

Program til udarbejdelse af systemuafhængig arkiveringsversion af administrationssystemer (ESDH-systemer¹, E-journaler i henhold til Statens Arkivers regelsæt)

Objektorienteret programmering foråret 2002. Gruppe: Claus Balslev, Tommy Balle

I forbindelse med indførelse af den digitale forvaltning inden for den offentlige sektor er der opstået et behov for at sikre, at arkivalier kan bevares i elektronisk form. Statens arkiver har i den forbindelse udarbejdet et regelsæt, som senest er ændret pr. 8 marts 2002.

Regelsættet går i grundtræk ud på, at alle databærende filer skal afleveres som "flade filer"², samt at informationerne indlejres i en metadata fil som et XML dokument. Derudover skal den elektroniske arkiveringsversion, som afleveres på CD-R, indeholde en række bestemte hjælpefiler/tabeller, som er med til at dokumentere afleveringen.

Det program som ønskes fremstillet, skal således være i stand til udfra "csv"³ baserede datafiler at omdanne disse filer til det gyldige dataformat jf. omtalte bestemmelser, samt udfra de informationer

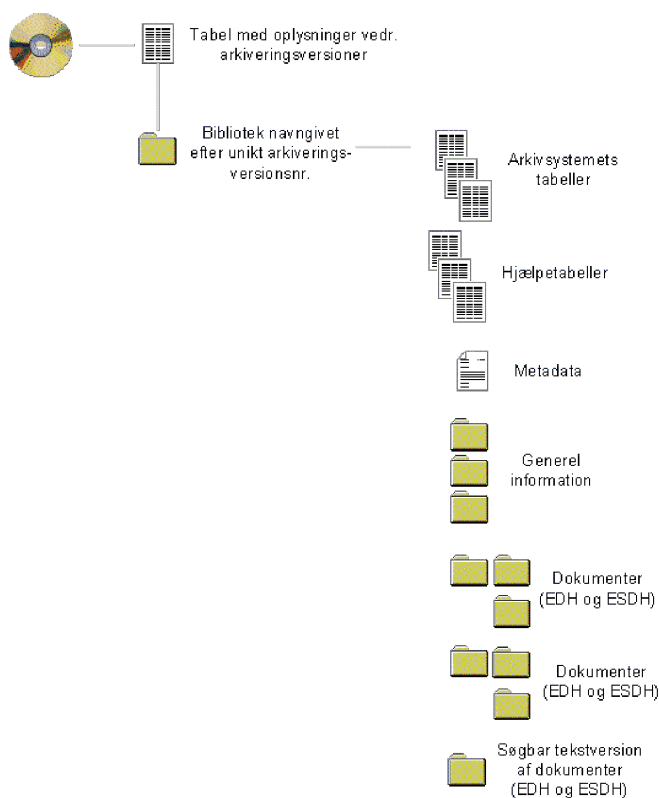


Illustration af systemuafhængig afleveringsversion af et ESDH-system

der medfølger, at oprette de tabeller som skal indgå i afleveringen, herunder automatisk oprettelse og indsættelse af data i metadatatokumentet.

Det forudsættes at skemaet til databasens relationer og entiteter er kendte, samt at domænerne struktur på tabellernes tupler, med tilhørende attributnavne, er givet på forhånd. Disse oplysninger indtastes i en eller flere grænseflader af bruger. (Hvor mange grænseflader der vil være behov for er endnu ikke klarlagt)

Programmet, efterfølgende kaldet for afleveringsværktøjet, er tænkt udført i et objektorienteret design. Således at funktioner/ klasser kan genbruges i flere forskellige sammenhæng i programmet.

Afleveringsværktøjet skal desuden kunne kontrollere filernes dataformater således at de overholder standarder og krav som Statens Arkiver foreskriver⁴.

Vores omdrejningspunkt bliver dannelsen af de filer som i illustrationen benævnes: "a: Tabel med oplysninger vedr. arkiveringsversioner, b: Arkivsystemets tabeller, c: Hjælpetabeller, d: Metadata".

¹ Elektronisk Sags og Dokument Håndteringssystem. Softwarepakke til fuldt elektronisk sagsbehandling inden for det offentlige område.

² "Flat file", datafil som består af en fortløbende struktureret tekststreng uden lineskift.

³ Datafiler, hvor strukturen i tuplerne er fastlagt udfra en separering af attributterne med eks. komma eller semikolon.

⁴ Statens Arkivers cirkulære nr. 25 af 8. marts 2002. <http://www.sa.dk/sa/statamtkom/arklov/nyaflicirk.htm#b2>

Afleveringsværktøjet skal være i stand til at generere følgende tabeller:

a.) Tabel som indeholder oplysninger om arkiveringsversionen, disse oplysninger er givet på forhånd og skal efterfølgende gemmes i et bestemt format i en "flad fil". Skemaet skal bestå af følgende 7 kolonner, tallene i parentes angiver længden på attributten:

ARKVER.TAB (ArkivID(8), AflType(1), MedieID(8), TidligereAfl(8), Systemnavn (256), Startdato(8), Slutdato(8))

Derudover skal det sikres, at de enkelte domæner i tabellens tupper overholder de konventioner, som findes i regelsættet. Eks. skal datoformat være opmærket på følgende måde: CCYYMMDD (hvor CC = århundrede, YY = år, MM = måned og DD = dag)□

b.) Arkivsystemets tabeller er udtræk af tabeller fra en database, som udgør arkiveringsversionens dataindhold. Tabellerne leveres med oplysninger om indhold i skema for de enkelte tabeller. Tabellerne afleveres i csv format. Slutproduktet skal være en "flad fil", hvor tekstindholdet er struktureret efter oplysningerne i det medfølgende skema, og skilletegn er fjernet. Derudover skal det sikres, at filernes tegnsæt overholder de krav, som stilles i cirkulæret. (Udsnit af tegnsæt DS/ISO 8859-1:1987 (Latin 1)).

Som udgangspunkt er de csv baserede filer afleveret således, at der er indsat blanktegn, der hvor tuppernes domæner ikke følger skemaets angivelse af domænelængde. Som en ekstrafunktion vil det være ønskeligt, at systemet udvides med en funktion der muliggør en sådan indsættelse af blanktegn i tabeller som ikke overholder dette.

c.) Hjælpetabeller er tabeller som på samme måde som beskrevet under punkt a. Der kontrolleres på samme måde som i a.

FILMAP.TAB, SKABER.TAB, GENINFO.TAB, DOKMAP.TAB:

FILMAP.TAB (Fil(12) , MedieID(8))

SKABER.TAB (Skaber(256) Startdato(8) Slutdato(8));

GENINFO.TAB (MedieID(8) DokBib(8) Beskrivelse(256))

DOKMAP.TAB(DokID(8), MedieID(8), UnderDir(8))

d.) Metadata, der skal udfra de oplysninger, som medlevers sammen med de databærende tabeller, beskrevet i punkt b, udskrives en XML fil kaldet metadata. Metadata opmærkes udfra en beskrevet syntaks som findes i EBNF. Der kontrolleres om denne syntaks bliver overholdt. Det ville være ønskeligt om de oplysninger, som indtastes i brugergrænsefladen under punkt b. kan genereres direkte i metadata.

Det er ikke sikkert at vi i gruppen er i stand til at kunne nå at kunne udarbejde den fulde version af programmet, men at vi i den anledning foretager en udvælgelse af enkeltdele i programmet, som vi så skriver.

Bilag 1

Der er ikke vist alle skærbilleder, idet funktionen og sammenhængen endnu ikke er helt klarlagt.

Aktuelt kunne det tænkes at afleveringsværktøjets brugergrænseflade kunne udformes som følger: Her er eks. et par af skærbillederne, som det fremgår af faner er det udtænkt at fungere i to niveauer.

Aflæveringsværktøj fantastika

Hoved menu | Hjælpotabeller | CSV filer | Metadata | Hjælp

Test Ny Post Slet Post << < > >> ? Hjælp

Arkver Filmap Geninfo Skaber Test

Arkvid (8)

Aflæveringstype (1)

Medfield(8)

TidligereAfl. (8)

SystemNavn (50)

StartDato (8)

SlutDato (8)

Aflæveringsværktøj fantastika

Hoved menu | Hjælpotabeller | CSV filer | Metadata | Hjælp

Test Ny Post Slet Post << < > >> ? Hjælp

Arkver Filmap Geninfo Skaber Test

Fil (12)

Medfield(8)