Filformater

Produktmøte 2019.09.12

Referat fra møtet:

*1. Velkomst og presentasjon av produktet:*

* Kort intro og presentasjon av
* Presentasjonen blir sendt ut til interessenter sammen med referatet når dette er ferdigstilt.

*2. Innspill, behov og ønsker fra deltakere:*

* Om ny forskrift:
  + Stor forbedring, men skulle ønske forskriften ikke sa konkret hvilke filformater som var tillatt, men at det i lovverket ble vist til en portal for gjeldende standarder som kunne oppdateres ved behov av Arkivverket. Enighet om dette fra flere, både produktsjef og andre deltakere, men en slik løsning har vi foreløpig dessverre ikke fått juridisk gjennomslag for.
  + Savner SIARD nevnt som filformat. Det er nevnt spesifikt i forskriften som «database dump», men ikke som filformat.
  + Samme gjelder Zip-filer og .tar filer. Disse er nevnt og godkjent med noen unntak og forbehold om at filene inni Zip/tar er godkjente i forskriften.
  + Det kan nevnes at Forskriften kun er rådgivende for ikke-statlige institusjoner
* Kriterier for valg av «arkivverdige» dokumentformater ble gjennomgått.
  + For detaljer – se vedlagt Powerpoint
* Kategorier av dokumentformater - - -:
  + - - ble kort gjennomgått, der endring/liberalisering fra 1/1-2018 ble nevnt,  
    spesielt nyere Office formater, GML for kart, IFC for 3-D data, MPEG-4 for video  
    og WARC for HTML sider. For detaljer/referanser henvises til Forskriften.
* Om Pronom:
  + Greit å bruke Pronom koder i Noark5 metadata, men da må også Arkade 5 være justert i henhold til dette. Det løser faktisk et problem som var tidligere i Arkade med formater som ikke var dekket. Arkade bør på sikt også kunne klassifisere filer basert på Pronom koder.
  + Torbjørn Aasen har jobbet mye med Pronom koder, blant annet i KDRS sammenheng. Han påpekte at det er flere måter å kategorisere filer på: filendelser, Mime typer o.a. Men PRONOMs sett av koder er mer granulert enn andre varianter. Torbjørn og Erik Aa ble enige om et felles prosjekt for å «mappe» koder «begge veier» for de forskjellige varianter av kategorisering.
  + Alle formater i Forskriften er nå dekket av PRONOM koder. Det siste som kom inn var det spesifikt norske format for kart: SOSI. «Mapping» mellom Forskriften og PRONOM koder vil senere bli publisert på nett.
* Testing og konvertering av forskjellige filtyper
  + Behov for anbefalinger fra Arkivverket om hvordan dette gjøres
  + Behov for tester for alle typer/kategorier av filer
* Filnavn: Behov for at kjøreregler om filnavn
  + Nå lager dårlige filnavn problemer i testing som gjerne havarerer der filnavn er lange eller har spesialtegn i seg.
  + Behov for at regler for dette utarbeides og kommuniseres bredt fra Arkivverket
* Tolkning av regelverk:
  + Det ble fortalt om eksempler der regelverket rundt filformater har blitt tolket for strengt, spesielt når det gjaldt den forrige forskriften.
  + Det er fortsatt mulig å søke om unntak fra Forskriften. Viktigst i en slik vurdering vil være i hvilken grad formatet/formatene tilfredsstiller de krav til «Lang tids forståelse av innhold» som skissert i møtet. Informasjonstap ved evt. konvertering tas også hensyn til.
  + Ny forskrift er mye bedre enn den gamle, men den må også kommuniseres slik at tolkningen følger etter der arkiv blir til! Og forhåpentligvis går det ikke 10 år til neste revisjon.