås av prosjektdeltagerne på <https://beta.kulturnav.org>. Prosjektdeltagerne utførte en kvalitetssjekk av både presentasjon av dataene, samt eventuelle feilkoblinger i mappingen av data. Disse ble utbedret i mappingscript. Deretter ble det gjort klart for eksport av de resterende tabellene fra de utestående databasene (Mjøsmuseet, Folkemusikkarkivet for Nord-Gudbrandsdalen og Sør- og Midt - Gudbrandsdalen, Gudbrandsdalsmusea og Norsk senter for folkemusikk og folkedans).

For å gjøre folkemusikkarkivarene kjent med programvaren KulturNav, samt diskutere faglige spørsmål rundt publisering og etterarbeid med FIOL dataene ble det arrangert flere samarbeidsmøter, samt Workshop for prosjektdeltagere den 30. September 2019. Workshopen konsentrerte seg om administrasjon og redigering av KulturNav poster ispedd diskusjon og avklaringer rundt aktuelle temaer.

For å illustrere prosessen rundt konvertering av FIOL data til et annet system legges saklisten for workshop med i noe redigert tilstand. Alle punktene i listen har senere blitt avklart, og diskuteres videre i avsnittet «Ferdigstilling, publisering og etterarbeid» (vedlegg 6).

Det ble satt et mål om følgende innen januar 2020:

1: Ha datasett for personer innen folkemusikk dans (vil ikke inneholde alle oppføringene fra FIOL) publisert i KulturNav med minimum fornavn, etternavn, livsrolle. Sett må ha eier og redaksjon.

2: Ha datasett for teknikk/medium, typologi, organisasjoner (grupper/band/lag), rolle publisert med eier og redaksjon i KulturNav.

Ferdigstilling, publisering og etterarbeid

Avklaringspunkter rundt presentasjon av data:

For å kunne publisere de utvalgte FIOl dataene i KulturNav måtte prosjektgruppa ta beslutninger rundt krav til kvalitet, personinformasjon, ordning og duplikater i dataene som skulle publiseres via KulturNav. Under er en liste av avklaringspunkter som ble tatt opp og avklart i fasen med etterarbeid.

Kvalitet: Hva er minstekravet til en personpost eller organisasjonspost (gruppe/band/forening) for at denne skal ha nytte og kunne identifiseres unikt av brukere av et framtidig fagsystem eller besøkende i KulturNav? Hvilke knytninger kan vi legge på personposten som gir mer informasjon om utøverens rolle og virkeområde?

Personinformasjon: Hvilken personinformasjon kan publiseres i KulturNav? Hvilke deler av personinformasjonen i FIOl må utelates fra publisering?

Ordning: Hvordan skal postene i KulturNav kunne spores til opprinnelig arkiv? Skal man beholde ordningen med tanke på hierarki i datasettet for FIOl roller, FIOl typologi og FIOl instrumenter, eller skal man forbedre ordningen?

Duplikater: Dataene som skulle publiseres i KulturNav inneholder et uttrekk fra flere databaser. Man ender derfor opp med en rekke duplikater som ikke alle maskinelt kan sorteres ut. Man må ha rettesnorer for hvilke personposter og gruppe/band poster som er duplikater, og disse må manuelt gås gjennom av arkivarene.

Bestemmelser og publisering:

Arbeidet med avklaringer ovenfor resulterte i følgende publiserte datasett i KulturNav. Disse er samlet under prosjektet Folkemusikk og folkedans:

https://kulturnav.org/project/folkemusikk_folkedans

Datasettet Personer: <https://kulturnav.org/5780336d-3b8e-4a5a-8b1c-25dc9571a1c2>

Det ble besluttet at en god personpost som kunne publiseres minimum måtte inneholde fornavn, etternavn, fødselsdato, virksomhet, emne, kjønn, virksomhet og emne. Et stort utvalg av postene inneholder også ytterligere informasjon.

I tillegg ble det besluttet å krysskoble personpostene mot rolle-datasettet, samt å presentere lenken mot virkested/tilholdssted som er registrert i FIOl. Dette for å sørge for ytterligere søkeinn ganger og sorteringsmuligheter for forskere, arkivarer og interessenter i framtiden.

Det ble tidlig løftet at noe av informasjonen i datasettet hadde karakter av personlig opplysninger, som ikke nødvendigvis ville gi stor verdi til autoritetsregisteret. Disse typene «ekstrainformasjon» ble derfor skjult fra publiserte poster, men er tilgjengelig i KulturNav for datasettets redaksjon. Det ble også drøftet om nålevende personer, som ikke er kjent utenfor en avgrenset del av folkemusikken/dansen skulle publiseres. Foreløpig er det ikke tatt en beslutning på dette. Disse personpostene er altså ikke publisert, men ligger tilgjengelig for redaksjonen.

I forbindelse med publisering av personposter fra FIOl databasene har det blitt lagt ned et større arbeid med manuell utplukking av duplikatposter fra de forskjellige databasene i tillegg til maskinell utplukking. Arbeidet med personposter er ikke ferdig. Det ligger i dag 14929 personposter fra FIOl i KulturNav. Av disse er kun 901 publisert for publikum. Alle er tilgjengelig for forskning.

Organisasjoner: <https://kulturnav.org/db709e06-d0c3-4903-bd88-aef485e3007d>

Et samlet uttrekk av alle organisasjoner fra FIOL databasene ble gjennomgått, der duplikater har blitt fjernet. Først maskinelt, og deretter gjennomgått av arkivarene. Det ble ikke lagt opp til krysskoblinger mot personer i FIOL, da datagrunnlaget var ansett som for dårlig.

Roller: (<https://kulturnav.org/2aa7c2c9-c309-4a0a-bcd6-669535166660>)

Typologi: <https://kulturnav.org/48986163-39d9-4ed3-9e51-5b636f61e792>

For datasettet roller ble det bestemt å beholde ordningen fra FIOL og publisere i KulturNav som den var i FIOL. For typologien rundt folkemusikk og folkedans hadde gruppen er diskusjon rundt å lage to grener (musikk/dans) som i utgangspunktet hadde de samme kategoriene, siden likhetstrekkene for kategoriseringen av de to er flere enn ulikhetene. Det ble likevel ikke foretatt noen endring her på grunn av størrelsen på jobben med ny mapping av FIOL personer til nytt rolle-hierarki, samt manglende diskusjon rundt dette i fagmiljøene i Norge.

Musikkinstrument: <https://beta.kulturnav.org/e0a3622b-74ab-423d-93b7-4be5815902d8> (Foreløpig bare beta)

Det ble besluttet å gjøre et utvidet arbeid rundt eksport av musikkinstrumentkatalogen fra FIOL. Denne ble ansett å ha spesielt stor verdi for fremtidige forvaltningssystem for folkemusikk, men også potensielt for en rekke fagområder utenfor folkemusikk genren. Det finnes flere internasjonalt anerkjente åpne musikkinstrumentkataloger tilgjengelig på internett. En av disse er MIMO (<https://mimo-international.com/MIMO/instrument-families.aspx>). Nyten av å ha en komplett internasjonal termliste for instrumenter som inkluderer mapping mot instrument fra FIOL ble ansett som stor. MIMO listen vil derfor publiseres i KN som et samarbeidsprosjekt mellom FIOL prosjektet og Ringve musikkmuseum. Ringve arbeider, mens vi skriver rapporten med oversettelse av MIMO til norsk. Alle henvisninger fra person i FIOL vil likevel bli beholdt da MIMO listen vil ha pekere fra «MIMO-instrument» til «FIOL-instrument». Denne jobben er allerede utført, men listen er ikke publisert på grunn av manglende oversettelse. Målet er at instrumentlisten publiseres ila. 2020-2021.

Publisitet og utadrettet virksomhet rundt prosjektet

14.05.2019: Riksarkivet. Innlegg i forbindelse med «Erfaringsseminar om bruk av arkivmidler»

23.10.2019: eKultur seminar Sandnes. Innlegg om prosjektet for museer som benytter KulturNav.

10.03.2020: Arkivseminar, Norsk senter for folkemusikk og dans.

Erfaringer og resultater

Erfaringer:

- Eksport av data fra FIOL i form av tekstbackup er relativt enkelt, og kan gjennomføres uten store forkunnskaper. Dette letter en eventuell videre prosess med utdrag fra framtidige databaser.
- Folkemusikkarkivarenes tid med å jobbe med FIOL dataene (etterarbeid) er liten. Arkivarene bør ha et tydelig mandat og avsatt tid til prosjektarbeid.
- Enkelte av folkemusikkarkivene hadde ikke tilfredsstillende lagring og sikring av FIOL databasene.
- Det kreves mer arbeid med personpostene for folkemusikk og folkedans. Her kan arkivarene som er redaktører for datasettet samarbeide med folkemusikkmiljøet i Norge med å publisere flere av de importerte postene. Det finnes også funksjonalitet og muligheter i KulturNav som gir både privatpersoner og ansatte i andre arkiv anledning til å bidra aktivt.

Resultater:

- Det har blitt utviklet en metodikk for eksport av personer, roller, steder, instrumenter, typologi og grupper/foreninger fra FIOL til KulturNav.

-
- De eksporterte dataene fra FIOl er tilgjengelig både i menneske-lesbart og maskinlesbart format i form av nettsiden <https://kulturnav.org> og KulturNavs API (json objekter).
 - Det vil være en enkel oppgave for programvareutviklere å benytte seg av FIOl dataene i KulturNav i et framtidig fagsystem for folkemusikk og dans.
 - Det har blitt opprettet en egen ansvarlig redaksjon for forvaltning av felles norske autoriteter innen folkemusikk og dans.
 - Lagring og sikring av eksisterende FIOl databaser hos prosjektdeltagere er bedret.

Anbefalinger:

- Det bør settes av midler til å vedlikeholde og arbeide aktivt med spesielt datasettet «Personer folkemusikk og folkedans». Dette datasettet er unikt i Norge, og inneholder informasjon som antas å være interessant både for forskning og publikum.
- Ved overgang til et framtidig fagsystem bør autoritetslister fra KulturNav benyttes for kobling av metainformasjon til objektposter som «opptak». På denne måten kan informasjon fra alle objektpostene samles på ett sted, og gi ytterligere søkeinnnganger i formidlingsløsninger.

Gjøvik /Lillehammer, november 2020

Torgeir Forseth
Produktansvarlig forvaltningssystemer, KulturIT
Prosjektleder

Silje Cathrin Fylkesnes
Skr. leder, Opplandsarkivet
Prosjektkoordinator