

Produktkategori - Standarder:

Dato: 29.10.2020 kl: 13.00-15.00

Agenda (med tentative tidsangivelser)

- 13:00 – 13:10 Introduksjon og status for ADDML
- 13:10 – 14:00 Anna Malmø-Lund og Petr Vasilev presenterer arbeidet med Metadata i Arkiv og POC med Felles Datakatalog (40 min presentasjon – 10 min diskusjon)
- 14:00 – 14:20 Ole Aldric presenterer LaTeX som potensiell fremtidig standard for lagring (10 min presentasjon + 10 min diskusjon)
- 14.20 – 14.30 Filformater: Erik Aaberg gir en Oppdatering/status – endring av Riksarkivarens forskrift.
- 14:30 – 14:50 Dialog rundt SIARD i DIAS-pakker. Er det behov for å ha en felles struktur? Om så, hva er hensiktsmessig strukturering? Hvilken informasjon skal tas med for å supplere SIARD? (20 min diskusjon)
- 14:50 – 15:00 Evt

Referat fra møtet

Antall deltakere: 34

Presentasjonene som har blitt holdt ligger ved referatet som vedlegg.

Ane Hovdan presenterte status for ADDML fra Arkivverket. Det har ikke skjedd noe nytt på dette området den siste tiden. Arkivverket forholder seg fremdeles til ADDML, men det har ikke vært fokus på å oppdatere standarden eller tilhørende verktøy. Samtidig utforskes nye måter å ta imot arkiv og beskrive metadata. Målet er å utvide antall standarder som kan brukes til å beskrive ulike metadata vi har behov for, og her tror man ADDML vil være en av dem som kan benyttes.

- Det ble i etterkant stilt spørsmål til hvor i riksarkivarens forskrift det stilles krav om at man må bruke ADDML for strukturbeskrivelse.

Anna Malmø-Lund og Petr Vasilev presenterte arbeidet med Metadata for arkiv og POC med felles datakatalog, et prosjekt som foregår i Arkivverket. Det har til nå blitt identifisert et utvalg metadata som dekker behovene til langtidslagring og tilgjengeliggjøring på datasett- og hovedobjektnivå. Disse har blitt testet ut i POC og man ser at det korresponderer bra på datasettnivå, men har mangler på hovedobjektnivå. Arbeidet har så langt vært internt, men skal nå utvides med eksterne.

- Det ble spurt om metadataene som er identifisert burde kategoriseres som strukturelle, administrative og deskriptive.
 - Det ble foreslått å bruke entiteter på samme måte som i produksjonslinjemetoden. Ønskelig med en forvaltning og utvidelse av disse i regi av dette arbeidet
 - Prosjektet sier at det er minimumsbehov som er identifisert i prosjektet og understreker at dette er første skritt på veien
-

Ole Aldric ga en introduksjon til hvite (tekstbaserte), brune (“åpne og dokumenterte”) og sorte (proprietære) formater, og presenterte et forslag til hvordan LaTeX kan få en mer sentral rolle i lagring i fremtiden som et hvitt format.

- Ole mente at hvite formater er å foretrekke. Mange brune formater er også tillatt i henhold til Riksarkivarene forskrift. Svarte ønskes ikke med mindre det er strengt nødvendig.
- LaTeX er definitivt et godt format (en god syntaks) - spesielt for en del tekniske beskrivelser
- LaTeX kan antagelig godkjennes som format ihht riksarkivarens forskrift som "Ren tekst" Det er mulig at LaTeX bør nevnes eksplisitt i neste versjon av forskriften.
- Det ble nevnt at dette på sikt kunne være interessant som substitutt for enkelte brune formater, men at veien dit vil være lang. Dette skyldes både kompetanse og tyngden av aktuelle formaters utbredelse.

Erik Aaberg presenterte status om filformater fra Arkivverket

Viktigste poeng: Det arbeides med ny versjon av Riksarkivarens forskrift – der bare generelle kriteriene vil ligge i selve forskriften, men detaljer (nåværende § 5-17) vil bli skilt ut som anbefalinger/påbud - og offentliggjort på Arkivverkets nettsider. Tidsplan: På høring fom. Tidlig i 2021, ferdigstilt og publisert 1. juli 2021. Men innspill mottas når som helst - også i år.

- **Det er et ønske om strukturert (maskinlesbar) oversikt over gyldige formater som kan brukes til maskinell behandling/parsing, inklusive PRONOM formatkoder.**
Vi antar at vi tar dette til følge.
- Ole mente at de beste formatene for lyd er FLAC og OGG, og at MKV var å foretrekke som container format. Det ble også en diskusjon om “codecs”, dvs. Den kodede audiovisuelle “innmaten” i containeren. Det vil komme endringer (tillegg) her i neste versjon av forskriften.
- Det ble stilt spørsmål om arkivformatene har vært sett i sammenheng med "Forskrift om IT-standarder i offentlig forvaltning"? <https://lovdata.no/dokument/SF/forskrift/2013-04-05-959>.
Vi som ser på dette har ikke gått inn i denne forskriften, men vil se på det.
- **Ett** nytt format – JSON – er allerede besluttet at kommer med i Riksarkivarens forskrift. Det inkluderer varianter som GeoJSON. Muligens vil vi også da kreve et medfølgende JSON Schema, slik som vi nå krever en XML-mal (.xsd) for XML-filer.

Se ellers den distribuerte presentasjon.

Det ble åpnet for dialog rundt SIARD i DIAS-pakker. Er det behov for å ha en felles struktur? Om så, hva er hensiktsmessig strukturering? Hvilken informasjon skal tas med for å supplere SIARD? (20 min diskusjon)

- Det var lite tid igjen til dette punktet, men det kom frem at det er behov for felles struktur og ønskelig med en workshop for å følge opp. En workshop bør ikke gape over for mye
- Det ble delt tanker rundt dette for Noark 5, SIARD og ADDML:
https://github.com/IKAMR/digitalt-depot-prosesser/blob/master/depot/DIAS-struktur_v1.1_IKAMR_2020-07-07.pdf

- Det foreslås at views kan legges i en "dependencies" mappe
-

Eventuelt

Det var ikke tid til dette punktet.