

# **Arkivskabte beskrivelser og administrationshistorisk dokumentation af e-arkivalier**

Tilgængeliggørelsesprojektet, delprojekt 12

Else Hansen

Rigsarkivet maj 2007

Indholdsfortegnelse:

<b>Arkivskabte beskrivelser og administrationshistorisk dokumentation af e-arkivalier .....</b>	<b>1</b>
<b>Beskrivelse af e-arkivalier. Problemer og forslag. Resumé.....</b>	<b>4</b>
Tilgængeliggørelsesprojektet og TGP12 .....	7
Arkivskabte beskrivelser .....	7
Dokumentation af e-arkivalier.....	8
Forslag til justeringer af interne procedurer og til videre initiativer .....	9
Med- og modspil.....	10
<b>Arkivtraditioner og e-arkivalier .....</b>	<b>12</b>
<i>Proveniensprincippet – kort fortalt</i> .....	12
Proveniens, arkivskaber og Daisy.....	14
Kontekst og proveniens.....	17
<i>Kontekst, præsentationsstruktur, logisk struktur og data</i> .....	18
<i>Emnebaseret søgning</i> .....	22
<b>Arkivskabte beskrivelser.....</b>	<b>24</b>
Kontekstbaseret søgning?.....	26
Informationsnetværk eller adresseringsnetværk.....	26
Registrering af informationsnetværk .....	29
Reference, informationsnetværk og beskrivelse.....	30
<i>Reference</i> .....	33
E-arkivaliets navn og eventuelle kaldenavne .....	33
Yderår.....	33
<i>Informationsnetværk</i> .....	34
Dataskaber: Hvor kommer data fra? (Aktør og heuristisk enhed).....	36
Databrunder: Hvem har adgang til data? (Aktør og heuristisk enhed).....	37
Kontinuitet til forgænger-/efterfølgerregistre.....	38
Basisregistre og standardidentifikatorer.....	39
<i>Beskrivelse mhp brugers identifikation af e-arkivalier</i> .....	39
Formål: hvorfor dette e-system .....	42
Indhold: hvad finder man i dette e-system? .....	43
E-arkivaliets type og afleveringsmodel.....	45
Systemansvarlig.....	45
Myndighedernes egen kvalitetskontrol .....	46
Sprog .....	47
Publikation af og om data.....	47
Visning af dokumentation fra geninfo .....	48

<b>Dokumentation .....</b>	<b>49</b>
Processuelle aspekter – papir- og e-arkivalier.....	49
Tidligere diskussion om administrationshistorisk dokumentation .....	50
<i>Hvad skal myndighederne dokumentere?</i> .....	52
E-systemets formål .....	52
E-systemets lovgrundlag .....	53
E-systemets indhold.....	53
Administrative funktioner .....	54
Præsentationsstruktur.....	55
Sletning eller overskrivning af data i e-systemet.....	56
Læsning af data efter aflevering .....	56
Tilvejebringelse og videregivelse af data.....	56
Forgænger- og efterfølgerregistre.....	58
Myndighedernes egen kvalitetskontrol .....	58
Publikation af og om data.....	58
Udvikling og drift af e-systemet.....	60
<b>Supplerende forslag og overvejelser .....</b>	<b>61</b>
DOK-IT: Pilotprojekt om indsamling af dokumentation .....	61
Opdeling af Geninfo .....	61
Kravrevision og interne procedurer .....	63
IT-arkitektur: TGP og Daisy .....	64
Retrospektiv brug af modellen for registrering og dokumentation .....	65
Hvordan bevares e-systemernes præsentationsstruktur?.....	65
Hvordan tilpasses registrering og dokumentation fortsat virkelighedens e-systemer?.....	66
<b>Bilag - oversigt .....</b>	<b>67</b>

## ***Beskrivelse af e-arkivalier. Problemer og forslag. Resumé***

Elektroniske arkivalier – e-arkivalier – bevares i dag af Statens Arkiver således, at arkivaliernes data sammen med de tabeller, de er opdelt i, er sikret for eftertiden. Denne metode – konverteringsstrategien – sikrer ikke i sig selv, at oplysninger om, hvordan myndighederne har anvendt de elektroniske systemer – e-systemer – eller hvordan data er tilvejebragt, indsamles og bevares sammen med data fra e-systemerne. Det betyder, at det vil blive vanskeligt for fremtidens brugere at få det fulde udbytte af e-arkivaliernes mange data. Uden sikker viden om hvordan og hvorfor data er indsamlet og brugt af såvel den myndighed, der har afleveret e-arkivaliet til Statens Arkiver, som af andre myndigheder, vil fremtidens arkivbrugere stå over for en særdeles vanskelig opgave, når disse datas udsagnskraft skal vurderes. Mens papirarkivalier ofte i selve deres fysiske udformning indeholder udsagn om deres anvendelse og om sagsgangen, vil e-arkivalier på grund af konverteringsstrategien ikke i sig selv indeholde den slags udsagn, men oplysninger om e-arkivaliernes brug og dermed om e-systemernes funktionaliteter kan sikres gennem målrettet dokumentation. En systematisk dokumentation af datas oprindelse og brug vil have flere fordele:

- Det vil opleves som *god service* af arkivbrugerne, hvis de i tilknytning til e-arkivaliet får oplysninger om datas oprindelse og brug (eller eventuelt referencer til disse oplysninger)
- Det vil betyde en *effektiv udnyttelse af ressourcerne*, hvis der ikke senere skal bruges tid på at undersøge om og hvor disse oplysninger findes, specielt fordi de er nemmest tilgængelige for såvel arkivet som den afleverende myndighed i forbindelse med anmeldelse og aflevering af e-arkivaliet.

Arkivbrugernes muligheder for at fremsøge og identificere e-arkivalier i det generelle søgesystem, Daisy, afhænger helt af, hvordan e-arkivalierne er beskrevet i dette system. Søgeselementet bør give mulighed for at bruge de oplysninger, som arkivbrugeren på forhånd har om e-arkivalierne, ligesom der som et umiddelbart resultat af søgningen bør være mulighed for at vurdere, om det fremfundne arkivalie vil indeholde de oplysninger, som efterspørges.

Derfor bør der i eller i tilknytning til søgesystemet findes indholdsmæssige og kontekstuelle beskrivelser af e-arkivalierne.

E-arkivalier vil typisk og i stigende grad indeholde oplysninger om større sagsbehandlingskomplekser, der måske strækker sig over flere myndigheder, og som – hvis der havde været tale

om papirarkivalier – ville have omfattet adskillige arkivalier (heuristiske enheder). Der er derfor behov for en mere detaljeret beskrivelse af e-arkivalier i søgesystemet.

Tidligt i projektforsøget blev der stillet spørgsmål ved, om arkivskabte beskrivelser af e-arkivalier skal være søgbare. Fordelen ved at gøre de arkivskabte beskrivelser søgbare er, at de derved vil få større anvendelsesmuligheder, ikke blot som en beskrivelse af e-arkivaliet, når det er søgt frem i søgesystemet, men også som en hjælp i selve søgeprocessen. Det foreslås derfor, at de arkivskabte beskrivelser udformes som udfyldelse af en række felter (besvarelse af en række spørgsmål). Alternativt kan de arkivskabte beskrivelser udformes som større enheder af prosatekst, men det vil mindske mulighederne for at bruge dem i søgningen betydeligt. Det foreslås tillige, at de umiddelbart tilgængelige dele af dokumentationen for e-arkivalier kan vises i tilknytning til søgningen af arkivalier. Der gives i rapporten en skitse til udformning af *et effektivt integreret søge- og beskrivelsessystem*, som vil have flere fordele:

- Det vil opleves som *god service* af arkivbrugerne, at de nemt kan søge og identificere de relevante e-arkivalier.
- Det vil betyde en *effektiv udnyttelse af ressourcerne*, hvis der hverken skal bruges person- eller maskinressourcer på at søge om adgang til eller tilgængeliggøre og genrejse e-arkivalier, som alligevel ikke er de rette.

Løsning af problemerne vedrørende dokumentation og søgning/beskrivelse har en fælles berøringsflade, idet at Statens Arkiver ved at sikre sig viden om, hvordan myndighederne har tilvejebragt og anvendt data både kan sikre, at en sådan viden vil være til rådighed for fremtidige arkivbrugere, og at e-arkivalierne kan beskrives på en måde, så identifikation af det rette arkivalie bliver mulig. Krav om tilstrækkelig dokumentation af e-arkivaliernes brug er et emne, som med fordel kan detaljeres i Statens Arkivers regelgrundlag (afleveringsbekendtgørelsen).

De ønskede oplysninger til brug for registrering og dokumentation vil ved anmeldelse og aflevering være den afleverende myndighed og sikkert også den sagsbehandlende arkivar bekendt. Denne hidtil ”tavse viden” bør formaliseres og struktureres, så oplysningerne bliver søgbare og sikret for fremtidens arkivbrugere. Den foreslåede løsning lægger op til, at IT-arkivarernes ressourcer snarere bruges i arbejdet med at skaffe tilstrækkelig dokumentation og søgbar registrering end med at udarbejde udførlige prosabeskrivelser af e-arkivalierne.

Denne rapport tager udgangspunkt i to grundlæggende spørgsmål om søgning og dokumentation af e-arkivalier:

- Hvilke oplysninger skal registreres (og i hvilken form) for, at brugerne af e-arkivalier nogenlunde sikkert kan identificere de rette arkivalier i søgesystemet?
- Hvilken dokumentation er nødvendig for, at brugerne kan vurdere om e-arkivalier indeholder oplysninger, der vurderes som troværdige i den konkrete arkivsituation?

Desuden er der på baggrund af rapporten iværksat et pilotprojekt, hvor der afprøves nye rutiner til indsamling af viden om myndighedernes tilvejebringelse og anvendelse af e-systemer.

Spørgsmålene behandles ud fra en administrationshistorisk synsvinkel – ikke ud fra en IT-faglig. Derfor lægges der vægt på en kontekst-orienteret undersøgelse af, hvorfor og hvordan data er indsamlet, hvilke myndigheder, der har været involveret i brugen af e-systemet og med hvilke formål og hvordan data er blevet anvendt. Rapporten behandler ikke spørgsmålet om kriterier for bevaring og kassation af e-arkivalier, men forholder sig til (dele af) den proces, der finder sted efter, at Statens Arkiver har besluttet at bevare et e-system.

Rapporten giver nogle meget konkrete bud på besvarelsen af ovenstående spørgsmål – med håb om at kunne medvirke til, at der inden længe kan findes holdbare løsninger på de rejste problemer. Det er klart, at der kan findes andre løsninger end de foreslåede, men det er mit håb, at der kan være enighed om, at de opstillede problemer – og måske også præmisserne for at løse dem – er relevante.

## **Tilgængeliggørelsesprojektet og TGP12**

Projektet om tilgængeliggørelse af elektroniske arkivalier har til formål i årene 2006-2008 at udvikle værktøjer og metoder, som gør det muligt for brugere uden særlige tekniske forudsætninger at benytte e-arkivalierne. Data skal omformes til et format, der vil gøre det muligt at søge i data, bl.a. ved at indlæse dem i en SQL database. Et mål er at brugeren skal kunne benytte foruddefinerede søgninger i databaserne. Et andet mål er at udvikle et program, der understøtter brugerens egne søgninger og dataudtræk uden, at det kræver særlige IT tekniske forudsætninger.

Et tredje mål er, at brugeren kan finde frem til de relevante arkivalier og kan afgøre om de fund, der leveres af søgesystemerne, er relevante for brugerens undersøgelser. Dette aspekt behandles i nærværende rapport. Projektinitieringsdokumentet for Tilgængeliggørelsesprojektet har fastsat, at delprojekt 12 skal ”undersøge behov og muligheder for kortfattede arkivskabte beskrivelser af e-arkivalier med henblik på at lette identifikationen”<sup>1</sup>, og har dermed angivet en delvis løsning på det rejste problem. Begrundelsen for Tilgængeliggørelsesprojektets delprojekt om arkivskabte beskrivelser er, at umiddelbart tilgængelige beskrivelser af det enkelte e-arkivalie vil betyde, at arkivbrugerne hurtigt og nemt kan afgøre om det fremfundne e-arkivalie er det rette, og det vil være nemt at orientere sig i flere e-arkivalier, at sammenligne dem og derigennem finde de(t) relevante. For arkivet rummer brugerens præcise identifikation af e-arkivalierne den umiddelbare fordel, at der ikke bruges ressourcer på at genrejse arkivalier, som efter blot en overfladisk undersøgelse alligevel ikke er interessante for brugerens formål.

### **Arkivskabte beskrivelser**

I den arkivskabte beskrivelse, som anvendes til både søgning og identifikation af e-arkivaliet, forslås tre niveauer indført:

- Reference. Hvordan arkivet finder e-arkivaliet – kvikreferencer for brugeren
- Informationsnetværk. Identificere hvordan data flyder mellem aktørerne gennem arkivalierne. Oplysninger til brugeren om, hvor data er kommet fra, og hvem der har anvendt dem.
- Beskrivelse/Identifikation af e-arkivaliet. Nøjere beskrivelse af e-arkivaliet, som skal sætte brugeren i stand til at identificere det e-arkivalie, som vil indeholde de efterspurgte oplysninger.

---

<sup>1</sup> Tilgængeliggørelse af elektroniske arkivalier 2006-2008. Projektinitieringsdokument. August 2005. side 26. J.nr. 2005-550-0001 <<http://saww/sa/udvproj/tp/tilgængprojini.doc>>

Referenceoplysninger som e-arkivaliets navn, e-nummer og arkiveringsversionsidentifikation er i forvejen kendte oplysninger i søgesystemet, og fremskaffes gennem kendte rutiner. Indhentning af oplysninger om informationsnetværk og til beskrivelse af e-arkivalierne vil kræve nye rutiner, nye spørgsmål og krav til de afleverende myndigheder, men vil også betyde, at e-arkivalier fremover kan søges og identificeres med stor sikkerhed. Disse oplysninger vil i hovedsagen være bekendte for de afleverende myndigheder, som derfor næppe vil finde det vanskeligt at videregive dem til Statens Arkiver.

Det vurderes, at emneregistrering af arkivskabere med fordel kan ske med anvendelse af Hof- og Statskalenderen som kilde. De ressortbeskrivelser, som gives der, vil oftest være så detaljerede, at de også kan dække brugen af e-systemer.

### **Dokumentation af e-arkivalier**

Struktureret tilvejebringelse af administrationshistorisk dokumentation bør ske ved, at Statens Arkiver identificerer en række områder, som bør dokumenteres. Derimod kan der ikke opstilles en "facitliste" for, hvilke dokumenter der skal indgå i dokumentationen, da sagsdannelsen omkring e-systemer vil variere fra system til system og fra myndighed til myndighed. Følgende områder bør dokumenteres:

- E-systemets formål og indhold
- Overførsel af data til og fra andre e-systemer
- E-systemets administrative funktioner, herunder retningslinier for indtastning i e-systemet
- Lovmateriale som baggrund for e-systemets oprettelse og brug
- E-systemets fælles funktioner med ældre og nyere systemer
- Myndighedernes egen kvalitetskontrol
- E-systemets tekniske udformning

Realiseres dette forslag vil mængden af dokumentation til hver arkiveringsversion stige, og det foreslås derfor, at der foretages en opmærkning af dokumentationen / en opdeling af geninfo, så dokumentation af forskellig oprindelse og med forskelligt indhold vil kunne identificeres maskinelt.



## **Forslag til justeringer af interne procedurer og til videre initiativer**

Udredningen af hvilke oplysninger, der skal registreres for at sikre identifikation af e-arkivalier, og af hvilken dokumentation, der skal give mulighed for at vurdere e-arkivaliers udsagnskraft, er et projekt, som ikke kun peger ind ad mod tilgængeliggørelse af e-arkivalier, men som også peger ud ad mod andre funktioner i Statens Arkiver. Derfor indeholder rapporten tillige forslag til justeringer af enkelte interne procedurer og til videre initiativer.

- Pilotprojekt med struktureret indsamling af administrationshistorisk dokumentation i projekt A og B fra juni 2007 og ca. et år frem. Resultaterne fra dette projekt bør indgå i kravsrvisionsprojektet samt i TGP's videre arbejde
- Etablering af passende registreringsfaciliteter i Daisy, TGP eller ..?
- Løbende at følge myndighedernes arkivdannelse med fokus på IT-baserede processer – og påse at arkivernes egne systemer løbende tilpasses de aktuelle forhold hos myndighederne.
- Projekt til afdækning af hvordan (dele af) e-systemers oprindelige funktionaliteter kan bevares og/eller dokumenteres, eventuelt med anvendelse af emuleringsstrategi.
- Opdeling af dokumenterne i geninfo med henblik på at kunne identificere de umiddelbart tilgængelige dokumenter samt opdeling efter forfatterkategori (myndighed eller arkiv) og hensigt (til myndighedens egen administration eller specielt udformet ved aflevering til arkiv).
- Afleveringsbekendtgørelsen gøres mere præcis mht. krav til hvilke områder, der skal omfattes af administrationshistorisk dokumentation
- Bevaring af dokumentation af e-arkivalier fastsættes i BK-bestemmelserne
- Bevaring af dokumentation om datas tilvejebringelse fra e-systemer, som kasseres med den begrundelse af data bevares andetsteds
- Retrospektiv registrering af allerede afleverede e-arkivalier med brug af oplysninger fra geninfo og fra Statens Arkivers journalsager.

## Med- og modspil

Arbejdet med TGP12 har taget udgangspunkt i mit arbejde i ERICA-projektet, der resulterede i bogen "Ny viden – gamle idéer. Elektroniske registres indførelse i centraladministrationen" (2006).<sup>2</sup> Konklusionerne fra denne undersøgelse er gradvist blevet skærpet og tilpasset TGP12's problemstillinger. Desuden er der inddraget forskellig litteratur, som der henvises til i noterne. Specielt skal fremhæves min kollega, Lone Smith Jespersens rapport: "En informationsvidenskabelig og arkivteoretisk undersøgelse af arkivalier og forslag til dokumentation af data i systemafhængige arkiveringsversioner af elektroniske arkivalier" (Statens Arkiver. Januar 2005)<sup>3</sup>, som fra et infologisk og arkivistisk synspunkt diskuterer registrering og dokumentation af e-arkivalier. Vi har i vid udstrækning arbejdet med de samme emner – konklusionerne ligger heller ikke så langt fra hinanden – og for yderligere diskussioner af infologiske og arkivistiske indfaldsvinkler til e-arkivaliers registrering og dokumentation, kan jeg med stor respekt henvise til Lones rapport.

Undervejs i arbejdet med TGP12 har der været mulighed for at præsentere disse idéer i forskellige fora: mundtlige oplæg ved IT-ekspertseminar på Island april 2006, Arkivforeningens seminar om elektroniske registre maj 2006, Nordiske Arkivdage august 2006 samt ved interne møder i Rigsarkivet november 2006 og januar 2007. Projektets resultater er desuden blevet diskuteret i e-styregruppen og i Koordineringsgruppen flere gange i 2007.

I forbindelse med belysning af diskussioner og beslutninger i Daisyudviklingen takker jeg arkivcheferne Inge Bundsgaard og Kirsten Villadsen Kristmar for at have åbnet deres private arkiver for mig.

I arbejdet med TGP12 er der gennemgået geninfo og journalsager for en række afleverede e-arkivalier. I bilag 1 præsenteres en række e-arkivalier, som så vidt muligt er "registreret" efter denne rapport's forslag. Disse "registreringer" er udarbejdet på grundlag af det materiale, som allerede findes som dokumentation, i Statens Arkivers journalsager og som sagsbehandlernes "tavs viden" om e-systemerne. Der er derimod ikke indhentet supplerende oplysninger hos de afleverende myndigheder. Jeg har selv udarbejdet en del af disse cases, andre er udarbejdet af IT-

---

<sup>2</sup> Else Hansen, Mette Hall-Andersen, Asbjørn Romvig Thomsen og Asger Svane-Knudsen. *Ny viden – gamle idéer. Elektroniske registres indførelse i centraladministrationen*. 272 sider. Syddansk Universitetsforlag. 2006.

Hansen, Else. "Ny viden – gamle idéer: elektroniske registre i den danske centraladministration." *Human IT* 9.1 (2007): 38–66. <<http://www.hb.se/bhs/ith/1-9/eh.pdf>>

Else Hansen: Elektroniske registres indførelse i den danske centraladministration. *Arkiv, Samhælle och Forskning* 2006:2

<sup>3</sup> Lone Smith Jespersen: *En informationsvidenskabelig og arkivteoretisk undersøgelse af arkivalier og forslag til dokumentation af data i systemafhængige arkiveringsversioner af elektroniske arkivalier*. Statens Arkiver. Januar 2005 <<http://www.sa.dk/sa/itogarkiv/teknologi/erfa/edokrapport.pdf>>

arkivarerne Birgitte Vedel-Larsen, Lise Qwist Nielsen og Mette Hall-Andersen. Som forventet gav det værdifuld inspiration til projektet at få deres 'fremmede øjne' på det foreløbige udkast til arkivskabte beskrivelser, som blev tilrettet på en lang række punkter efter samtaler med de tre IT-arkivarer.

Forskellige arbejdspapirer og versioner af udkast til arkivskabte beskrivelser er blevet diskuteret ved adskillige møder i TGP-gruppen, og jeg skylder såvel tidligere som nuværende medlemmer af gruppen stor tak for megen inspiration og konstruktiv kritik. I vil sikkert kunne genfinde jeres "fingeraftryk" rundt om i rapporten.

## **Arkivtraditioner og e-arkivalier**

E-arkivalier udgør en relativt ny gruppe af arkivalier over for de ”konventionelle” papirarkivalier og adskiller sig på en lang række punkter fra disse. Som forvaltningsakter kan e-systemer anvendes af flere myndigheder samtidig, idet forskellige myndigheder kan nedlægge oplysninger i eller hente oplysninger fra de samme e-systemer på samme tid. Desuden adskiller e-systemer sig fra papirarkivalier ved at bearbejdning af deres oplysninger kan ske uden at efterlade sig spor. Som arkivalier er e-arkivalier langt mere komplekse end papirarkivalier, idet de ofte vil rumme flere oplysninger og flere arbejds gange end det typiske papirarkivalie.

### **Proveniensprincippet – kort fortalt**

Proveniens har gennem de sidste hundrede år nok været det mest centrale begreb til forståelse af arkivaliers ordning og brug.

Princippet om proveniens har haft overordentlig stor gennemslagskraft. Der er tale om principper for arkivaliernes ordning (og dermed for deres opstilling) og beskrivelse. I 1800-tallet havde man arbejdet med forskellige former for emnemæssig ordning af arkivalierne, men det havde vist sig vanskeligt – uoverkommeligt – at udarbejde dækkende emnemæssige beskrivelser af arkivalierne. I stedet introduceredes proveniensprincippet omkring århundredeskiftet:

- arkivalierne ordnes efter, hvilken myndighed der har skabt dem
- den myndighedsskabte (oprindelige) orden opretholdes

Proveniensprincippet blev i Danmark introduceret af rigsarkivar V.A.Secher i 1918. Hans artikel kom til at udgøre grundlaget for udvikling af proveniensprincippet indførelse i Danmark, og den blev så sent som i 1977 stadig citeret som den bedste kilde til forståelse af proveniensprincippet:

”Arkivalierne skulle ... anbringes, hvor de i Virkeligheden høre hjemme, d.v.s. der, hvor de ad den regelmæssige administrative Forretningsgangs Vej ere ankomne for at forblive.”<sup>4</sup> (*EHA's understregninger*)

Man antog altså, at arkivalier ”i virkeligheden” hører hjemme hos en bestemt myndighed – og at de forbliver hos denne. Det er som bekendt ikke tilfældet med oplysninger i e-systemer. Flere myndigheder kan lægge oplysninger ind i det samme register; flere (og måske andre) myndigheder kan hente oplysninger fra samme register; en myndighed (eller privat firma) kan være an-

---

<sup>4</sup> *Ordlister til brug for medarbejdere ved Rigsarkivets 2.afdeling*. 1977. Side 191, citerer V.A.Secher: Om Proveniensen- (Hjemmehørs-) Princippet. Meddelelser fra det danske Rigsarkiv 1906-18.

svarlig for e-systemets struktur og udformning, mens andre myndigheder er ansvarlige for indhold og brug; oplysninger i e-systemer kommer ikke nødvendigvis umiddelbart fra mennesker, man kan komme fra andre e-systemer. I den verden bliver det særdeles vanskeligt at afgøre, hvor oplysningerne ”i virkeligheden” hører hjemme, og hvor de ender ”for at forblive”.

I indledningen til ”Rigsarkivet og hjælpemidlerne til dets benyttelse” fra 1983 bringes også en definition af proveniensprincippet:

”Ifølge proveniensprincippet defineres et arkiv som en bestand af aktstykker (der kan have vidt forskellig materiel form), som er vokset frem hos den arkivskabende person eller institution som følge af dennes virksomhed og bevareret (arkiveret) af denne.”<sup>5</sup> (*EHA's understregning*)

Med pointeringen af at der skal være tale om et arkiv, der ikke alene opbevares hos en bestemt myndighed, men som også er skabt ”som følge af dennes virksomhed”, er det endnu mere tydeligt, at e-arkivalier udfordrer proveniensbegrebet, idet en af de store administrative fordele ved e-systemer netop er, at data kan skabes (indsamles) hos en myndighed, men gemmes og bearbejdes i et e-system hos en anden myndighed, sammen med andre data, som måske kommer fra en tredje myndighed.

Der blev tidligt i den danske arkivdiskussion gjort opmærksom på, at der var problemer med at få e-arkivalier til at passe ind i den traditionelle opfattelse af proveniens. I 1982-83 skrev arkivar Poul Thestrup således artiklen Proveniensbegrebets afskaffelse, hvor han påpegede, at ”multi-acces-registre” ikke kan forstås med det traditionelle proveniensprincip. Som eksempel nævnte han motorregistret, ”som en række organer i fællesskab opdaterer og trækker oplysninger fra”.

”Man kan naturligvis formelt opretholde fonds-proveniensprincippet ved at sige, at registret tilhører en bestemt institution, nemlig den systemansvarlige institution, men det vil kun være en formel omgåelse af problemet. ... generelt er der ingen grund til at beslutte, at fonds-proveniensprincippet skal ophæves. Det vil den tekniske udvikling i vidt omfang nemlig gøre for os. ... Det er vel information, og ikke arkiver, der skal opbevares til eftertiden.”<sup>6</sup>

Som reaktion på dette skrev arkivaren Michael Hertz i 1984, at myndighedernes stigende brug af e-systemer ikke i sig selv betød, at proveniensbegrebet kom under pres. Han påpegede, at ved at sammenholde arkivaliet med dets oprindelse ville brugeren kunne få vigtige oplysninger om de forhold, som arkivaliet beskæftiger sig med. Identifikationen af én bestemt proveniens anså han ikke for væsentlig, men derimod opretholdelse af muligheden for at lade arkivalierne være kilder

<sup>5</sup> *Rigsarkivet og hjælpemidlerne til dets benyttelse*. I. 1. bind. Red. Af Wilhelm von Rosen. 1983. side 17

<sup>6</sup> Proveniensprincippets afskaffelse. Poul Thestrup. *Arkiv*. Bd. 9, nr. 2/3 (1982/83). Side 166-170

til forståelse og analyse af myndighedernes aktiviteter. En ny vurdering af proveniensbegrebet ville gøre det anvendeligt, også over for e-arkivalier:

”Arkiverne kan bevidne retstilstande, forløbet af forvaltningsmæssige ærinder, af de aktiviteter som myndighederne har udfoldet i samspil med det omgivende samfund og derpå dokumenteret. Dette er arkivernes opgave. Den lettes gennem anvendelsen af proveniensprincippet. Det er ikke ’information’ som interesserer, men de opgaver, som er blevet dokumenteret under udførelsen og hvorom der foreligger information.”<sup>7</sup>

Hertz fremhævede, at ”information” kan man finde så mange steder, ikke mindst på bibliotekerne, men at det, som adskiller netop arkivalier fra andre typer af informationer, er, at arkivalierne – ud over at give oplysning om forhold uden for myndighederne – også giver oplysning om, under hvilke vilkår ”information” er dannet, og herunder også med hvilket formål, den er dannet. Arkivalier kan altså karakteriseres ved at indeholde *både* informationer *og* oplysninger om, hvordan disse informationer er blevet skabt.

Netop dette er et problem for e-arkivalier bevaret efter konverteringsstrategien. I den nødvendige konvertering mistes (næsten) alle referencer til e-arkivaliets oprindelse, formål og brug. Det er derfor nødvendigt med en ekstra dokumentationsindsats for at kompensere for dette tab.

### **Proveniens, arkivskaber og Daisy**

Med beslutningen om at etablere en database til registrering af Statens Arkivers samlinger (Daisy) kom der sidst i 1990’erne igen fokus på en dansk forståelse af proveniensbegrebet, idet et af Daisys centrale begreber: aktør, blev præsenteret som ”en præcisering og nuancering” af proveniensbegrebet.<sup>8</sup> Aktør blev hurtigt anvendt synonymt med arkivskaber, og blev forstået som den myndighed, ”den aktør eller de aktører, der enten dannede den heuristiske enhed eller systematisk henlagde sine arkivalier i den pågældende heuristiske enhed.”<sup>9</sup> Den enkelte arkivskaber kan forbindes hierarkisk med over- og underordnede myndigheder, ikke ved deres forhold til arkivalierne, men ved myndighedernes indbyrdes forhold. Daisy giver mulighed for at knytte flere arkivskabere i sammenfaldende eller forskellige perioder sammen med det enkelte arkivalie. Der er derimod ikke tale om en nuancering af proveniensbegrebet, som tager højde for, at forskellige myndigheder kan have forskellige roller over for det samme arkivalie. Aktørbegrebet er udtryk for en meget enkel og faktisk endimensional opfattelse af proveniensbegrebet:

---

<sup>7</sup> Proveniensprincippet forståelse. Michael Hertz. *Arkiv*. Bd. 10, nr. 2 (1984). Side 107-121. side 119

<sup>8</sup> Katekismus. 2. udgave. Side 7. *Statens Arkivers intranet*. <<http://saww/sa/daisy/phj/Doc/kat.doc#Aktør>>

<sup>9</sup> Katekismus. 2. udgave. Side 6. *Statens Arkivers intranet*. <<http://saww/sa/daisy/phj/Doc/kat.doc#Aktør>>

” Arkivskaber er med andre ord den organisatoriske enhed, som skaber den heuristiske enhed, og ikke dem, der skaber de enkelte HE-elementer (breve, sager mv.), det være sig andre organisatoriske enheder eller enkelte ansatte hos en aktør.”<sup>10</sup>

Daisy kan registrere hierarkiske relationer mellem de enkelte arkivskabende myndigheder, men kan ikke skelne mellem de roller forskellige myndigheder kan have over for de enkelte arkivalier. I de første diskussioner om udvikling af Daisy’s begrebsapparat fandtes dog en mere nuanceret opfattelse af relationer mellem myndigheder og arkivalier, idet man diskuterede begrebet Arkivalieskaber til forskel fra Arkivskaber. De to begreber blev i 1997 defineret således:

”Arkivalieskaber er den myndighed mv. der skaber arkivalier, dvs danner og tilvejebringer dokumenter til en eller flere arkivkomponenter. Disse arkivalier kan eventuelt ligge som en del af en anden arkivskabers arkivdannelse.

Arkivskaber er den myndighed mv. der på skabelsestidspunktet er arkivansvarlig for de(n) arkivkomponent(er), hvori arkivalierne samles.”<sup>11</sup>

Opdelingen blev begrundet med, at den ville gøre det muligt dels at opretholde en arkivfundsdefinition, hvor der kun er én arkivskaber til hvert arkivfond; dels at skabe mulighed for at registrere, at der rent faktisk er flere myndigheder mv., der producerer arkivalier til en arkivkomponent (og dermed til et arkivfond). Der synes at have været uenighed om det hensigtsmæssige i at opretholde arkivfundsbegrebet, altså at samle de arkivalier, som en myndighed havde skabt. Begrebet Arkivalieskaber var nok ikke først og fremmest tænkt til den situation, hvor flere myndigheder lægger data ind i det samme elektroniske register, men snarere til den situation, hvor flere myndigheder lægger papirsager i det samme journalsagsarkiv (det kunne f.eks. være udvalg og råd, som journaliserer i departementets arkiv). På den baggrund blev det senere besluttet primært at knytte arkivalieskaber til arkivskabere, men dog at registrere det, hvis man ”tilfældigvis” opdagede en relation mellem en arkivalieskaber og et arkivalie.<sup>12</sup> Endnu senere nåede man frem til at betragte arkivalieskaber som aktører, idet det blev bemærket, at ”en arkivkomponent sagtens kan have flere arkivskabere i forlængelse af hinanden, men kun én arkivskaber af gangen.”<sup>13</sup> Slutteligt blev begrebet arkivalieskaber afskaffet.<sup>14</sup> Der blev også diskuteret roller som arkivansvarlig, arkivstrukturansvarlig, arkivopbevarer og arkivafleveringsansvarlig, men disse blev afvist fra begrebsapparatet med den begrundelse, at registrering af disse roller udelukkende ville

---

<sup>10</sup> Katekismus. 2. udgave. Side 6. *Statens Arkivers intranet*. <<http://saww/sa/daisy/phj/Doc/kat.doc#Aktør>>

<sup>11</sup> Centrale begreber i de kommende sags- og arkivdatabaser. 24. januar 1997. SM/- j.nr. 51-11. *KVK's arkiv*

<sup>12</sup> Notat. Kirsten Villadsen Kristmar. 16. marts 1999. *KVK's arkiv*

<sup>13</sup> Notat. Kirsten Villadsen Kristmar. 27. marts 1999. *KVK's arkiv*

<sup>14</sup> Håndskrevet notat. Kirsten Villadsen Kristmar. 26.maj 1999. *KVK's arkiv*

ske af hensyn til sagsbehandlingen, men at de efter endt aflevering ikke længere ville have nogen relevans.<sup>15</sup>

Samtidig fortsatte en diskussion om, hvorvidt det ved elektroniske registre ville være hensigtsmæssigt at registrere en arkivskaber og en række arkivalieskabere, eller om alle, som havde lagt data ned i registret, skulle betragtes som arkivskabere. Som argument for det sidste blev det hævdet, at de alle måtte betragtes som ligeværdige bidragydere.<sup>16</sup> Forskelle mellem flere arkivskabere til det samme arkivalie skulle altså ikke registreres; der blev i stedet lagt vægt på det, som arkivskaberne havde til fælles.

Ud over denne interne arkivfaglige diskussion skal det også bemærkes, at brugen af e-systemer har ændret sig gennem tiden. Frem til starten af 1990'erne fandtes der kun ringe tilskyndelse for myndighederne til at sammenlægge (eller blot sammenholde) deres ressortområder pga. e-systemer. E-systemer blev anset for at høre til den enkelte myndighed, og det var op til den enkelte myndighed selv at beslutte, hvilke e-systemer der skulle indføres, og hvordan de skulle udformes.<sup>17</sup> Med Finansministeriets udgivelse af "Effektiv edb i staten 1992" blev der imidlertid lagt op til at ændre denne opfattelse. Et perspektiv som "Arbejdsprocesser smelter sammen" og indsatsområder som "Data som fælles ressource: genbrug i datafællesskaber af elektroniske oplysninger på tværs af institutionsgrænser" og "Fælles opgaveløsninger – fælles programmel: nye koncepter for effektiv administration" var tydelige signaler om, at statslige e-systemer fremover ikke skulle begrænses til de enkelte myndigheder, men at e-systemerne burde anvendes på tværs af myndighedsgrænser og til nye arbejdsprocesser i myndighederne.<sup>18</sup> Dette blev yderligere styrket med ændringer i Registerloven i 1990, som betød, at samkøring af registre inden for områderne arbejdsløshedsforsikring, sygedagpenge, social pension, social bistand, delpension og uddannelsesstøtte blev tilladt med det formål at kontrollere udbetaling af offentlige ydelser.<sup>19</sup>

Med disse signaler om ny administrativ brug af e-systemer var bolden også givet op til, at arkivvæsenet kunne overveje om registreringen af e-arkivalier i tilstrækkelig grad afspejlede de administrative forhold, som havde skabt arkivalierne. På den anden side tog det nogen tid, før de

---

<sup>15</sup> PK's første kommentarer til MiniSvupsa, emneliste version 2. 26.januar 1999. *KVK's arkiv*

<sup>16</sup> Notat. Kirsten Villadsen Kristmar. 11. maj 1999. j.nr. 1998-082-0007. *KVK's arkiv*

<sup>17</sup> Else Hansen et al.: *Ny viden – gamle idéer*. 2006. Side 20-38

<sup>18</sup> *Effektiv edb i staten. Rapport fra det edb-politiske udvalg om statens brug af informationsteknologi i 90'erne*. Finansministeriet. København. 1992. side 6-7

<sup>19</sup> Peter Blume: *Personregistrering*. København. 1996. side 162 -167



nye signaler fra Finansministeriet var omsat til fungerende e-systemer hos de statslige myndigheder, og den nye brug af e-systemer var næppe slået fuldt ud igennem i slutningen af 1990'erne.

### Kontekst og proveniens

Allerede lidt før 1990'ernes danske diskussion om proveniens var der i det canadiske tidsskrift *Archivaria* blevet publiceret en artikel af den canadiske arkivar Terry Cook: "What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift".<sup>20</sup> I denne artikel, som siden er citeret i utallige tekster om arkiver og arkivbegreber, lægger Cook op til en ny forståelse af arkivteoriens centrale begreber, bl.a. proveniensprincippet. Han mener, at proveniens ikke længere vil blive opfattet som et statisk begreb, der sætter et arkivalie direkte i sammenhæng med dets arkivskaber, men at det vil udvikle sig til et begreb, der rummer arkivskabers funktioner og processer ved arkivaliets skabelse. Noget lignende bliver fremført af Lone Smith Jespersen:

”... det vil være frugtbart at anvende proveniens som et begreb, der drejer fokus mod oprindelsen af struktur og data med det formål at sikre den nødvendige forudviden for at fortolke data”<sup>21</sup>

Lone Smith Jespersen tager i sin analyse udgangspunkt i den infologiske model, som bl.a. lægger vægt på, at data i sig selv ikke kan forstås, men at data først kan forstås i sammenhæng med den nødvendige forudviden, altså viden om forholdene ved e-systemets oprettelse, brug m.m. Hun fremhæver, at oplysninger om oprindelsen af data er nødvendig for at kunne vurdere deres pålidelighed og endelig, at det ud fra et historisk – kildekritisk – synspunkt er væsentligt at have viden om datas oprindelse.<sup>22</sup>

OAIS-modellen, som er udviklet med henblik på e-arkivalier, præsenterer begreber til forståelse af såvel data (både struktur og indhold) som administrationshistorisk og teknisk dokumentation. I denne forbindelse er det væsentligt at bemærke modellens skelnen mellem Content Information (som nogenlunde svarer til, hvad vi kalder ”data”)<sup>23</sup> og Preservation Description Information (som nogenlunde svarer til dokumentation). OAIS-modellens proveniens-begreb, som

---

<sup>20</sup> *Archivaria*, nr. 43, Spring 1997

<sup>21</sup> Lone Smith Jespersen (2005). Side 22

<sup>22</sup> Lone Smith Jespersen (2005). Side 24

<sup>23</sup> *Reference Model for an Open Archival Information System (OAIS)*, January 2002. P 1-8

<<http://public.ccsds.org/publications/archive/650x0b1.pdf>>:

**Content Information:** The set of information that is the original target of preservation. It is an Information Object comprised of its Content Data Object and its Representation Information. An example of Content Information could be a single table of numbers representing, and understandable as, temperatures, but excluding the documentation that would explain its history and origin, how it relates to other observations, etc.

hører til Preservation Description Information, ligger på linie med de sidstnævnte forståelser af proveniens:

**“Provenance Information:** This information documents the history of the Content Information. This tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. This gives future users some assurance as to the likely reliability of the Content Information. Provenance can be viewed as a special type of context information.

**Context Information:** This information documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects existing elsewhere.”<sup>24</sup>

Der lægges ikke længere vægt på at etablere redskaber til opstilling og genfindning af arkivalierne – hvilket nok var relevant for papirarkivalier for 100 år siden, men som ikke er relevant for elektroniske arkivalier. Derimod lægges der vægt på, at det er arkivets opgave at sikre sammenhængende viden om arkivaliet og dets oprindelse og brug (kontekst). Når der i den oprindelige definition af proveniensbegrebet blev lagt vægt på at opretholde arkivets oprindelige orden, skyldtes det også, at ved at studere netop denne oprindelige orden ville den senere arkivbruger kunne få viden om, i hvilken sammenhæng arkivaliet var skabt, og hvilke forretningsgange det var indgået i.

I stedet for at lægge så stor vægt på den ”oprindelige” orden og proveniens, foreslås det i nyere arkivteori, at der i arkivets beskrivelsesarbejde anlægges et bredere perspektiv på de processer, der har skabt arkivaliet, et perspektiv som tillader flere arkivskabere med forskellige roller over for arkivaliet og som giver plads til at beskæftige sig med forskellig administrativ brug af det samme arkivalie. Hele beskrivelsen af arkivalierne skal snarere hjælpe brugeren til at forstå arkivaliets kompleksitet end til blot at gengive en engang fastlagt viden om arkivaliet.<sup>25</sup>

### **Kontekst, præsentationsstruktur, logisk struktur og data**

Oplysninger om et e-arkivalies kontekst er ikke tilstrækkeligt til at beskrive et e-arkivalie fuldt ud. Et e-arkivalie kan beskrives med en opdeling i fire elementer. *Kontekst* er de administrative vilkår for e-systemet: hvem har lagt data ind i systemet, hvorfor og hvordan er data indsamlet,

---

<sup>24</sup> *Reference Model for an Open Archival Information System* (OAIS), January 2002. P 1-8  
<<http://public.ccsds.org/publications/archive/650x0b1.pdf>>: Page 4-28

<sup>25</sup> Tom Nesmith: ”Reopening Archives: Brining New Contextualities into Archival Theory and Practice”, *Archivaria*, number 60. Fall 2005. Side 266, hvor Nesmith bl.a. henviser til Peter Horsmann og Laura Millar.

hvordan er data blevet kontrolleret, og hvad er data blevet brugt til af hvem? *Præsentationsstrukturen* er den præsentation af data, som e-systemets brugere har arbejdet med, altså skærbilleder og rapporter med pluk fra udvalgte dele af databasen, hvor der kan være foretaget beregninger el. lign. på data (som stored procedures mv.), og hvor der kan være hentet data fra andre databaser visning på skærmen, til kontrol af data eller til at tilføje ekstra værdier til data.

Hertil kommer e-systemers *logiske struktur*, som er selve sammenhængen mellem databasens felter og tabeller, udtrykt ved primærnøgler og fremmednøgler, altså hvordan data relaterer sig til hinanden. Endelig står bevaring af *data* naturligvis helt centralt i bevaringsarbejdet. Data er jo selve indholdet af e-arkivaliet, de oplysninger om personer, hændelser og fænomener, som gør e-arkivaliet bevaringsværdigt. Her adskiller e-arkivalier sig ikke fra papirarkivalier. Modellen kan udtrykkes i nedenstående skema:

### E-arkivaliers sammensætning

	OAIS	e-arkivalier	e-arkivaliers behandling efter Statens Arkivers 2004-regler	papirarkivalier
<b>Kontekst</b>	Provenance Information. Context Information	Hvorfor e-systemet er oprettet, hvem har brugt det og hvilke behov har det opfyldt	Belyses usystematisk i Geninfo	Vil ofte fremgå af papirarkivaliets udformning og af dets sammenhæng med øvrige arkivalier
<b>Præsentationsstruktur</b>	Semantic Information. Structure Information	Defineres i "program", applikationer, stored procedures mm. Det som præsenteres for e-systemets brugere. Består af et pluk af felter fra databasen, felter der beregnes og faste tekster. Kan også være pluk fra andre databaser. <sup>26</sup>	Går tabt, men beskrives usystematisk i Geninfo	Vil fremgå af papirarkivaliets udformning og eventuelt af tilknyttede procedurer m.m.
<b>Logisk struktur</b>	Representation Information	Databasens struktur, forstået som en samling af relaterede tabeller, opdelt i felter.	Bevares i en version tillempt Statens Arkivers afleveringskrav	Arkivaliernes oprindelige orden
<b>Data</b>	Data Object	De informationer som findes i databasens tabeller og felter	Bevares i henhold til bevaringsbestemmelsen	De informationer, som findes skrevet ned på papiret

Bevaring af *præsentationsstrukturen* gennem dokumentation eller emulering betyder, at viden om den maskinelle bearbejdning og præsentation af data vil blive sikret for eftertiden. En afgørende forskel mellem data i papirbaserede processer og i elektroniske processer er, at mens de papirbaserede data kun kan ændres ved menneskelig indgriben, vil de elektroniske data ofte blive forandret maskinelt. De elektronisk baserede data er dynamiske – og de kan skabe ny viden i forvaltningen.

<sup>26</sup> Se f.eks. Smith Jespersen (2005), side 24-25, hvor præsentationsstruktur vises med et eksempel fra den Ny Kirkebog

”Indholdet i papirregistre var *statisk* i den forstand, at når en oplysning først var ført ind i registret, ville den forblive der, om end den kunne slettes, eller der kunne tilføjes supplerende oplysninger. Derimod er de elektroniske registres indhold *dynamisk*, idet der kan foretages beregninger på grundlag af de en gang registrerede oplysninger, ligesom oplysningerne kan samkøres med andre oplysninger, og der derved kan fremkomme helt nye informationer. Hertil kommer naturligvis muligheden for opdateringer, der kan sikre, at et register altid indeholder aktuelle data. Den dynamiske brug af de elektroniske registres data er en mulighed, som forvaltningen har udnyttet i stort omfang, reguleret af Registerloven. De elektroniske registre betød en vældig fornyelse i anvendelsen af de offentlige registre, fordi den nye teknik gav mulighed for at

- sammenholde oplysningerne fra flere registre
- foretage beregninger på registrets oplysninger
- nye typer af kontrol med data
- at udtrække og sammensætte oplysninger til helt specifikke formål”<sup>27</sup>

Rigtig mange e-systemer kan foretage beregninger på data. Det gælder bl.a. e-systemer, der skal beregne tildeling til borgerne af forskellige ydelser efter bestemte kriterier. Her vil man finde lovgivningen omsat til database-programmer. Sagsbehandleren hos myndigheden skal så indtaste de fastlagte basisoplysninger, hvorefter programmet beregner ydelsens størrelse, men sagsbehandleren vil ikke nødvendigvis få redskaber til at kunne gennemskue de beregninger, der fører til udregningen af en konkret ydelse.<sup>28</sup>

I e-systemer kan der laves forespørgsler efter behov. Det er en af e-systemernes store fordele, at data kan ”sorteres”, så kun de i sammenhængen relevante data vises. Behov for hurtigt at udskille relevante data fra et registers øvrige data er ikke nyt. F.eks. finder vi i det papirbaserede Udlændingeregister eksempler på, hvordan man der har oprettet særlige registre til at dække aktuelle opslagsbehov, f.eks. ved indrejse af flygtninge fra Ungarn i 1956, eller registre over udlændingenes nationalitet. I et elektronisk register kan den type opslagsbehov klares med en forespørgsel. Fra et arkivsynspunkt er det imidlertid en stor forskel, at de særlige papirregistre bevares i arkivet, så den senere forsker umiddelbart kan se, at her var et særligt opslagsbehov vedrørende indrejsende fra Ungarn i 1956. Derimod behøver en forespørgsel i databasen ikke at efterlade sig noget spor. Den del af historien om registrets brug forsvinder simpelt hen, når det kun er data, som bevares. Medmindre det er beskrevet i dokumentationen, naturligvis!

---

<sup>27</sup> Else Hansen et al: *Ny viden – gamle idéer*, 2006. Side 248

<sup>28</sup> se f.eks. Nina v. Hielmcrone: ”Retssikkerhed, effektivitet og borgerinddragelse i forvaltning af sociale kontant-ydelser”, *Nyhedsbrev Socialsektor*. 2004 ; Nr. 9. s. 8-15 og Orla Damkjær: Statens Arkivers bevaring af IT-arkivalier. *Den hemmelige rapport. Bevarings- og kassationsarbejdet i Rigsarkivet*. Bevarings- og Kassationsafdelingen 2002. side 64- 65. Statens Arkivers intranet <<http://sawww/RABK03BKafdred/Seminarer/20-11-2002/SamletRapport10.doc>>

Så længe registret er i brug, vil viden om dets funktioner og i en vis udstrækning også om datas validitet være til stede blandt de, som arbejder med systemet, muligvis som en "tavs viden". Men når registret er blevet til et elektronisk arkivalie, vil denne "tavse viden" kun være tilgængelig, hvis den er beskrevet i dokumentationen.

"Ved manuel sagsbehandling kan procedurer være beskrevet i manualer eller vejledninger. Beskrivelserne vil imidlertid altid være ukomplette, idet de dels forudsætter oplæring af de personer, som skal udføre procedurene, dels forudsætter en vis almen viden om administrativ praksis og dels forudsætter, at mennesker er intelligente selvindlærende enheder, der kan lave fortolkninger af utilstrækkelige specifikationer, så de er tilpasset de daglige arbejdsrutiner. Ved maskinel sagsbehandling skal procedurene derimod være fuldstændigt specificerede i form af et computerprogram. Computere er hverken selvindlærende, intelligente eller i stand til at foretage fortolkning af utilstrækkelige specifikationer."<sup>29</sup>

E-systemer har den oplagte fordel frem for papirbaserede arbejdsgange, at alle deres procedurer vil være nøje beskrevet i programmet. Oplysninger, der kan belyse de omstændigheder, hvorunder e-arkivaliets oplysninger blev skabt og anvendt, vil for papirarkivaliers vedkommende fremgå af papirarkivaliet selv i mange tilfælde, f.eks. viser varieret håndskrift, at der har været flere forfattere, eller at dokumentet er blevet til over en længere periode. Desuden vil forhold om datas anvendelse ofte fremgå af andre arkivalier, f.eks. hvis der er blevet givet oplysninger videre til andre myndigheder, eller hvis der er blevet lavet statistik på grundlag af oplysningerne. Dette er anderledes for e-arkivaliers vedkommende. Data vil se ens ud, uanset hvem der har lagt dem ind i e-systemet, og hvornår det er sket, rettelser kan foretages uden, at det kan ses i data, overførsel af data til andre e-systemer (hos andre myndigheder) kan ske uden at efterlade sig spor, og det samme gælder beregninger på data. (Disse forhold kan dog i nogle e-systemer blive gjort til genstand for særlig registrering, så der f.eks. gemmes historiske data, eller der logges oplysninger om brugere af systemet.) Disse forskellige former for maskinel bearbejdning af data er ikke blot tekniske trivialiteter. Det er simpelt hen den måde, hvorpå de elektroniske data indgår i myndighedernes sagsbehandling og i myndighedernes beslutningsgrundlag.

Det ligger uden for denne rapports rammer nærmere at diskutere, hvordan alle dele af præsentationsstrukturen skal bevares, jf. forslaget sidst i rapporten om at iværksætte et særligt projekt om bevaring af præsentationsstruktur.

---

<sup>29</sup> Orla Damkjær: Statens Arkivers bevaring af IT-arkivalier. *Den hemmelige rapport. Bevarings- og kassationsarbejdet i Rigsarkivet*. Bevarings- og Kassationsafdelingen 2002. Side 64- 65. *Statens Arkivers intranet* <<http://saww/RABK03BKafdred/Seminarer/20-11-2002/SamletRapport10.doc>>

## Emnebaseret søgning

Denne rapport inkluderer ikke et selvstændigt punkt om emnebaseret søgning. Forslag hertil findes i Emneordsrapporten af Leon Jespersen<sup>30</sup>, ligesom der er et forslag om emneindexering under udarbejdelse i Daisy-regi.<sup>31</sup>

Emneordsregistrering i tilknytning til aktører – som foreslået af Leon Jespersen – vil give brugeren en god mulighed for at komme *i nærheden af* det e-arkivalie, der indeholder de efterspurgte informationer. TGP12 har udført et mindre forsøg med emneord efter disse retningslinier, og det har vist sig, at emneord fra Hof- og Statskalenderen knyttet til arkivskabere i kombination med e-arkivaliets formålsbeskrivelse, population og variabler giver et godt grundlag for emnesøgning og præcis identifikation. Følges forslaget om ikke blot at registrere den standardiserede arkivskaber, men også arkivskabere med rollerne dataskaber, databrunder og dataansvarlig, vil der hermed fremkomme et bredt net af emneord, som vil lede hen til grupper af e-arkivalier inden for samme emne.

Emneordsregistrering af e-arkivalier på niveauet heuristisk enhed vil ikke kunne erstatte den foreslåede registrering af e-arkivaliernes formål og indhold (udtrykt i population og variabler). For at opfylde projektets mål om at give rimelig sandsynlighed for, at brugeren kan identificere det rette arkivalie uden at anvende e-arkivaliets data, må e-arkivaliets formål og indhold beskrives præcist og dermed med den variation, der er mellem de enkelte e-arkivalier. Hensigten med emneordsregistrering er så at sige den modsatte, nemlig at beskrive e-arkivaliets indhold i standardiserede termer fra en kontrolleret liste.

Det skal for en ordens skyld bemærkes, at en senere emnemæssig registrering af e-arkivalier efter retningslinier, som måtte blive udviklet med henblik på papir- og e-arkivalier, ikke vil forstyrre den beskrivelsesmodel, som her er foreslået, men derimod yderligere udbygge den.

Det skal ligeledes bemærkes, at flere andre systemer til registrering af e-arkivalier har emneregistrering som et helt centralt punkt. Det gælder f.eks. Netherlands Historical Data Archive (NHDA)<sup>32</sup> og The National Digital Archive of Datasets (NDAD).<sup>33</sup> Kort kan det siges om NDAD's beskrivelser af e-arkivalier, at disse er struktureret med emne-tesauri og navnelister som udgangspunkt for søgning. Derimod er de arkivskabte beskrivelser ikke søgbare, men meget læsevenlige. NDAD's on-line thesaurus (UKAT) til emnesøgning bygger på UNESCO's thesau-

<sup>30</sup> *Projekt struktureret søgning. RAPPORT.* Leon Jespersen. Rigsarkivet marts 2005

<sup>31</sup> Referat af SAL-møde 25.4.2006. Notat: Handlingsplan for projekt struktureret emnesøgning. Til SA-L fra Inge Bundsgaard. 20. april 2006

<sup>32</sup> <<http://www.dans.knaw.nl/en/data/nhda/>>

<sup>33</sup> <<http://www.ndad.nationalarchives.gov.uk/>>

rus 1995. Thesaurus fører til alle beskrivelsens niveauer: administrationshistorie, series-level, dataset-level and dataset documentation. Derimod fører thesaurus ikke til selve data "as this would allow the researcher to access the data without reading through the contextual information provided."<sup>34</sup>

Den beskrivelsesstandard, som foreslås i denne rapport, er udarbejdet med henblik på at identificere en kreds af arkivalier, der opfylder nærmere definerede krav og i denne kreds at udvælge netop det arkivalie, som brugeren ønsker at arbejde med, altså en proces hvor resultatet bliver stadig mere snævert. Andre beskrivelsesstandarder tilbyder at identificere "de nærmeste" lignende enheder. Det gælder f.eks. forskellige bibliotekssystemer, hvor der tilbydes "Find andre bøger om dette emne", eller Internet-boghandlen Amazon, der ikke blot finder den bog, man søger, men også tilbyder beslægtede titler. Dette sker imidlertid på grundlag af en emneregistrering af hver enkelt bog, og det, der tilbydes som "de nærmeste" fund, er andre bøger, hvor emneregistreringen på et eller flere punkter er sammenfaldende med den fremfundne. Den foreslåede beskrivelsesmetode baserer sig imidlertid ikke på faste emneordslister til beskrivelse af e-arkivaliernes indhold, men tilbyder derimod de længere fritekstfelter, der kan give mere præcise (og "skæve" eller uventede) oplysninger. Derfor vil det næppe være realistisk at forestille sig, at den foreslåede standard vil kunne udvikles til at give forslag til "de nærmeste" arkivalier. De nærmeste arkivalier vil derimod kunne findes gennem den emneregistrering, som tilknyttes arkivskaberne, uanset arkivaliernes medium.

---

<sup>34</sup> Patricia Sleeman. It's Public Knowledge: The National Digital Archive of Datasets. *Archivaria*, nr. 58, Fall 2004, p.195

## **Arkivskabte beskrivelser**

Det er bærende for dette forslag til udformning af arkivskabte beskrivelser af e-arkivalier, at disse beskrivelser for det første skal gøre rede for såvel e-arkivaliets kontekst (datas oprindelse og brug) som dets formål og indhold, og for det andet at disse oplysninger skal struktureres sådan, at de kan udnyttes maksimalt.

For at kunne gøre rede for e-arkivaliets kontekst må de arkivskabte beskrivelser nødvendigvis omfatte ikke blot det enkelte e-arkivalie, men også de øvrige e-systemer, som dette e-system refererede til. Det er netop karakteristisk for e-systemer, at der kan udveksles data mellem systemerne; det er en af forudsætningerne for at e-systemer opfattes som så effektive af forvaltningen. Arkivets beskrivelse af denne dataudveksling kan ske gennem registrering af *Informationsnetværk*, som præsenteres nedenfor.

For at give mulighed for at anvende beskrivelsernes oplysninger på flere måder forslås det, at de struktureres ensartet og så vidt muligt i maskinlæsbar form, altså ved at give afgrænsede svarmuligheder på de enkelte punkter. Det primære formål med de arkivskabte beskrivelser er, at brugeren skal kunne identificere de(t) rette arkivalie(r) i et afgrænset udvalg af arkivalier. Sekundært skal det foreslås, at de arkivskabte beskrivelser kan indgå i søgningen. Dette er nærliggende af to grunde, dels fordi de arkivskabte beskrivelser indeholder strukturerede oplysninger, som vil være søgbare, dels fordi de arkivskabte beskrivelser for at opfylde formålet om at identificere det rette arkivalie alligevel bør vises i tilknytning til søgeresultatet.

Den arkivskabte beskrivelse tilvejebringes ved arkivarernes indsats. Det er væsentligt, at arkivarerne får mulighed for at give den viden videre, som de opnår i forbindelse med sagsbehandling, der bl.a. omfatter kontakt med myndigheder, som har anvendt e-systemet.

Arkivskabte beskrivelser synes at kunne antage principielt (mindst) to former. Der kan være tale om lange, lærde afhandlinger, altså lange tekstforløb, hvori beskrives relevante forhold vedrørende e-arkivaliet. Eller der kan være tale om at opdele beskrivelsen i en række nærmere definerede underpunkter (felter).

Lange tekstforløb har den fordel, at her kan beskrives mere komplicerede sammenhænge og forhold, som "falder uden for" den opstillede beskrivelsesmodel. Til gengæld er det vanskeligere at strukturere sådanne mere frie beskrivelser, hvad der medfører, at de kan ende med at blive ret forskelligartede. De afgørende ulemper er dog, at brugernes søgemuligheder i en lang tekst er



begrænset til fritekstsøgning med de usikkerheder som sådanne søgninger indebærer,<sup>35</sup> at de lange tekster vil ikke forbinde aktører og heuristiske enheder i Daisy, og endelig at det kan blive ganske ressourcekrævende at udarbejde dem.

En opdeling af beskrivelsen i felter vil have den ulempe, at der kun levnes ringe plads til at beskrive komplicerede forhold og specielt forhold, der ikke er taget højde for i udarbejdelsen af modellen. Der kan også stilles spørgsmål ved, hvilken nytte brugeren vil have af at vide, at en myndighed måske er dataskaber til 300 registre? Vil brugeren drukne i oplysninger – eller vil brugeren netop få hjælp til med få oplysninger at finde det rette arkivalie? Fordelen ved at opdele beskrivelsen af arkivaliet i felter er, at beskrivelsen så er struktureret fra starten, hvad der vil give en større ensartethed i beskrivelserne. De helt afgørende fordele er dog, at der på denne måde gives forskellige søgemuligheder til de enkelte dele af beskrivelsen, og at en sådan strukturering vil lette udarbejdelsen af beskrivelserne.

De lange tekstforløb vil formentlig kræve mere arbejde fra arkivarens side – sætninger skal formes og korrigeres. Derimod burde en del af udfyldelsen af på forhånd definerede felter kræve mindre anstrengelser – forudsat der tilvejebringes gode skærbilleder til indtastning, og at de relevante oplysninger er indsamlet. Denne vurdering hviler naturligvis på antagelsen om, at det vil være de samme obligatoriske oplysninger, som skal indgå i begge beskrivelsestyper. Ved at vælge den – formentlig – mindre arbejdskrævende type med foruddefinerede felter burde der blive ressourcer til, at den anden vigtige del af dette forslag – dokumentationen – kan få relativt mere opmærksomhed.

I den foreslåede model arbejdes med en hybrid mellem de to modeller, idet den foreslåede beskrivelse består af en række på forhånd definerede felter, og hvoraf nogle giver mulighed for længere svar. Desuden skal det understreges, at den foreslåede model fungerer i sammenhæng med forslaget om at foretage en mere dybtgående og systematisk indsamling af administrationshistorisk dokumentation end hidtil, idet registreringens mere summariske oplysninger tænkes suppleret med de mere varierede oplysninger i dokumentationen. Det er således vigtigt, at der etableres tekniske muligheder for visning af dokumentation i sammenhæng med søgeresultatet.

---

<sup>35</sup> Se f.eks. Projekt struktureret søgning. RAPPORT. Leon Jespersen. *Rigsarkivet marts 2005*. side 91 ff. og *Whitepaper om klassifikation. Convergens inspirationsbog no.3*. Af Peter Busch, seniorkonsulent cand. mag., Jens Bruntt, seniorkonsulent, cand. scient. bibl. November 2000. Side 11-12  
<[http://www.crealize.dk/pbusch/crealize.nsf/f819a23c1afb8c1ec1256c830075e839/b747fbbf70b2010dc1256c7e0048fde2/\\$FILE/Whitepaper%20om%20Klassifikation11.doc](http://www.crealize.dk/pbusch/crealize.nsf/f819a23c1afb8c1ec1256c830075e839/b747fbbf70b2010dc1256c7e0048fde2/$FILE/Whitepaper%20om%20Klassifikation11.doc)>

## Kontekstbaseret søgning?

Søgning af arkivalier er traditionelt sket med udgangspunkt i arkivaliets proveniens, der i tidligere arkivteori er blevet betragtet som den grundlæggende determinant for arkivalierne.

”Ved brug af og søgning efter arkivalier i Rigsarkivet og andre arkiver er det uomgængeligt nødvendigt at kende det hovedprincip, hvorefter arkiver i moderne tid ordnes, dvs. proveniensprincippet...”<sup>36</sup>

Dette gentages i en let modificeret version i den aktuelle brugerrettede udgave af Daisy, hvor det i den centrale vejledningstekst på søgebilledets forside hedder. ”Søg efter arkivskaber og/eller arkivserie.”<sup>37</sup> Proveniensen (arkivskaberen) står stadig centralt i søgningen, der dog også kan omfatte de betegnelser, som Statens Arkiver har givet de enkelte arkivalier (arkivserier).

Som der tidligere er gjort rede for, er det i de senere års arkivteori blevet foreslået, at proveniensbegrebet udvides med en stærkt øget opmærksomhed på arkivaliernes kontekst. Der er ligeledes gjort rede for, at det traditionelle proveniensbegreb kun er dårligt egnet til at bestemme e-arkivalier, idet e-arkivalier er i mange tilfælde ikke skabt af én, men flere myndigheder, ligesom data i et e-system hos en myndighed kan være skabt af en anden myndighed og overført derfra.

Når der lægges så stor vægt på arkivaliernes kontekst, skyldes det ikke mindst et stigende behov for begrebmæssigt at tilpasse arkivteori til den kendsgerning, at det enkelte arkivalie ikke kan fortolkes endeligt en gang og for altid. Dette bevises vel til overmål af forskellige tiders historieforskning. Det er så spørgsmålet, om det er rimeligt fortsat at lade søgemulighederne i arkivets søgesystem tage udgangspunkt alene i den diskutale proveniens – eller om der kan åbnes mulighed for også at søge med udgangspunkt i arkivaliernes kontekst. Den arkivteoretiske vending mod et kontekst-orienteret proveniensbegreb betyder, at også registrering må genovervejes med henblik på om, kontekst er tilstrækkeligt belyst i den konkrete registrering af e-arkivalierne.

## Informationsnetværk eller adresseringsnetværk

Hvordan svarer nu Daisys begrebsapparat til de udfordringer, som søges imødekommet med nye proveniensbegreber? Som tidligere nævnt opererer Daisy med et endimensionalt arkivskaberbegreb, hvor en arkivskaber altid vil have samme rolle over for arkivaliet. Dog tillades flere arkivskabere til samme arkivalie, men stadig uden at det er muligt at skelne mellem forskellige roller over for arkivaliet. De enkelte arkivalier, de heuristiske enheder, kan knyttes sammen af

---

<sup>36</sup> Rigsarkivet og hjælpemidlerne til dets benyttelse I. 1. bind. 1983. side 17

<sup>37</sup> < <http://www.daisy.sa.dk/default.asp>>

adresseringsnetværker, som beskriver sammenhænge mellem arkivalier, der også som forvaltningsakter hørte sammen:

”Det er velkendt, at nogle heuristiske enheder er forbundet med andre heuristiske enheder. Vi plejer at sige, at en heuristisk enhed er søgemiddel til en anden heuristisk enhed. For eksempel er et journalregister søgemiddel til journal og journalsager. ... Et journalregister indeholder typisk en journalnummerhenvisning til journal og journalsager, og forbindelsen mellem disse tre heuristiske enheder består i, at der alle steder optræder en HIF, som er bygget op af de samme DIF'er, for eksempel År og Løbenr.”<sup>38</sup>

Denne beskrivelse af netværker er – som det fremgår – udviklet til at beskrive sammenhænge mellem papirarkivalier.<sup>39</sup> I de tilfælde hvor e-arkivalier er søgemiddel til andre arkivalier, f.eks. hvor en elektronisk journal (eller et elektronisk register) er søgemiddel til papirsager, kan disse begreber naturligvis også anvendes for e-arkivalier.

Derimod har Daisy hidtil afstået fra at beskrive e-arkivalier, idet man har henvist til at e-arkivalier er selvbeskrivende:

”Et E-system er et afgrænset elektronisk arkivalie, hvor metadata beskriver felter og henvisninger (adresseringer). Et E-system rummer således allerede sin beskrivelse i sig selv og behøver ingen yderligere beskrivelse i Daisy, med undtagelse af eventuelle identifikatorer til for eksempel en E-journal, der henviser til de tilhørende papirsager.”<sup>40</sup>

E-arkivalierne indeholder ganske rigtigt dokumentation i Geninfo og andre beskrivende oplysninger i metadata og forskellige hjælpetabeller. Disse oplysninger er imidlertid svært tilgængelige, efter gældende regler og praksis er de først tilgængelige, når en bruger har opnået adgang i

---

<sup>38</sup> Katekismus. 2. udgave, side 33 *Statens Arkivers Intranet* <<http://saww/sa/daisy/phj/Doc/kat.doc#HEer>>

<sup>39</sup> Den eksklusive henvisning til papirarkivalier var også tydelig i den forklaring, man gav i Katekismus 1. udgave: ”Nogle arkivalier er søgemidler til andre. F.eks. kan et navneregister være indgangen til en journal, det samme kan en kronologisk postbog, og en journal kan pege på (adressere) en bestemt sag i en sagsrække. Til at ”pege” rundt i arkivalierne har myndighederne brugt forskellige numre, tekst og årstal eller et utal af kombinationer heraf. De offentlige myndigheder har i århundreder anvendt sådanne numre i deres arkivhåndtering til at henvise til sager og breve, og dermed sikre at disse kunne findes igen, når man fik brug for det. De har med andre ord anvendt identifikatorer. En sådan henvisning til eller sammenknytning af en myndigheds arkivalier kalder vi et **netværk**. De fleste myndighedsarkiver (fra 17.-1800-tallet til i dag) er i varierende grad struktureret som netværk. Nogle meget udbygget, f.eks. centralforvaltningens arkiver, hvor mange komponenter kan være knyttet sammen med identifikatorer. Andre blot med en enkelt protokol eller fortegnelse over indkomne breve, som henviser til brevet med et brevnnummer eller en dato. Netværket (søgevejene) bindes sammen af identifikatorerne. I de tilfælde hvor en komponents identifikator peger ud (**ud-identifikator**) på en eller flere komponenters identifikatorer (**ind-identifikatorer**), er der tale om et netværk.” <<http://saww/DAISY05kat/3HE.doc>> Katekismus. Heuristiske enheder. side 3 (Katekismus’ understregninger)

<sup>40</sup> Katekismus. 2. udgave, side 13 *Statens Arkivers Intranet* <<http://saww/sa/daisy/phj/Doc/kat.doc#HEer>>

henhold til Arkivloven og eventuelt Persondataloven til det konkrete arkivalie. Disse oplysninger kan således hverken anvendes ved søgning eller identifikation af e-arkivaliet.

Hertil kommer, at det for mange e-arkivalier ikke er relevant at pege på andre arkivalier gennem adresseringsnetværker, idet e-systemerne oftest vil være opbygget således, at de i sig selv indeholder både registre og det, som registrene henviser til. Bestemte dele af databasen vil kunne fremsøges med særlige forespørgsler. Dermed bliver det i mange tilfælde uvæsentligt at kunne skabe denne type adresseringsnetværk mellem e-arkivalier.

Det er derimod relevant at pege på de sammenhænge mellem e-arkivalier, der *understøtter en kontekstorienteret forståelse af arkivalierne*, altså hvordan data gennem e-systemer kommer fra myndighed til myndighed (eller fra afdeling til afdeling i den enkelte myndighed).<sup>41</sup> Vi kunne kalde disse sammenhænge for *Informationsnetværk*. Her skal udpeges tre former for sådanne sammenhænge:

- når en myndighed har adgang til at indtaste i eller benytte data fra et e-system, som beror hos en anden myndighed
- når data overføres fra et e-system til et andet (som eventuelt beror hos en anden myndighed)
- når bestemte sagsforhold indtil et givent tidspunkt registreres i et e-system (eventuelt et papirregister), og derefter registreres i (en del af) et andet e-system

Registreringen af Informationsnetværk må således indeholde oplysninger om, hvem der har leveret data til og modtaget data fra det pågældende e-system samt hvilke systemer, der tidligere og efterfølgende varetager de samme funktioner.

Det skal bemærkes, at Informationsnetværk naturligvis ikke kun findes mellem e-systemer, men også kan findes i en papirbaseret organisation. Her kan man nævne et brev, der både har en afsender og en modtager, som et let genkendeligt eksempel. Typisk vil et sådant brev have afsat sig spor i såvel afsenders som modtagers arkiv, f.eks. i form af en kopibog hos afsender og en postbog hos modtager eller i form af henholdsvis et koncept til eller en afskrift af brevet hos afsender og det modtagne brev hos modtager. I et papirbaseret system vil denne udveksling af data, som et brev ofte vil være, således afsætte sig spor i arkivet, og dermed kunne erkendes af senere arkivbrugere. Det er anderledes med elektroniske forbindelser mellem myndigheder og mellem e-systemer. Når en myndighed lægger data ind i et e-system, der beror hos en anden myndighed, kan dette ske uden, at e-arkivaliet bærer nogen som helst spor efter det (medmindre det fremgår af dokumentationen). Tilsvarende med overførsel af data fra det ene e-system til det andet. Man

---

<sup>41</sup> Nesmith (2005) taler om "reorienting archival work to the more expansive contextuality". Side 262

kan altså påvise informationsnetværk i en papirbaseret arbejdsgang, men det er ikke relevant at registrere og dokumentere dette som ved e-arkivalier, fordi data ikke kan ”flyde” udokumenteret mellem papirarkivalier.

Informationsnetværk er et begreb, der kan kvalificere forståelsen af ”multiproveniens”, idet der ikke blot peges på, at mere end en myndighed har medvirket til skabelsen af dette arkivalie, men de forskellige arkivskaberes roller over for arkivaliet fastslås samtidig. Der kan være involveret så mange dataskabere og dataleverandører omkring et e-arkivalie, at det ikke giver mening at tale om én arkivskaber eller om flere (ligestillede) arkivskabere. Når e-systemer er så udbredte, skyldes det jo bl.a., at de gør det nemt at dele og overføre data, og at der kan være flere (samtidige) brugere: adgangen til data er ikke bundet af geografi, og der er mulighed for at give forskellige rettigheder til de enkelte brugere; arbejdsprocesser som indsamling, bearbejdning og brug af data kan adskilles og udføres af forskellige myndigheder.

Ved at registrere Informationsnetværk skaber arkivet flere indgange til e-arkivaliet, åbner for flere anvendelser af e-arkivaliet og giver mulighed for at forskelligartet viden om e-arkivaliet ved søgningens begyndelse kan anvendes. På den anden side udgør Informationsnetværket en afgrænset samling af informationer og bliver dermed ikke en uendelig beskrivelse af hvad-somhelst vedrørende e-arkivaliet.

Den hidtidige bevaringsindsats er gået ud på at bevare data, men med aktiv anvendelse af informationsnetværk vil både data og kontekst kunne bevares.

### **Registrering af informationsnetværk**

Informationsnetværkets oplysninger vil være en central del af de arkivskabte beskrivelser, som skal give arkivbrugerne mulighed for med rimelig sikkerhed at identificere det rette e-arkivalie. Viden om udveksling af data vil give væsentlige informationer om brugen af data.

Ved at udvikle mulighed for at foretage en systematisk registrering af oplysninger om informationsnetværk kan disse oplysninger også gøres søgbare og dermed give langt *bedre søgemuligheder* end den nuværende endimensionale registrering af arkivskabere, idet det ikke længere vil være nødvendigt præcis at vide, hvem Statens Arkiver har udpeget til at være arkivskaber for et e-arkivalie. E-arkivaliet vil også kunne søges frem, hvis brugeren kender en dataskaber eller en databrunder. Oplysninger om informationsnetværk skal under alle omstændigheder indgå i de arkivskabte beskrivelser og bør også være en del af den søgbare beskrivelse.

Beskrivelsen af de enkelte myndigheders udveksling af data gennem e-systemer kan beskrives på flere niveauer. Det skal her foreslås, at det registreres *mellem hvilke e-systemer og mellem*

*hvilke myndigheder, der er sket overførsel af data.* Disse oplysninger vil formentligt forholdsvis nemt kunne fås hos den afleverende myndighed, og det vurderes, at de vil give tilstrækkelige informationer i brugerens søge- og identifikationsproces. Relationer mellem et register og flere myndigheder kan beskrives ved at tilknytte ”roller” til arkivskaberne: dataskaber, databruget og dataansvarlig. Relationer mellem registre kan muligvis(!) beskrives gennem en modifikation af de begreber om identifikatorer, som allerede findes i Daisy. Der kunne *måske* udvikles en DIF-skabelon til formålet. Den skulle i så fald være ganske simpel, idet den kun skulle oplyse, at her findes relationen ”overførsel af data”, men ikke sige noget om, hvilke (typer af) data der er overført.

Alternativt kan man nøje fastslå, *hvilke data der er overført* mellem e-systemer og mellem myndigheder (svarende til Daisy-adresseringsnetværkernes udpegning af identifikatorer). Det vil typisk dreje sig om mange forskellige data, der så i registreringen skal identificeres på tabel- og feltniveau i de enkelte e-arkivalier, hvilket vil blive et meget omfattende registreringsarbejde samtidig med, at det er tvivlsomt, om en sådan detaljeret registrering ville opfylde væsentlige behov i forbindelse med søgning og identifikation?

Endelig kan man naturligvis afstå fra at foretage en systematisk registrering og blot *beskrive relationerne i notefelter*. Denne vej kan synes tillokkende, fordi det ikke kræver udvikling af nye begreber eller af registrerings- og søgefaciliteter. Desuden kan en beskrivelse af mellem hvilke e-systemer og hvilke myndigheder, der er sket overførsel eller udveksling af data placeres såvel i tilknytning til søgesystemet (Daisy) som i tilknytning til Tilgængeliggørelsessystemet (TGP) eller et helt tredje sted. Dette alternativ vil til gengæld mangle alle de fordele, som et begrebssæt og en registrering i felter giver, nemlig at relationen oprettes mellem flere e-arkivalier, at relationen dermed bliver synlig i visningen af flere e-arkivalier og arkivskabere, og at der gives mulighed for henvisning mellem flere e-arkivalier, hvilket vil have oplagte søgemæssige fordele.

### **Reference, informationsnetværk og beskrivelse**

Registrering af e-arkivalier med henblik på søgning og identifikation kan opdeles i tre formål, nemlig reference, informationsnetværk og beskrivelse:

1. *Reference*: Referenceoplysninger er helt basale oplysninger om arkivaliet som navn og datering samt f.eks. e-nummer og arkiveringsversionsID. Formålet er at kunne referere til det rette arkivalie, så det kan genfindes i arkivets systemer. Referenceoplysninger vil i sig selv kun være af interesse for den, der kender arkivaliet i forvejen. Det vil især være interne brugere, men også eksterne brugere vil kunne anvende referenceoplysninger til ”kviksøgninger”.

2. *Informationsnetværk*: Formålet med at registrere det informationsnetværk, som e-arkivaliet indgår i, er at klarlægge, hvor e-arkivaliets oplysninger er kommet fra og hvem der har brugt dem. Informationsnetværket er en videreudvikling af det traditionelle proveniens-begreb og spiller – ligesom dette – en central rolle ved brugerens søgning af det relevante arkivalie. Det er derfor vigtigt, at de oplysninger, der beskriver Informationsnetværket, er søgbare.
3. *Beskrivelse*: Formålet med beskrivelse er at identificere det *rette* arkivalie blandt de arkivalier, som er identificeret gennem søgning i Reference og Informationsnetværk. Her uddybes altså Informationsnetværkets oplysninger.

Informationer med funktionen Beskrivelse skal derfor have den kvalitet, at de sætter brugeren i stand til at adskille indholdet i det ene arkivalie fra indholdet i andre arkivalier, og at brugeren kan foretage en positiv identifikation af det rette arkivalie. Denne funktion varetages i dag i beskedent omfang gennem Daisys forskellige felter til registrering af arkivaliet, men især af feltet ”Publikumsnote”, for hvis udfyldelse der dog ikke findes retningslinier.

Selv om et bestemt forhold registreres, *fordi* det er nødvendigt for beskrivelsen og dermed for den endelige identifikation af arkivaliet, vil man ved at placere det i søgedatabasen opnå den fordel, at oplysningerne også bliver søgbare. Hermed fås mulighed for at kombinere Informationsnetværkets søge-oplysninger (der typisk vil være af mere basal art) med beskrivelses-oplysningerne (der typisk vil være mere komplicerede og forskelligartede)

I bilag 2 og 3 er fremlagt to scenarier, der illustrerer heuristiske aspekter ved den arkivskabte beskrivelse af e-arkivalier.

## TGP 12 – model for registrering og beskrivelse af elektroniske arkivalier

TGP12 - forslag	TGP12-niveau	Forklaring	Daisy - uændret	Daisy (evt. andet registreringsværktøj)
Aktør				
Arkivskaber	Inf.netværk	Afleverer e-arkivaliet, har evt. haft hovedansvaret for dets drift	X	Uændret = registrere og vise rollen "arkivskaber"
Dataskabere	Inf.netværk	Alle andre myndigheder, som har lagt data direkte ned i e-systemet. Mulighed for at registrere, at ikke-myndighed (borger, virksomhed) har leveret data til systemet. Bør have mulighed for angivelse af tider		Registrere og vise rollen "dataskaber" og dens udstrækning i tid.
Databrugere (og administrativ brug)	Inf.netværk	Alle andre myndigheder, som har brugt data direkte fra e-systemet. Mulighed for at registrere, at ikke-myndighed (borger, virksomhed) bruger systemet. Bør have mulighed for angivelse af tider		Registrere og vise rollen "databrunder" og dens udstrækning i tid.
Ansvarlig for e-arkivaliets struktur	Beskrivelse	Hvem har skabt og haft ansvaret for strukturen (IT-systemet) og for overførslen til arkiveringsversion? De(n) teknisk ansvarlige (myndighed eller privat virksomhed). Bør have mulighed for angivelse af tider		Registrere og vise rollen "dataansvarlig" og dens udstrækning i tid
Heuristisk enhed				
e-arkivaliets navn	Reference	Mulighed for angivelse af tider	X	
e-arkivaliets kaldenavne	Reference	Mulighed for angivelse af tider	X	
Heuristisk enheds yderår (dannelse og oplysning)	Reference	Mulighed for angivelse af tider	X	
e-nummer	Reference		X	
e-arkivaliets type	Beskrivelse		X	Bør kunne vises på HE-niveau
Afleveringsmodel	Beskrivelse		X	Bør kunne vises på HE-niveau
Dataskaber	Inf.net	Andre registre som har leveret data til e-systemet (opslag, overførsel, samkøring osv.). Bør have mulighed for angivelse af tider		Registrere og vise relationen "dataskaber" og dens udstrækning i tid
Databrunder	Inf.netværk	Andre registre som har brugt data fra e-systemet (opslag, overførsel, samkøring osv.). Bør have mulighed for angivelse af tider		Registrere og vise relationen "databrunder" og dens udstrækning i tid
Henviser til andre heuristiske enheder	Inf.netværk	Viser e-arkivaliet til andre arkivalier (journalsager?)	X	Addresseringsnetværk
Forgænger-/efterfølgerregistre	Inf.netværk	Hvilket register har tidligere varetaget samme funktion?		Registrere og vise rollen "forgængerregister"
e-arkivaliets formål	Beskrivelse	Formålet med at oprette og drive registret		Fritekst
e-arkivaliets indhold	Beskrivelse	Registrets centrale population, centrale variabler og anvendelse af standardvariabler		1 fritekstfelt til population 3 fritekstfelter til variabler Afkrydsningsfelter til standardvariabler
Sprog	Beskrivelse	Hvilket sprog er anvendt i e-arkivaliet?		Default=dansk, valgmuligheder: engelsk, andet
Arkiveringsversion				
Arkiveringsversions yderår (dannelse og oplysning)	Reference		X	
ArkiveringsversionsID	Reference		X	
Arkivskabers kvalitetskontrol	Beskrivelse	Er arkivskabers kvalitetskontrol dokumenteret?		Ja/nej
Publiceret materiale	Beskrivelse	Er data publiceret?		Ja/nej. Visning af publikationsliste. Mulighed for at åbne de enkelte publikationer
Oversigt over dokumentation	Beskrivelse	Henvisning til liste over arkiveringsversionens dokumentation		Visning af liste over dokumentation. Mulighed for at åbne dokumentation



## Reference

Formålet med reference er at kunne referere til det rette arkivalie, så det kan genfindes i arkivets systemer. Det vil sige, at der er tale om referencer ind ad i arkivets systemer, men *ikke* referencer mellem arkivalier og heller *ikke* referencer ud af systemet, f.eks. til ISBN-numre. *E-nummer og ArkiveringsversionsID* er interne referenceoplysninger, som især bruges af de interne brugere, men som også kan være nyttige som ”kvikreferencer” for eksterne brugere, der leder efter et bestemt og allerede identificeret arkivalie. E-nummer knyttes til den heuristiske enhed, mens ArkiveringsversionsID knyttes til arkiveringsversionen.

## E-arkivaliets navn og eventuelle kaldenavne

Det anbefales at følge Daisys hidtidige praksis, hvor der for hver heuristisk enhed udpeges et officielt navn, men hvor der tillige er mulighed for at registrere andre navne (kaldenavne) og forkortelser.<sup>42</sup> Da der ikke i Daisy gives mulighed for at angive den periode, hvor arkivaliet har haft et bestemt navn, må der vælges ét officielt navn til hvert e-arkivalie på heuristisk enhed-niveau, mens øvrige navne bliver kaldenavne, også selv om de i en periode har været officielle navne. Registrering af e-arkivaliets officielle navn og eventuelle kaldenavne vil naturligvis være af betydning for såvel søgning som identifikation af e-arkivalier.

## Yderår

Det anbefales at følge de definitioner på oplysningstid og dannelsesetid for heuristisk enhed, som er fastlagt i Daisy<sup>43</sup> og i bekendtgørelsens bestemmelser om ARKVER.TAB og SKABER.TAB.<sup>44</sup> Oplysningstid svarer til angivelserne i ARKVER.TAB og dannelsesetid til oplysningerne i SKABER.TAB. Myndigheden må ved anmeldelse eller aflevering oplyse dato for ibrugtagning og dato for ældste oplysning i e-systemet. Hvis e-systemet fortsat anvendes efter aflevering, kan der ikke angives data for sidste opdatering eller indtastning af data eller for nyeste oplysning.

---

<sup>42</sup> Indtastningsvejledning for arkivalier i Daisy: Navnetyper. *Statens Arkivers Intranet* <[http://saww/sa/daisy/phj/doc/In\\_Ark.doc](http://saww/sa/daisy/phj/doc/In_Ark.doc)>

<sup>43</sup> Katekismus. 2. udgave. side 22-23. *Statens Arkivers Intranet* <<http://saww/sa/daisy/phj/Doc/kat.doc#HEer>>

<sup>44</sup> *Bekendtgørelse nr. 342 af 11. marts 2004*, bilag 4, pkt. H

## Informationsnetværk

Tidligere i denne rapport er der introduceret en idé om at registrere *informationsnetværk* som en metode til at beskrive, hvordan data flyder mellem e-systemer. Her skal jeg give et forslag til, hvordan informationsnetværk kan registreres, så de kan anvendes til søgning og beskrivelse af e-arkivalier. Mit forslag vil tage udgangspunkt i, at det beskrives *mellem hvilke aktører og mellem hvilke heuristiske enheder, der er udvekslet informationer*. Derimod søges det ikke beskrevet, *hvilke* informationer der er udvekslet. Der er altså tale om en mindre detaljeret og mindre dybtgående registrering, end der kendes fra Daisys begreber om HIF'ere.

Der er tidligere i denne rapport givet en kort oversigt over proveniensbegrebets udvikling, herunder hvordan aktørbegrebet er blevet diskuteret og nu anvendes i Daisy. Det følger af denne redegørelse, at Daisys bestemmelse af arkivskaber-begrebet ikke er fuldt tilstrækkelig til at beskrive arkivskaberforholdene for e-arkivalier. For at nuancere arkivskaber-begrebet – og for at gøre det mere kontekstorienteret – foreslås det at etablere tre nye begreber, nemlig "dataskaber", "databrunder" og "dataansvarlig". Disse begreber præsenteres nærmere nedenfor.

De nye begreber vil sætte spørgsmålstegn ved, hvad det standardiserede daisy-begreb "arkivskaber" skal dække over, idet de tre begreber repræsenterer en raffinering af det standardiserede arkivskaber-begreb. Arkivskaberen til et e-arkivalie vil i mange tilfælde være en myndighed, som både er dataskaber, databrunder og dataansvarlig. I andre tilfælde kan der være tale om, at arkivskaberen kun har haft det driftsmæssige og administrative ansvar for e-systemet, men hverken været dataskaber eller -bruger, f.eks. ved CPR-registret. Eller at en myndighed nok har været dataskaber og -bruger, men at en række andre myndigheder også har været dataskabere og -brugere til det samme system. Hvem er så "arkivskaber"?

I Daisy-sammenhæng er "arkivskaber" imidlertid et meget centralt begreb, der for papirarkivalier er alene om at angive, hvilken myndighed et arkivalie kommer fra. For at give gode muligheder for at søge et arkivalie uden på forhånd at vide, om det er et papirarkivalie eller et e-arkivalie, er det vigtigt at e-arkivalier og papirarkivalier er registreret på samme måde. Det foreslås derfor, at der fortsat registreres en "arkivskaber" for e-arkivalier efter de samme retningslinier som hidtil, altså som oftest at den myndighed, der afleverer e-systemet til arkivet er arkivskaber, i enkelte tilfælde hvor den afleverende myndighed har overtaget e-systemet fra f.eks. en nedlagt myndighed eller ved ressortomlægninger kan den afleverende myndighed og arkivskaberen imidlertid være forskellige. Disse problematikker er velkendte i Statens Arkivers sagsbehandling.

Daisy giver mulighed for at registrere hierarkiske sammenhænge mellem aktører og heuristiske sammenhænge mellem heuristisk enheder. Derimod gives der ingen mulighed for at registrere kronologiske sammenhænge mellem heuristiske enheder, altså at registrere forgængere og efterfølgere.<sup>45</sup>

Almindeligvis tilknyttes e-arkivaliet (den heuristiske enhed) arkivskabers højeste administrative niveau. Det er en gammel praksis, som forklares i Blåmejsebogen i et notat fra 1997:

”[IT-SAG registrerer] ikke den inddeling i undernumre, som OFFARK har brugt til nummerering f.eks. af afdelinger i en myndighed. Dette skyldes dels, at edb-opgaver typisk vedrører hele myndigheden, dels at bevaring/aflevering mest hensigtsmæssigt forhandles passende ’centralt’.”<sup>46</sup>

En mere præcis registrering af arkivskaber, altså registrering af hvilken afdeling/kontor, der har skabt eller brugt e-arkivaliet, vil imidlertid i mange tilfælde kunne give både heuristiske og administrationshistoriske fordele. Registrering på lavere organisatoriske niveauer vil give en mere præcis identifikation og medvirke til at synliggøre relationer mellem en myndigheds forskellige registre. Et e-system kan sagtens vedrøre en afgrænset del af myndigheden, men til gengæld have dataskabere og -brugere uden for myndigheden. Registrering af arkivskabere på lavere niveauer er særligt relevant, fordi større og større dele af arkivmassen bliver elektronisk. Hvis denne kun registreres på overordnet niveau, betyder det, at den heuristiske hjælp, der findes i Daisys registrering af arkivskabere på lavere niveauer, vil forsvinde.<sup>47</sup> En sådan registrering vil underbygge ønsket om at registrere informationsnetværker, ikke kun mellem myndigheder, men også inden for de enkelte myndigheder, idet arkivskaber skal forstås som dataskaber, -bruger og -ansvarlig.

---

<sup>45</sup> Det er ikke lykkedes at dokumentere denne beslutning – eller argumenterne bag den. Jeg har dog hørt, at beslutningen skulle være begrundet med, at det ikke er lykkedes at give en konsistent definition af, hvornår der er tale om en forgænger samt at eventuelle forgængere skulle kunne findes gennem emnesøgning. (mundtlige meddelelser)

<sup>46</sup> IT-SAGSdatabasen (den midlertidige) En kort vejledning vedr. indhold og søgemuligheder. Juni 1997. IT-afdelingen. *Blåmejsebogen*, bd. 2, faneblad 7 (Inger Jensens eksemplar)

<sup>47</sup> På dette punkt anbefales det, at også e-arkivalier fremover registreres efter de anbefalinger, der findes i Daisys Katekismus 2. udgave, side 6

### **Dataskaber: Hvor kommer data fra? (Aktør og heuristisk enhed)**

Dataskaber er den, der leverer data til e-systemet. Der kan være flere dataskabere til et e-system:

- den myndighed som "ejer" e-systemet (=arkivskaber)
- andre myndigheder, som har adgang til direkte at indføre data i e-systemet
- andre e-systemer hos den myndighed, der er ansvarlig for e-systemet
- e-systemer hos andre myndigheder

Data kan altså enten komme fra en myndighed (aktør), det vil sige være indtastet (eller kvalitetskontrolleret hvis borgerne taster data ind i systemet) af en myndighed. Eller data kan komme fra et andet e-system (heuristisk enhed), der kan være "ejet" af en anden myndighed. Hvis data kommer fra et andet e-system registreres dette e-system (e-arkivalie) med relationen "dataskaber" (måske med begreber udviklet fra HIF- og DIF-registreringen), systemets aktører i rollerne som dataskaber, -bruger, -ansvarlig og arkivskaber berøres derimod ikke.

Viden om, hvor data kommer fra, kan spille en rolle ved brugerens endelige identifikation af e-arkivaliet, hvis brugeren søger oplysninger fra en bestemt kilde (der har leveret data til e-systemet) – eller modsat kan brugeren udelukke netop dette e-arkivalie, fordi oplysningerne ikke kommer fra denne kilde eller kommer fra en kilde, brugeren har udelukket.

Registreres det, hvor data kommer fra, vil det blive muligt at finde den omspurgte information uden på forhånd at kende det bevarede register og dets arkivskaber, blot man kender en dataskaber. Det vil ligeledes blive muligt at lede en oplysning frem, selv om den er bevaret på et højere aggregeringsniveau, end den er indsamlet.

Oplysning om dataskaber til et e-arkivalie vil være relevant i Daisy. Registrering heraf forudsætter, at der i Daisy gives mulighed for at registrere aktør-rolle "Dataskaber", og at der gives mulighed for at registrere informationsnetværks-relationer mellem heuristiske enheder. Der må gerne kunne sættes tid på disse relationer. Det ligger ikke i forslaget, at informationsnetværks-relationer mellem heuristiske enheder skal give oplysning om, hvilke data der udveksles mellem de to e-systemer, blot at der udveksles data. En præcis identifikation af hvilke data, der udveksles mellem e-systemerne, vil kunne blive ganske omfattende, og dermed dels betyde meget merarbejde ved registreringen, dels vil mange informationer medvirke til at give et "sløret" billede af relationen mellem e-systemerne. (Derimod skal det naturligvis fremgå af dokumentationen og eventuelt af metadata-filen, hvilke data der overføres mellem systemerne).

Registreringen kan ske efter en liste med faste svarmuligheder. Der må vælges flere svar:

<i>Faste svarmuligheder:</i>	<i>Fra (tid)</i>	<i>Til (tid)</i>	<i>(svarmulighed)</i>
Indsamlet af myndigheder (hvilke)			Valg fra Daisy
Indtastet af borger / virksomhed (ikke myndighed)			Afkrydsning
Overført fra / opslag i andre e-systemer (hvilke)			Valg fra Daisy
Andet (hvad)			Notefelt
Ved ikke			Afkrydsning

### **Databrunder: Hvem har adgang til data? (Aktør og heuristisk enhed)**

Databrunder er den, der har adgang til at benytte data i e-systemet. Der kan være flere databrunder til et e-system:

- den myndighed som ”ejer” e-systemet (=arkivskaber)
- andre myndigheder, som har adgang til direkte at hente data fra e-systemet
- andre e-systemer hos den myndighed, der er ansvarlig for e-systemet
- e-systemer hos andre myndigheder

Data kan således benyttes af flere myndigheder, enten ved at de har adgang til e-systemet eller ved at de får overført data fra e-systemet til andre e-systemer.

Begreberne ”dataskaber” og ”databrunder” er parallelt opbygget, men omhandler forskellige forhold. Viden om brugen af data kan være vigtig for brugerens vurdering af datas udsagnskraft, idet det vil være myndighedens administrative brug af data, som har afgjort afgrænsningen af data. Der vil ofte være flere brugere til ét system, mange internt hos myndigheden, men også brugere hos andre myndigheder – og måske også i offentligheden, f.eks. gennem såkaldte ”portaler”. Desuden kan data være blevet givet videre til andre e-systemer ved overførsel eller ved at disse har haft mulighed for opslag i e-systemet.

Registreres det, hvem der har brugt data (fået overført data), vil det blive muligt at finde den omspurgte information uden på forhånd at kende det bevarede register og dets arkivskaber, men blot ved at kende til brugeren. Ligeledes vil viden om, hvem der har brugt data, spille en rolle ved den endelige identifikation

Oplysningen om databrunder vil være relevant i Daisy, og vil kunne tilføjes, hvis der i Daisy gives mulighed for at registrere arkivskaber-rolle ”Databrunder”, og der gives mulighed for at registrere informationsnetværks-relationer mellem heuristiske enheder. Informationsnetværks-relationer mellem heuristiske enheder skal ikke give oplysning om, hvilke data der udveksles mellem de to e-systemer, blot at der udveksles data.

Registreringen skal ske efter en liste med faste svarmuligheder. Der må vælges flere svar:

<i>Faste svarmuligheder:</i>	<i>Fra (tid)</i>	<i>Til (tid)</i>	<i>(Svarmulighed)</i>
Myndigheder der henter, rekvirerer eller får indberettet oplysninger fra registret (hvilke)			Valg fra Daisy
Videregiver data til eller samkøres med andre systemer (hvilke)			Valg fra Daisy
Andre			Notefelt
Ved ikke			Afkrydsning

Det kan overvejes også at registrere, hvilke administrative funktioner e-systemet har haft, jo senere om dokumentation af brugen af e-arkivalier.

### **Kontinuitet til forgænger-/efterfølgerregistre**

Et register kan siges at være forgængerregister for et andet register, når de to registre

- anvendes til de samme funktioner

og/eller

- registrerer de samme personer, hændelser eller fænomener med samme formål

Et nyt register er nemlig ikke altid udtryk for nye registreringer. Ofte vil de samme fænomener også tidligere være registreret i andre registre, som kan være elektroniske eller på papir.<sup>48</sup>

Når en myndighed anmelder et nyt register, vil det oftest være nemt at få oplysning om, hvor de samme oplysninger tidligere er blevet registreret. Her til kommer, at hvis data er konverteret ind i det aktuelle e-system fra et ældre e-system, bør dette under alle omstændigheder dokumenteres, da konvertering kan have betydning for datas kvalitet.

Registrering af forgængerregister vil kunne skabe en forbindelse mellem papirarkivalier og e-arkivalier. Det vil i mange brugersituationer være relevant at kunne følge en serie af oplysninger fra register til register, f.eks. når man ønsker at etablere en tidsserie eller blot at følge et fænomen gennem en længere periode. Registrering af forgængerregistre vil tjene heuristiske formål, når man f.eks. formoder, at en datarække må fortsætte ud over den heuristiske enheds yderår. Forgænger- eller efterfølgerregister adskiller sig *kronologisk* fra hovedregistret, men omhandler (til dels) samme emne.

En sådan registrering skal ikke nødvendigvis indebære en registrering af, *hvilke* oplysninger der er fælles for registrene, eller *hvorfor* der er sket en overgang fra det ene e-system til det næste (årsagen til skift fra det ene system til det næste kunne være 1) skift i ressortområde, 2) skift af system, og 3) ny administrativ brug af samme data).

<sup>48</sup> Else Hansen et al.: *Ny viden - gamle idéer* indeholder flere eksempler på, at elektroniske registre har afløst papirregistre.

Efter dette forslag vil registreringen fungere, hvis det blot registreres, at der findes fælles oplysninger eller funktioner for registrene. Det vil sige, at det ikke skal registres (men dokumenteres), hvilke data der er fælles for forgænger- og efterfølgersystemerne.

### **Basisregistre og standardidentifikatorer**

Der henvises meget ofte til enkelte e-systemer som cpr-registret og cvr-registret. Med rette kan der rejses den indvending mod at lade arkivet registrere *samlige* forbindelser mellem alle e-arkivalier, at det vil give et uoverskueligt stort antal referencer til sådanne e-systemer. For at undgå dette bør der derfor gives mulighed for at registrere sådanne e-systemer som *basisregistre*, altså registre som indeholder et fælles datagrundlag for flere grene af forvaltningen, f.eks. cpr-registret.<sup>49</sup> Når visse registre kan udpeges som basisregistre, skyldes det, at de indeholder nogle meget centrale og ofte anvendte oplysninger, f.eks. cpr-nummer med henvisning til den enkeltes navn, bopæl m.v. Sådanne *standardidentifikatorer* anvendes jo så i mange e-systemer, og vil følgelig nu og senere kunne anvendes til at sammenholde oplysninger fra e-arkivalier med standardidentifikatorer. Ved at registrere om der i et bestemt e-arkivalie findes standardidentifikatorer – frem for at registrere forbindelser til basisregistre – vil arkivet sætte brugerne i stand til allerede i søgebilledet at afgøre, om flere e-arkivalier kan køres sammen på grundlag af disse identifikatorer.

### **Beskrivelse mhp brugers identifikation af e-arkivalier**

I dette forslag gives flere veje til at *frem søge* e-arkivalier: gennem referencer, informationsnetværk eller med tiden emnesøgning i tilknytning til aktører. Disse søgeveje vil kunne give flere mulige e-arkivalier som resultat; de vil ikke nødvendigvis lede direkte hen til ét e-arkivalie. For at kunne *identificere netop det rette e-arkivalie* i et søgeresultat med flere fund skal hvert enkelt e-arkivalie beskrives nærmere, end det er sket ved registrering af informationsnetværk. Der mangler stadig vigtige oplysninger om e-arkivaliets formål, indhold m.m. Sådanne forhold vil i sagens natur variere fra e-arkivalie til e-arkivalie. Samtidig er det vigtigt for den præcise identifikation af e-arkivaliet, at netop disse forhold beskrives meget præcist, da det oftest vil være disse forhold, der kan afgøre, om man har fundet "det rigtige" register.

Ud over identifikation skal beskrivelser af e-arkivalier også give arkivbrugerens oplysninger om arkivaliets historie, administrative brug m.m. I de ældre papirregistraturer kunne dette integreres med de relativt stringente beskrivelser af indholdet i de enkelte arkivfonde, når disse blev

---

<sup>49</sup> "Basisregister" er et gammelt begreb, f.eks. defineret i *Statens edb til administrative formål 1974/75*, side 52

suppleret med ”kommentarer” eller ”anmærkninger”. Denne ordning kan kun dårligt overføres direkte til elektroniske arkiv-søgesystemer, der som alle andre e-systemer fordrer en høj grad af konsistens i registreringerne, hvorfor anmærkningernes ad hoc prægede form bliver tydelig og vanskelig at anvende. I stedet foreslås en systematisering af de administrationshistoriske oplysninger, for at de kan beskrives ensartet, og også for at de kan indgå i søgningen.

Inden forslaget til søgbare beskrivelser af e-arkivalier præsenteres, skal sammenhængen mellem søgning og oplysninger om arkivaliet diskuteres på baggrund af konkrete eksempler.<sup>50</sup> The National Archives i England er et eksempel på en elektronisk arkivregistrering med detaljerede beskrivelser af e-arkivalier, opdelt i forskellige felter til beskrivelse af administrationshistoriske (kontekst-orienterede) forhold, tekniske forhold og arkivmæssige forhold (f.eks. vedrørende adgang til arkivaliet).<sup>51</sup> *Disse ganske omfattende beskrivelser kan imidlertid ikke anvendes til søgning, men får først værdi, når det relevante e-arkivalie er fremsøgt.* Søgning kan derimod ske gennem emneord, fastlagt i en tesaurus, fritekstsøgninger i udvalgte felter og gennem navneregistre over virksomheder, personer og steder.

Canadiske Tom Nesmith foreslår endnu mere vidtgående brug af omfattende arkivbeskrivelser. For at sikre arkivbrugerne dyb og varieret forståelse af det enkelte arkivalies kontekst foreslår han nemlig, at arkivarerne udarbejder *essays*, der belyser forskellige aspekter ved det enkelte arkivalie. Nesmith henviser til et eksempel på det canadiske arkivvæsens hjemmeside, hvor en digitaliseret udgave af premierminister King's (1874-1950) dagbøger er præsenteret sammen med essays, omhandlende bl.a. dagbogsgenren og en biografi om King.<sup>52</sup> Ved at strukturere beskrivelsen i essays undgår arkivaren ifølge Nesmith de mange snærende regler for en mere præcis beskrivelse og får samtidig frihed til at beskrive det enkelte arkivalie på dets egne præmisser.

”Description should be seen as the action mediated by archivists of researching and representing the multi-faceted contextuality (or history of records or 'archival narrative' about them) which enables records and knowledge to be made through archiving.”<sup>53</sup>

Hermed sikres det, at der videregives en omfattende viden om det enkelte arkivalie, der sikkert vil sætte kommende brugere i stand til at vurdere de udsagn, som måtte kunne læses ud af dette arkivalie. Det må imidlertid stå klart – selv efter denne helt kortfattede omtale – at der er tale om

---

<sup>50</sup> For diskussion af forskellige internationale erfaringer kan henvises til Lone Smith Jespersen (2005), side 31-44

<sup>51</sup> <<http://www.ndad.nationalarchives.gov.uk/>>

<sup>52</sup> <<http://www.collectionscanada.ca/05/0532/>>

<sup>53</sup> Nesmith (2005) side 270



en særdeles ressourcekrævende procedure, der nærmest må omfatte et selvstændigt forskningsprojekt i forbindelse med udarbejdelsen af de beskrivende essays om hvert enkelt arkivalie.

Disse omfattende metoder til beskrivelse af arkivalier er spændende, men ikke uproblematisk. Hvis arkivet anvender mange ressourcer på at skabe beskrivelser af arkivalierne, bør disse beskrivelser så ikke også kunne indgå i søgningen? Skal alle de ledsagende tekster nødvendigvis udformes som essays, forfattet af arkivarer – eller kan de eventuelt bestå af ”genbrug” fra myndighedernes (eller arkivets egen) sagsbehandling? I denne rapportes forslag besvares begge disse spørgsmål med et Ja.

Den beskrivende del af registreringen, som præsenteres nedenfor, skal kunne vises for brugere i sammenhæng med oplysningerne om interne referencer og informationsnetværk *og tillige* i sammenhæng med visninger af udvalgte dele af e-arkivaliets dokumentation. Det kan f.eks. dreje sig om lister over publikationer om e-arkivaliet og lister over hvilken dokumentation, der findes i arkiveringsversionen samt mere indholdsrige dokumenter som myndighedernes dokumentation af e-arkivaliet og dokumenter fra arkivets sagsbehandling. Spørgsmålene om forskellig restriktion i adgangen til sådanne dokumenter skal diskuteres senere. Endelig bør det overvejes, om denne beskrivelse kan knyttes tæt sammen med den egentlige tilgængeliggørelse af e-arkivalierne, således at beskrivelsen er i nøje overensstemmelse med opmærkningen af data i selve e-arkivaliet og med de prædefinerede forespørgsler, som tilbydes brugerne.

Beskrivelser af e-arkivalier skal indgå i flere sammenhænge:

- identifikation af de enkelte e-arkivalier
- som element i søgningen af e-arkivalier
- som ”indgang” til de tilgængeliggjorte data i sig selv

Nedenfor gennemgås de enkelte dele af den foreslåede beskrivelse, som foreslås opdelt i formål, indhold, type og afleveringsmodel, systemansvarlig, myndighedens egen kvalitetskontrol, e-systemets sprog samt oplysninger om publikation af og om data. Disse elementer i beskrivelsen af e-arkivalier skal ses i sammenhæng med de oplysninger om e-arkivaliet, der er registreret som informationsnetværk og reference samt i sammenhæng med oversigter og dokumenter fra arkiveringsversionens dokumentation

### **Formål: hvorfor dette e-system**

E-systemets formål kan udtrykkes i besvarelsen af spørgsmål som, hvorfor blev e-systemet oprettet, og hvilke behov det skal opfylde. Svarene skal kunne hjælpe arkivbrugeren til at afklare, om der i dette e-arkivalie vil findes de oplysninger, der efterspørges. Det er kostbart og besværligt at oprette et e-system, og man må derfor antage, at det altid vil ske med et bestemt formål for øje.<sup>54</sup>

”[the Aim and Purpose element] is particularly suited to the description of datasets because the nature of databases means that they are rarely established in an *ad hoc* fashion, but are designed with specific reason and purpose in mind.”<sup>55</sup>

E-systemers formål angives allerede i systematisk form i flere forbindelser: i love og bestemmelser, i anmeldelser til Datatilsynet og i kravspecifikationer. Der findes dog ikke et sted, hvor formålet med alle e-systemer er angivet. Det meget varierede kildegrundlag lægger op til en meget pragmatisk løsning på problemet om, hvordan e-arkivaliers formål beskrives: nemlig "bedst muligt". I praksis vil det sige, at myndigheden besvarer anmeldelses-/dokumentationsskemaets spørgsmål om e-arkivaliets formål, arkivaren kontrollerer, at teksten er forståelig og tilstrækkeligt omfattende (bl.a. ved at konsultere de relevante kilder), og denne tekst anvendes derefter til registrering af e-arkivaliets formål.

Som det fremgår af bilag 7 er der tale om ganske store forskelle på, hvordan e-systemers formål angives, ikke blot mellem forskellige institutioner, men også inden for samme institution. Registreringen af e-systemers formål kan derfor ikke blot kopiere eksisterende beskrivelser af formål.

Muligheden for at registrere formål i fritekstfelter betyder, at det enkelte e-arkivalies formål kan beskrives temmelig præcis, men vil også indebære, at beskrivelserne kan blive ret forskellige. Det vil derfor ikke give mulighed for at oprette umiddelbare relationer mellem formålmæssigt beslægtede e-arkivalier eller på anden vis sammenknytte disse tematisk. En vis tematisk sammenknytning vil dog ske gennem den planlagte tilknytning af emneord til arkivskabere med forskellige roller.

---

<sup>54</sup> Hvilke forviklinger det kan medføre, når oprettelse ikke sker med et bestemt formål for øje, er 1970'ernes Fremmedarbejderregister et godt udtryk for (*Ny viden - gamle idéer*, side 202-214).

<sup>55</sup> Elizabeth Shepherd and Charlotte Smith: The Application of ISAD(G) to the Description of Archival Datasets. *Journal of the Society of Archivists*, Vol. 21, No. 1, 2000. Side 63.

Såfremt det ved aflevering konstateres, at e-systemets formål er blevet ændret (eller udvidet eller indskrænket) noteres dette i det samme fritekstfelt med angivelse af, hvornår ændringen er sket og hvad ændringen består i.

Specielt for journaler og esdh-systemer angives det primære formål: ”elektronisk registrering af sager og dokumenter og arkivering af alle dokumenter på papir”, ”elektronisk registrering af sager og dokumenter og arkivering af dokumenter både på papir og i elektronisk form” eller ”elektronisk registrering af sager og dokumenter og elektronisk arkivering af alle dokumenter.”<sup>56</sup> Ofte vil journaler og esdh-systemer tillige have andre funktioner i organisationen, f.eks. kan de danne grundlag for produktionsstatistik, som indberettes til de bevilgende myndigheder. Sådanne supplerende formål angives også.

### **Indhold: hvad finder man i dette e-system?**

En beskrivelse af registrets indhold øger muligheden for præcis identifikation af e-arkivaliet. Oplysningerne vil fra e-arkivalie til e-arkivalie være af så forskellig karakter, at det ikke vil være muligt at lave en liste, hvorfra der kan foretages valg. Det vil således kun være muligt at lave fritekstsøgninger på disse oplysninger. Indhold beskrives gennem ”central population”, ”variabler” og ”standardidentifikatorer”.

- *Central population: Hvem eller hvad er registreret i e-systemet?*

De personer, hændelser eller fænomener, som registreres i e-systemet. Registrets centrale population, f.eks. studerende ved institutionen, ”Personer, der benytter de institutioner eller boformer, der drives i henhold til lov om social service § 94” eller ”Omfang og art af anmeldelser til politiet om lovovertrædelser”. Ved store e-systemer kan der være flere centrale populationer.

For e-journaler og esdh-systemer vil den centrale population være ”alle sager” eller et afgrænset udsnit af sager, som så må beskrives nærmere.

- *Variabler: Hvad er der registreret om den centrale population?*

Oplysninger i *semantisk* afgrænsede datagrupper, som har central betydning i forhold til registrets formål og administrative brug. Der skal ikke nødvendigvis registreres henvisninger til konkrete felter i konkrete tabeller i e-arkivaliet, men snarere til semantisk afgrænsede grupper

---

<sup>56</sup> Disse formuleringer svarer til *Cirkulære af 15. marts 2002 om anmeldelse og godkendelse af elektroniske journaler og elektroniske dokumenthåndteringssystemer*, § 2

af felter, der tilsammen beskriver forhold, som varierer for den centrale population. (altså ikke fornavn + efternavn + adresse + postnummer + telefon, men derimod ”pensionsmodtager” eller ”patient” eller ...). For e-journaler (og ESDH-systemer) registreres f.eks. sagsidentifikation (journalnummer), sagstitel og dokumentdato.

- *Standardidentifikatorer*: Betydningsbærende identifikatorer, som bruges i adskillige e-systemer, og som derfor giver mulighed for at sammenholde oplysningerne i flere e-arkivalier. Det registreres altid, hvis CPR-nummer findes i e-arkivaliet, idet denne oplysning dels kan bruges *internt i arkivvæsenet* i forbindelse med adgangstilladelser, dels kan den give en bruger et hurtigt overblik over muligheder for at sammenholde registre. Tilsvarende registreres andre nationale felter som f.eks. CVR/SE-nummer, matrikelnummer, kommune- og amtsnumre (før og efter 1.1.2007), WHO's sygdomskoder og geografisk registrering.

Udpegning af ”central population” og variabler vil sikkert komme til at bero på sagsbehandlerens og myndighedens skøn: hvad er ”centralt” og hvad er ”vigtige variabler”? Det samme ville i øvrigt være tilfældet for en længere sammenhængende tekstbeskrivelse af, hvad e-systemet drejer sig om. Som bekendt vil de stadig større e-systemer dreje sig om en mængde forskellige forhold, og det vil næppe være muligt i en oversigtlig beskrivelse at gøre rede for alle disse forhold. Dog findes der i e-arkivaliernes dokumentation en liste over tabeller, felter, kodelister, relationer m.m. (XML-filen), som giver mere udførlig besked om indholdet af e-arkivaliet.<sup>57</sup> Da det kræves, at den overholder standarden for XML, er den dog ikke umiddelbart læsbar for brugere helt uden kendskab til XML, men kan vises sammen med søgeresultatet som andre dele af dokumentationen.

Det må overvejes nærmere, hvilke muligheder der er for at sammenknytte udvælgelsen af central population og vigtige variabler med de prædefinerede søgninger, som der er planer om at tilbyde brugerne ved tilgængeliggørelse af e-arkivalier. Det ville uden tvivl fremme brugervenligheden i såvel søge- og beskrivelsessystemet (som er emnet for nærværende rapport) som i fremvisningssystemet (som er emnet for arbejdet i de øvrige delprojekter i Tilgængeliggørelsesprojektet), hvis der bliver en høj grad af overensstemmelse mellem den population og de variabler, som præsenteres i søge- og beskrivelsessystemet – og den første prædefinerede søgning, som brugeren støder på i fremvisningssystemet. *Det foreslås altså, at den første prædefinerede*

---

<sup>57</sup> Bekendtgørelse nr. 342 af 11. marts 2004 om arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer. Bilag 3, pkt B

*søgning viser netop den centrale population og de centrale variabler, som brugeren har set i beskrivelsen.*

### **E-arkivaliets type og afleveringsmodel**

Begreberne E-arkivaliets type (EDH-system, ESDH-system, journal eller register) og afleveringsmodel for registre (engangsaflevering, årgangsaflevering og øjebliksbillede/akkumulerende aflevering) er fastlagt i afleveringsbekendtgørelsen.<sup>58</sup> Disse oplysninger er væsentlige for arkivets administrative sagsbehandling, men er også af betydning for en beskrivelse af e-arkivaliet. Med denne registrering udskilles journaler og ESDH-systemer fra den store, heterogene gruppe af e-arkivalier, men restgruppen af registre er stadig stor. Der er tale om oplysninger, som ikke vil variere fra arkiveringsversion til arkiveringsversion i den samme heuristiske enhed, og det vil derfor være nyttigt, hvis oplysningerne (også) kan vises for brugene på heuristisk enhed-niveau.

### **Systemansvarlig**

Som tidligere nævnt kan e-arkivalier betragtes som sammensat af data, struktur, præsentationsstruktur (bl.a. applikationer og stored procedures) samt kontekst. E-arkivalier adskiller sig bl.a. fra papirarkivalier ved, at der tydeligt kan skelnes mellem struktur og data (indhold), som ydermere ofte vil have forskellige arkivskabere.

”The two-dimensional nature of datasets means that there are usually at least two creators of note: one who is responsible for creating/designing and/or maintaining the structure of the database, and one who is responsible for ‘creating’ or compiling or analysing the contents.”<sup>59</sup>

Lone Schmidt Jespersen har påpeget, at Daisys begrebsapparat ikke er klart på dette punkt. Hun peger på det sted i (1. udgave af) Katekismus, hvor der står: ”Arkivskaber til en heuristisk enhed er den aktør eller de aktører, der enten danner den heuristiske enhed, eller systematisk nedlægger sine arkivalier i den pågældende heuristiske enhed.” Hun konkluderer derpå – med rette – at det i Daisy

”... tilsyneladende [står] åbent, om arkivskaberen er skaberen af strukturen eller skaberen af indhold. Af de gældende regler for e-arkivalier fremgår det derimod klart, at der ved skaber forstås skaberen / skaberne af indhold.”<sup>60</sup>

---

<sup>58</sup> *Bekendtgørelse nr. 342 af 11. marts 2004 om arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer.* bilag 4, pkt H om feltet AfIType

<sup>59</sup> Shepherd and Smith (2000) Side 63

<sup>60</sup> Lone Schmidt Jespersen (2005) side 22

Derefter henvises til bekendtgørelsen, hvor ”Skaber” defineres som:

”Navn på alle selvstændige myndigheder, som har indtastet data i systemet. Hvis en myndigheds organisatoriske enheder ikke alle har brugt systemet eller ikke har taget systemet i brug på én gang, anføres de relevante organisatoriske enheder som arkivskabere. (Bekendtgørelse nr. 342 af 11. marts 2004 arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer, bilag 4, punkt H,6)”<sup>61</sup>

Da arkiveringsversionerne er udformet med vægt på data, er det naturligvis relevant at registrere, hvem der har skabt data, men for at give senere brugere mulighed for at følge etablering og drift af systemet, er det også relevant at registrere, hvem der har skabt systemet. Det er ikke nyt, således ses der mange eksempler på, at der i navngivningen af journaler indgår leverandørens navn, f.eks. Finansministeriets ”E-journal, Scanjour” (0004:E003).

Det kan diskuteres, om det kun er relevant at registrere systemansvarlig, hvis det er to forskellige *myndigheder*, der har ansvar for henholdsvis system og data? Med den stadig større privatisering må Statens Arkiver overveje at acceptere, at registrering af private firmaer i stigende grad bliver relevant?

Registrering af denne oplysning er især af *beskrivende* karakter – oplysningen kan medvirke til at gøre identifikationen af et e-arkivalie mere sikker. Desuden vil oplysningen give brugeren mulighed for at søge videre efter yderligere materiale om registrets dannelse og vedligeholdelse hos den systemansvarlige. Oplysning om systemansvarlig kan tillige være til gavn for fremtidig format- og strukturkonvertering.

Denne registrering gælder kun den, som har skabt e-systemets struktur. Herudover kan der også stilles spørgsmål om, hvem der har stået for driften af e-systemet (og dermed muligvis for de nødvendige tilretninger undervejs i e-systemets levetid).

### **Myndighedernes egen kvalitetskontrol**

Registrering af myndighedernes egen kvalitetskontrol kan give et fingerpeg om datas troværdighed. Hvis data er omhyggeligt kontrolleret af myndighederne, vil data sandsynligvis være i god overensstemmelse med de forhold i samfundet, som de beskriver (men ikke nødvendigvis i overensstemmelse med, hvad der er blevet indberettet til myndighederne).<sup>62</sup>

---

<sup>61</sup> Ibid.

<sup>62</sup> ”Data capture and validation before transfer to the archive” nævnes som en del af beskrivelsen af e-arkivalier af Shepherd and Smith (2000), side 79.

Det vil være af interesse for en kommende arkivbruger hurtigt at kunne danne sig et overblik over, hvilken form for kvalitetskontrol data har været underkastet. Derimod må en nærmere redegørelse for kvalitetskontrollens udførelse henvises til dokumentationen. Derfor registreres følgende (flere svarmuligheder):

data kontrolleres ikke
systembaseret kontrol
manuel kontrol
kontrol for indholdsmæssige fejl
kontrol for fejl i datatyper o. lign.
andet

## Sprog

Registrering af sprog vil tjene flere formål. Dels vil det afklare om, der i e-arkivaliet anvendes et sprog, som brugeren er kendt med, dels vil angivelse af sprog give et fingerpeg om forholdene ved skabelse af arkivaliet, f.eks. om det er skabt i en international kontekst, hvad der vil være en stigende sandsynlighed for. Med den stigende integration mellem landene i EU er det sandsynligt, at et stigende antal registre vil være affattet på engelsk, mens det næppe er sandsynligt, at andre sprog vil være bærende i e-systemer fra danske myndigheder.

Spørgsmålet om sprog rejser også spørgsmålet, om metadata, feltværdier m.m. nødvendigvis skal være affattet på dansk.

Sprog registreres fra en liste, hvor "dansk" er angivet som default-værdi, men hvor der tillige er mulighed for at vælge "engelsk" og "andet".

## Publikation af og om data

I selve registreringen angives det blot i et *ja/nej-felt*, om der findes en publikationsliste. E-systemets data kan være anvendt i forbindelse med publikation allerede før aflevering til arkivet, og det kan være nyttigt for brugeren at blive bekendt med sådanne publikationer for i disse at kunne søge videre oplysninger om e-arkivaliet og dets data. I nogle tilfælde vil brugerens spørgsmål måske kunne besvares af publiceret materiale, fordi brugerens emne allerede er behandlet i litteraturen – eller måske har arkivet valgt at kassere e-systemet netop med henvisning til, at data er publiceret. Relationen kan også gå den anden vej: Kan oplysninger i publikationer uddybes gennem arkivstudier?

Det må gøres klart, at der ikke kan blive tale om en til enhver tid udtømmende litteraturliste (som henvises til bibliotekssystemet), men udelukkende om publikationer, som 1) er udkommet, når e-arkivaliet afleveres, 2) bygger direkte på e-arkivaliets data, og 3) er kendt af den afleveren-

de myndighed. Men også en liste, som ikke er komplet, kan rumme værdifulde oplysninger for en kommende bruger.

### **Visning af dokumentation fra geninfo**

I tilknytning til registreringen af de ovennævnte data om e-arkivaliet bør der etableres mulighed for at vise de umiddelbart tilgængelige dokumenter i dokumentationen af e-arkivaliet (se nærmere om disse under ”dokumentation”). Det foreslås, at der oprettes en liste over de enkelte dokumenters forfatter, dato, titel, kategori og omfang, at denne kan åbnes umiddelbart fra søgevinduet, og at de tilgængelige dokumenter kan åbnes fra denne liste (se ”Indsamling af dokumentation”).



## **Dokumentation**

Der skelnes mellem registrering og dokumentation. *Registrering* er det, som lægges ind i den (eller de) database(r), hvor oplysninger om arkivalier og aktører opbevares (f.eks. Daisy) med henblik på at sætte brugerne i stand til at fremsøge de relevante arkivalier og få tilstrækkelige oplysninger til at kunne sandsynliggøre, at et givent arkivalie indeholder netop de oplysninger, som efterspørges.

*Dokumentation* er de dokumenter, der giver den umiddelbare baggrundsviden om e-arkivaliet. Formålet med dokumentation er at give oplysning om forhold, der er nødvendige for en vurdering af validiteten i arkivaliets oplysninger, f.eks. forhold vedrørende indsamling og brug af oplysninger. Dokumentation kan desuden betragtes som en udvidet beskrivelse. Dokumentation skal støtte kildekritisk arbejde, men har ikke heuristiske formål, og skal først bruges, når det efterspurgte e-arkivalie er identificeret, og dets oplysninger skal fortolkes. Dokumentation er desuden en forudsætning for at tilvejebringe en fyldestgørende og korrekt heuristisk registrering.

Dokumentation kan være skabt af dataskaber, -bruger og -ansvarlig, mens e-systemet stadig var i brug (f.eks. kravspecifikation, brugervejledning, aftale om overførsel af data m.m.), af den afleverende myndighed i forbindelse med afleveringsprocessen (f.eks. i form af notater, der understøtter dokumentationsskemaets oplysninger) eller af arkivet i forbindelse med godkendelses- og afleveringssagsbehandling (bevarings- og kassationsbestemmelser, indstillinger m.m.).

Det skal bemærkes, at der i det efterfølgende *kun behandles spørgsmål vedrørende dokumentation af administrationshistoriske aspekter*, mens den vigtige dokumentation af e-arkivaliers tekniske aspekter ikke behandles her. Det må dog antages, at håndteringen af teknisk dokumentation tilvejebragt ved afleveringen i det store og hele vil kunne følge de retningslinier, som angives for administrationshistorisk dokumentation.

### **Processuelle aspekter – papir- og e-arkivalier**

Alle arkivalier adskiller sig fra forvaltningsakter – og mange andre dokumenter – ved at være løsrevet i tid og rum fra den sammenhæng, de er skabt i. For at kompensere for dette har det længe været anerkendt i arkivteori, at det er arkivets opgave at påse, at der bevares netop den type arkivalier, som kan belyse forholdene ved arkivaliets skabelse. I den administrationshistoriske dokumentation af e-arkivalier findes oplysninger af den type, som i fordanskningen af Schellenbergs terminologi kaldes for ”processuelle aspekter”. I den danske vejledning ”Kassation og bevaring af arkivalier” fra 1979 lægges der vægt på, at bevaring af arkivalier med processuelle

aspekter skal sikre den kommende arkivbruger viden om ”den arkivskabende institutions virksomhed.”<sup>63</sup> Vigtigheden af at bevare arkivalier, der belyser de processuelle aspekter, understreges med påpegnings af, at ”resultatet af institutionens virksomhed findes andre steder, men baggrunden for beslutningen findes kun i den arkivskabende institution.” Der lægges vægt på at dokumentere forhold som institutionens oprettelse og institutionens politik. Derimod hedder det om de interne arbejdsgange i institutionen:

”Institutionens procedure belyses som regel tilstrækkeligt gennem samlingen af administrative forskrifter, ved præcedensskabende sager og ved de arkivalier, der bevares af hensyn til det informative aspekt.”<sup>64</sup>

I 1979 kunne man nemlig stadig koncentrere interessen om papirarkivalier, der ofte i sig selv vidner om deres brug. Papirarkivalier kan være nærmest ”selvforklarende”, idet det vil være tydeligt, om de er skabt i én arbejdsgang eller i flere tempi (forskellig håndskrift m.m.) ligesom rettelser og sletninger oftest vil fremgå tydeligt (overstregning, tilføjelser m.m.). Derfor var det ikke nødvendigt at belyse institutionens procedurer nærmere end, hvad der var fastsat i overordnede retningslinier.

De generelle fordringer til bevaring af arkivalier med processuelle aspekter, som blev opstillet i ”Kassation og bevaring af arkivalier” i 1979, er stadig gældende. Samtidig betyder den systemuafhængige bevaring af e-arkivalier dog, at for at kunne indfri en forventning om at kunne forstå i hvilken kontekst e-arkivalierne i sin tid blev skabt og anvendt, må der bevares arkivalier (”dokumentation”), der gør rede for institutionernes procedurer.

### **Tidligere diskussion om administrationshistorisk dokumentation**

Det er langt fra at være et nyt forslag, at e-arkivalier skal ledsages af administrationshistorisk dokumentation. I et udkast til ”Statens Arkivers krav til elektroniske arkiver og anmeldelse heraf”, dateret 9. maj 1997, hedder det f.eks., at ”generelle oplysninger” (Geninfo) skal indeholde:

”... oplysninger, der har væsentlig betydning for arkivets indhold og forståelsen heraf: oplysninger om brugen af systemet (brugervejledninger mv.), om ændringer i systemet i den periode afleveringsversionen dækker, tilsvarende om ændringer i myndighedens ressort og organisation (herunder om dele af myndigh. ej har anvendt systemet), samt en overleveringshistorik for arkivet, hvis dette (eller dele heraf) er overført fra andre myndigheder til den afleveren-

---

<sup>63</sup> *Kassation og bevaring af arkivalier*. En intern vejledning. Rigsarkivets 2. afdeling. 1979. Side 23

<sup>64</sup> *Ibid.*

de myndighed - i den udstrækning dette har betydning for arkivets indhold og forståelsen heraf.”<sup>65</sup>

Heller ikke den idé at opmærke de dokumenter, som indgår i geninfo, som senere skal præsenteres, er ny. Således finder vi i ”Arkivversion til indlevering af Elektroniske Arkivalier. Statens Arkivers tekniske kravspecifikation. 23. december 1998” en anvisning på udformning af den tabel, der skal danne indholdsfortegnelsen til dokumenter i generel information. Heri er felter til dokumenttitel, kort beskrivelse af dokumentets indhold, person eller organisation, der er ansvarlig for det intellektuelle indhold af dokumentet, dato for dokumentoprettelse og henvisning til dokumentets filbibliotek.<sup>66</sup>

I forbindelse med arbejdet med kravrevision i 1999 blev udformning af og begrundelser for Generelle Informationer diskuteret. Der fandtes korte overordnede beskrivelser af registre i afleveringsfortegnelserne, men i 1999 var det et spørgsmål, hvordan de kunne overføres til eller integreres i Daisy. Kunne de standardiseres til nogle få klare spørgsmål? Ligeledes blev det fremhævet, at kun med klare krav til myndighederne ville man kunne opnå et ensartet niveau for den administrationshistoriske dokumentation, og at i og med at sådanne krav blev stillet, måtte det også kontrolleres, om kravene var overholdt ved aflevering. Det blev således diskuteret, om det skulle stilles krav om, at særlige områder og dimensioner blev dækket gennem den administrationshistoriske dokumentation, og om myndighederne kunne pålægges at udarbejde særlige dokumenter til generel information.<sup>67</sup> Overvejelserne blev opsummeret i et mødereferat fra foråret 1999:

”Det essentielle ved krav om generelle oplysninger, er at vi ikke stiller krav, der tvinger myndighederne til et merarbejde. Desuden skal vores egen kontrol af ovenstående begrænses til hvorvidt elementerne er afleveret eller ej. Hvis der mangler, skal vi spørge om det er en forglemmelse eller om de ikke eksisterer. Der skal ikke foretages en indholdsmæssig vurdering.”<sup>68</sup>

Erfaringen har imidlertid vist, at denne kurs resulterer i *særdeles forskelligartet* dokumentation fra e-arkivalie til e-arkivalie, og tilmed ofte i *utilstrækkelig* dokumentation. Denne rapport foreslår om administrationshistorisk dokumentation af e-arkivalier gør derfor op med den vurdering, der udtrykkes i ovennævnte citat fra 1999. I denne rapport foreslås det, at arkivet netop skal fore-

---

<sup>65</sup> Udkast til ”Statens Arkivers krav til elektroniske arkiver og anmeldelse heraf”, dateret 9. maj 1997. SM/- . Dokumentet har jnr. 121-10, men er fundet som bilag i sagen med jnr. 1997-225-0004(9).

<sup>66</sup> Jnr. 1998-227-0002(12)

<sup>67</sup> Diverse notater i jnr. 1999-220-0006

<sup>68</sup> Referat af 2. revisionsmøde – Afleveringsregler. 22.4.1999 og 4.5.1999. Jnr. 1999-220-0006(17)

tage en indholdsmæssig vurdering, idet bestemte forhold skal dokumenteres, enten ved materiale, som allerede findes hos myndigheden – eller ved materiale, som myndigheden fremstiller til formålet. At fremskaffe dette materiale kan indebære et merarbejde for myndighederne, hvorfor det er en fordel, hvis det er specificeret i lovhjemlen, hvad dokumentationen skal omfatte.

### **Hvad skal myndighederne dokumentere?**

I modsætning til heuristisk registrering, som må være fuldstændig for at give mulighed for en sikker søgning, må kravet til dokumentation af de administrationshistoriske aspekter ved et e-arkivalie være, at den er *så fuldstændig som muligt*. Dokumentation af e-arkivaliernes processuelle aspekter må nødvendigvis være heterogen, og der må accepteres varierende kvalitet fra e-arkivalie til e-arkivalie, idet myndighedernes ”produktion” af egnet dokumentation vil være forskellig. Grundstammen i dokumentation af e-arkivalier bør findes hos myndighederne selv i form af de dokumenter, som de har skabt i forbindelse med udvikling og brug af e-systemet. Er vigtige dele af udviklings- og brugsprocesserne sket uden at have efterladt sig fyldestgørende dokumentation, bør der imidlertid være mulighed for at udfylde de ”tomme huller” med notater, forfattet af myndighederne og skrevet specielt til arkivet i afleveringssituationen.

Ved siden af den myndighedsskabte dokumentation findes også den arkivskabte dokumentation, altså de dokumenter, som arkivet skaber i sin sagsbehandling. Det vil typisk være Bevarings- og kassationsbestemmelser og de tilhørende notater samt Afleveringsbestemmelsen med tilhørende notater. Hvis der findes mere overordnede analyser af myndighedernes arbejdsområde og -gang, f.eks. sektoranalyser, bør de også indgå i dokumentationen.

Der bør i lovhjemlen fastsættes et mindstemål for, hvilke forhold der skal være dokumenteret for et e-arkivalie. Hvis disse forhold ikke kan dokumenteres med arkivalier skabt af myndighederne eller arkivet, bør myndighederne udarbejde særlige dokumenter direkte til brug for arkivet om sådanne forhold.

Nedenfor uddybes det, hvorfor de enkelte punkter bør dokumenteres, og der peges på visse dilemmaer og problemer, som kan opstå, når den rette dokumentation skal udpeges.

### **E-systemets formål**

*Dokumenteres f.eks. med kravspecifikation, beslutning om oprettelse af e-arkivaliet, indledning til brugervejledning m.m.*

Dokumentation af e-systemets formål skal underbygge og nuancere den angivelse af e-arkivaliets formål, som er angivet i registreringen. Forhold som baggrunden for e-systemets oprettelse og

redegørelse for hvilke behov e-systemet skal opfylde, skal belyses i dokumentationen. Der kan findes flere formuleringer af e-systemets formål, til forskellige målgrupper. Bemærk at der kan ske ændringer i e-systemets formål, mens det er i brug. Sådanne ændringer skal dokumenteres.

### **E-systemets lovgrundlag**

*Dokumenteres med det relevante lovstof eller med en liste over dette.*

E-systemet kan være oprettet med grundlag i lovtekster, enten i love som decideret lægger rammerne for det pågældende system eller i love, som pålægger myndigheden at behandle data på en bestemt måde – og hvor dette sker i e-systemet. Disse lovtekster vil på en autoritativ måde afgrænse og bestemme e-systemets formål og i mange tilfælde også brugen af data. Geninfo (eller listen over relevant lovgivning) skal ikke indeholde persondataloven og lignende generelle bestemmelser, men kun de bestemmelser der omhandler netop dette IT-system. Placeres disse lovtekster i Geninfo, vil brugere altid have dem lige ved hånden. Alternativt kan man overveje blot at bede den afleverende myndighed udarbejde en liste over den relevante lovgivning.

### **E-systemets indhold**

*Dokumenteres med brugervejledninger, beskrivelser af e-arkivaliet. Særlige begreber skal forklares gennem dokumentation eller referenceliste*

Ved e-systemets indhold forstås de data, som findes struktureret i tabeller og felter. De findes beskrevet i den tekniske dokumentation, men må også gerne dokumenteres gennem mere administrativt rettet dokumentation, der f.eks. kan diskutere, hvad der skal medtages i databasen, og hvordan de enkelte grupper af data skal afgrænses og defineres. E-systemets indhold afspejles i registreringen af den centrale population og de vigtigste variabler.

Mange e-systemer vil beskæftige sig med stærkt specialiserede områder, og der vil i mange tilfælde blive brugt specialiserede fagudtryk – eller ligefrem udtryk, som er særligt udviklet til e-systemet. Identifikation af sådanne fagudtryk og begreber kan i nogle tilfælde ske gennem kodelister og opslagstabeller. Det er ønskeligt, at betydningen af særlige fagudtryk og begreber udviklet direkte til e-systemet dokumenteres, evt. med henvisninger til uddybende litteratur. Særlige begreber og udtryk afspejles ikke i registreringen.

Journalplaner for e-journaler og esdh-systemer kan også betragtes som en slags ”kodeliste”, og disse skal være tilgængelige i dokumentationen.

(Eksempel: Råstofadministrationsregister 2394:E001. Her findes et notat fra den afleverende myndighed, hvor det forklares, hvad de enkelte variabler betyder: ”data om indvindingen”, ”nyt-

tiggørelse” og ”klapning” – specielt er det vel nyttigt lige at få en forklaring på, hvad ”klapning” er.)

### **Administrative funktioner**

*Dokumenteres med beslutninger og procedurer vedr. brugen af data til forskellige formål, anmeldelse til Datatilsynet (hvis e-systemet er anmeldt) m.m.*

Der skelnes i forvejen mellem journaler og andre e-systemer på alle niveauer i Statens Arkivers sagsbehandling, men det bør overvejes, om det kan være hensigtsmæssigt at udskille endnu flere grupper af administrative funktioner med henblik på at udarbejde særlige krav til dokumentationen for forskellige administrative funktioner. Det vil dog kræve en liste over administrative funktioner, udviklet på grundlag af casestudies i myndighedernes administration, hvilket ligger uden for nærværende projekts rammer. Det skal bemærkes, at forskellige myndigheder eller forskellige dele af samme myndighed kan anvende det samme e-system til forskellige formål, og måske også på forskellige tidspunkter i løbet af e-systemets levetid. Fokus for denne del af undersøgelsen er altså på brugen af e-systemet og dermed også på distribution af data.

I oplægget til pilotprojekt i projekt A og B (se bilag 4, 5 og 6) peges der på en række administrative funktioner, som kan varetages gennem e-systemer. Listen af sådanne administrative funktioner vil uden tvivl blive udbygget, efterhånden som der indhøstes erfaringer med brug af dokumentationsskemaet.

Anmeldelse til Datatilsynet indeholder en række væsentlige oplysninger om e-systemets administrative funktioner, og bør derfor altid indgå i dokumentationen for de e-systemer, som er anmeldt.

### **Foreløbig liste over administrative funktioner:**

Er afgørelser truffet alene på grundlag af IT-systemets oplysninger?	Spørgsmålet er taget fra Datatilsynets anmeldelsesskema. Svares der ja, skal man være ekstra opmærksom på at få tilstrækkelig teknisk dokumentation, da væsentlige dele af sagsbehandlingen så foregår ”inden i” maskinen.
Tilsyn og kontrol	Hvis formålet med e-systemet er tilsyn og kontrol, vil det sikkert betyde, at data dels er meget pålidelige, dels kan være ret summariske. Det er vigtigt at få dokumenteret, hvad der er ført tilsyn eller kontrol med, hvordan dette er udøvet og med hvilken hjemmel.
Statistik	Statistikoplysninger vil ofte være videregivet til andre myndigheder, ligesom de ofte vil være ”normaliseret”. Begge dele skal dokumenteres.
Henvisning til sager på papir	Altså hvad vi kalder en journal. Vær opmærksom på, at også registre kan have journal-funktioner. Det dokumenteres, hvilke sager der henvises til og hvordan henvisningerne sker.
Henvisning til elektroniske sager	Altså hvad vi kalder et ESDH-system, men måske også henvisninger mellem flere elektroniske systemer.
Basisregister	Et basisregister er et register, der har som sin absolutte hovedopgave at indsamle basisoplysninger til brug for andre myndigheders sagsbehandling. Det er f.eks. registre som cpr-registret, cvr-registret og BBR. Her findes ”nøglerne” til nogle nationale og globale identifikatorer.
Forskning og udredning	Kan også defineres negativt, idet det ikke er sagsbehandling. Det skal dokumenteres, hvordan oplysningerne er brugt, f.eks. publiceret.

Internetpublicering / Internetportal	Er e-systemernes oplysninger (eller dele deraf) til rådighed på Internet? Nogle myndigheder har det som en del af deres opgave at stille oplysninger til rådighed, enten helt offentligt eller til en afgrænset gruppe. Det skal dokumenteres, hvilke af e-systemets oplysninger der indgår i en sådan offentliggørelse og hvilken form denne har, f.eks. gennem skærmdumps. Det skal endvidere dokumenteres, om data fra dette e-system præsenteres i sammenhæng med data fra andre e-systemer.
Udskrift af attester, beviser m.m.	Attester med retslig eller anden væsentlig betydning uden for myndigheden, f.eks. eksamensbeviser, straffeattester, børneattester. Det må sikres, at disse kan rekonstrueres ud fra de bevarede data, altså dokumenteres hvordan attesterne er dannet og nøjagtigt hvad de har skullet bevise eller attestere.
Registrering af genstandsmateriale el. l.	Hvorfor er det væsentligt at få disse genstande registreret?
Sagsbehandling	Varetagelse af myndighedens ressortområde

## Præsentationsstruktur

*Dokumenteres med brugervejledninger (dækker indtastning og opslag), redegørelser for systemets funktionaliteter (bl.a. beregninger, stored procedures, applikationer), skærmdumps af de mest anvendte (alle?) skærbilleder osv.*

Dokumentation af e-systemernes præsentationsstruktur udgør et grænseområde mellem denne rapports administrationshistoriske fokus og et andet fokus af mere teknisk karakter. Som beskrevet tidligere i denne rapport udgør præsentationsstrukturen myndighedernes faktiske brug af e-systemet, som det er blevet muliggjort med applikationer, stored procedures (beregninger) m.m. *Hensigten* med at dokumentere dette er at give senere arkivbrugere viden om, hvordan data faktisk er blevet anvendt, mens e-systemet blev brugt af myndighederne, og at give senere arkivbrugere mulighed for at genskabe et rimeligt nøjagtigt billede af, hvilke beslutninger myndigheden traf og på hvilket grundlag de blev truffet.

Derimod kommer denne rapport til kort, når det gælder at foreslå med *hvilke midler* denne beskrivelse kan ske. Det skal dog nævnes, at andre har foreslået at dokumentere e-systemernes præsentationsstruktur med skærmdumps, der indeholder henvisning fra skærbilledernes felter til konkrete felter i databasen og de oftest anvendte søgninger udtrykt i SQL og med ofte anvendte rapporter. Problematikken er diskuteret i Lone Smith Jespersens rapport, side 24-28, hvor det konkluderes, at følgende bør bevares:

- ”– Bevaring af dokumentation af de dele af applikationen, der har skabt dataelementerne ved på brugergrænsefladeniveau at definere struktur.
- Bevaring af dokumentation af de dele af software, der har haft betydning for data i databasen. Det vil sige de beregninger, der foretages automatisk, og de dataoverførsler der sker fra andre systemer.”<sup>69</sup>

<sup>69</sup> Lone Smith Jespersen (2005) side 27

### **Sletning eller overskrivning af data i e-systemet**

Hvis der sker rutinemæssig sletning af data i e-systemet, tilstræbes det, at bevaringsbestemmelsen tager højde for det, således at ingen data går tabt af den grund. Udover bevaring af data, også hvis de slettes eller overskrives af myndighederne, bør det også dokumenteres, hvorfor og hvornår der er sket sletninger eller overskrivning, da viden herom er vigtig for at kunne få et overblik over beslutninger og beslutningsgrundlag hos de enkelte myndigheder.

### **Læsning af data efter aflevering**

I enkelte tilfælde kan det forekomme, at en myndighed ikke længere opdaterer et e-system, men stadig anvender dets data administrativt, til opslag el.lign. Hvis e-systemet afleveres, fordi dets data ikke længere opdateres, men e-systemet stadig anvendes til andre formål, skal det ved aflevering dokumenteres til hvilke formål det afleveres efter aflevering, da viden herom kan have betydning for en senere vurdering af datas udsagnskraft.

### **Tilvejebringelse og videregivelse af data**

*Tilvejebringelse dokumenteres med instrukser og brugervejledninger for alle implicerede myndigheder, vejledningsmateriale til borgere/virksomheder/foreninger samt blanketter og skemaer, som er anvendt ved indberetning af data.*

*Videregivelse dokumenteres gennem aftaler og bestemmelser om flere myndigheders brug af e-systemets data og instrukser for, hvordan dette gennemføres. (På den tekniske side skal eventuelle konverteringer ved overførsel af data dokumenteres)*

Viden om, hvordan og med hvilket formål data er indsamlet, er særdeles væsentlig for en senere vurdering af datas udsagnskraft. Det kan dreje sig om oplysninger om, hvilke myndigheder der har nedlagt data i e-systemet, om e-systemets data er overført fra andre systemer eller hentet ved opslag i andre e-systemer (og hvor har disse så fået deres data fra?) eller om borgere (virksomheder, foreninger) har haft lejlighed til at lægge data direkte i e-systemet via særlige indtastningsmuligheder på Internettet. Efter hvilke retningslinier er data blevet indtastet/indlæst? Er der sket redigering og kvalitetskontrol?

Det er af afgørende betydning for en forskningsmæssig udnyttelse af data at have rede på ophavssituationen for e-arkivaliernes data.

Data i e-systemer vil altid være tilvejebragt som resultat af nogle bestemte arbejdsgange, der i mange tilfælde også vil være beskrevet i brugervejledninger og andet internt materiale, som bør indgå i dokumentationen. Hvis data er tilvejebragt ved indtastning skal retningslinier for indtast-



ning (brugervejledning) fremgå af dokumentationen. Findes der forskellige retningslinier for forskellige grupper, skal de alle indgå i dokumentationen.

Endelig skal det gennem dokumentationen kunne påvises, hvorfor (med hvilken berettigelse) de enkelte grupper (myndigheders afdelinger og kontorer, grupper af borgere) kan indtaste i dette IT-system.

Det skal fremgå af dokumentationen, hvilke data der er tale om, fra hvilken myndighed der er overført data og fra hvilket system hos denne myndighed. Samt om der eventuelt er sket en bearbejdning af oplysningerne ved overførslen. (Dette gælder dog ikke for standardidentifikatorerne fra basissystemer f.eks. cpr-nummer)

Ved anmeldelsen af systemet bør arkivet sikre sig viden om, hvordan myndigheden påtænker at indsamle data, og i den forbindelse rådgive om, hvordan denne indsamling skal dokumenteres. Allerede ved anmeldelse bør brugervejledninger med videre indsamles, dels for at sikre at de bevares til dokumentation af e-arkivalier, dels for senere at kunne dokumentere ændringer heri. Ved aflevering skal arkivet sikre, at det dokumenteres, hvis der er sket indsamling af data på måder, som ikke var forudset ved e-systemets ibrugtagning, herunder ved ændringer i de oprindelige retningslinier.

Hvis data er kommet ind i systemet via en blanket, skal denne blanket indgå i dokumentationen. Hvis data er kommet ind via indtastning på Internet, skal skærbilleder og instruktion indgå i dokumentationen.

Hvis data er overført fra et andet e-system, skal det dokumenteres, om det er alle oplysninger i det afleverende system, der er overført – eller eventuelt hvilket udsnit af oplysninger. Tilsvarende skal det fremgå af dokumentationen, om der er tilført nye oplysninger i forbindelse med overføring. Endelig kan data i enkelte tilfælde gives videre til andre myndigheder i papirform (som udskrift). Dette skal også dokumenteres.

Hvis data bevares på et højere administrativt niveau, vil data i mange tilfælde blive kasseret på lavere administrative niveauer, hvor data dog i sin tid er blevet modtaget. Når data bevares på et højere administrativt niveau, kan det således ikke umiddelbart dokumenteres, hvordan data er indsamlet. Det bør derfor overvejes at bevare dokumentation om, hvordan data indsamles for systemer, som i øvrigt kasseres? *Princippet om at bevare på højeste aggregeringsniveau giver store problemer set fra en dokumentationssynsvinkel.*

### **Forgænger- og efterfølgerregistre**

Det skal dokumenteres, hvilke data og funktioner, der er fælles for registret og dets forgænger/efterfølger, om der er sket overførsel af data og hvilken konvertering der eventuelt er sket i den forbindelse.

Dokumentation af, hvordan e-systemer afløser hinanden (og afløser papirregistre), vil være en hjælp til dels at etablere tidsserier af data til brug for kvantitative studier, et område hvor e-arkivalier har særlige fortrin frem for papirarkivalier, dels til at følge udviklingen af administrative procedurer. Oplysningerne om fælles funktioner, fælles datagrupper og eventuel konvertering mellem forgænger- og efterfølgersystemer skal dokumenteres. Hvis dannelses- og oplysningstid (jf. Daisys begreber) ikke er identiske, skal det dokumenteres, hvordan og hvorfra oplysninger, som er ældre end selve e-systemet, er kommet ind i e-systemet.

### **Myndighedernes egen kvalitetskontrol**

Det er ikke tilstrækkeligt at foretage registrering af, om myndighederne har foretaget kvalitetskontrol. Det må tillige dokumenteres, hvordan denne kontrol er udført, hvilke dele af data der er blevet kontrolleret og om der kun har været tale om en maskinel kontrol – eller om kontrol af data f.eks. også har været en del af myndighedernes sagsbehandling.

### **Publikation af og om data**

Oplysninger om, hvorvidt data er publiceret, burde kunne indhentes i forbindelse med aflevering af e-arkivaliet – der er tale om en viden, som burde være let tilgængelig hos den afleverende myndighed. Publiceret materiale vil ikke blot kunne præsentere (dele af) data på en overskuelig måde, men vil også kunne give værdifulde oplysninger om e-systemets brug.

## Typer af publikation af og om data

Publikationstype	Publikation af data	Publikation om data		Oplysning om publikation i arkiveringsversion	
		Forfatter er daskaber, -bruger eller -ansvarlig	Forfatter er forsker <sup>70</sup>	litteraturliste	Indgår med hele teksten i geninfo
Papir med ISBN/ISSN-nummer	Årsberetninger, statistikker m.m.	Årsberetninger, fortolkninger af data (undersøgelser) m.m.	Fortolkninger af data (undersøgelser) m.m.	Ja	Ikke nødvendigvis
Elektronisk (pligtafleveret)	Do.	Do.	Do.	Ja	Ikke nødvendigvis
Papir uden ISBN-nummer, rummer væsentlige oplysninger	Specialer, interne rapporter m.m.	Interne rapporter m.m.	Specialer m.m.	Ja	Vurderes i de enkelte tilfælde
Elektronisk (ikke pligtafleveret), rummer væsentlige oplysninger	Interne rapporter	Interne rapporter	Artikler, præsentationer, kommentarer m.m.	Ja	Vurderes i de enkelte tilfælde

For publikationer, som er underlagt bestemmelserne om pligtaflevering til Det Kongelige Bibliotek, kan det være tilstrækkeligt, at arkiveringsversionen indeholder en liste med fulde bibliografiske data, så en bruger senere søge materialet hos bibliotekerne. En bedre service – og nærmere målet om at arkiveringsversioner skal være selvdokumenterende – vil det naturligvis være at lade elektroniske versioner af disse publikationer indgå i arkiveringsversionen. Publikationer, som ikke er underlagt bestemmelserne om pligtaflevering, bør derimod altid indgå i arkiveringsversionen, hvis det skønnes, at de indeholder væsentlige oplysninger om e-systemet, dets data og brug. I forbindelse med internet-sider må det tages i betragtning, at internet-sider ikke bevares systematisk.

Det bør fremgå af dokumentationsskemaet, hvilke publikationer om e-systemet den afleverende myndighed har kendskab til. En sådan publikationsliste vil skulle indgå i arkiveringsversionens geninfo, og den vil være umiddelbart tilgængelig. Den bør derfor kunne tilgås umiddelbart fra visningen af registreringen (Daisy eller TGP). Om de enkelte publikationer også vil være umiddelbart tilgængelige må afhænge af en nærmere vurdering.

<sup>70</sup> Oplysninger om forfatter til publikationerne kan være en væsentlig oplysning ved vurdering af litteraturens relevans. Hvis forfatteren er den myndighed, som har indsamlet materialet, kan synsvinklen være helt anderledes, end hvis forfatteren er en forsker. Denne skelnen nævnes også af Shepherd og Smith (2000), side 82.

## **Udvikling og drift af e-systemet**

*Dokumenteres med tekniske beskrivelser af systemet samt vedligeholdelsesbestemmelser m.m.*

Dokumentation af tekniske forandringer i systemets levetid (nye funktionaliteter, nye skærmbil-  
leder, ændringer i databasestruktur, konverteringsoplysninger, skift af tegnsæt). Har ressortæn-  
dringer medført tekniske ændringer? Der fokuseres her på e-systemernes IT-side, ikke på forret-  
ningssiden, som vil blive dokumenteret gennem de øvrige spørgsmål.

## **Supplerende forslag og overvejelser**

Til slut skal der peges på forskellige aspekter, som vil kræve særlige overvejelser – og måske ændringer – for at forslaget om registrering og dokumentation af e-arkivalier kan gennemføres. Disse supplerende forslag og overvejelser må således medtænkes i en samlet vurdering af forslaget.

### **DOK-IT: Pilotprojekt om indsamling af dokumentation**

Det er åbenlyst, at der bør indhentes mere solid viden om, hvordan myndighedernes dokumentation kan være udformet, hvordan kravene bør beskrives over for myndighederne, hvilken form den har osv. osv. For at indhøste erfaringer om dette vil projekt A og B gennemføre et pilotprojekt med Lone Smith Jespersen som projektleder. Projektet skal omfatte (stort set) alle afleveringer fra juni 2007 og ca. et år frem. Ved hjælp af dokumentationsskemaet med tilhørende vejledning (se bilag 4 og 5) skal kommende afleveringer dokumenteres efter retningslinierne i dette forslag, med de tilretninger, som indhøstede erfaringer dikterer. Dokumentationsskemaet kan til lige bruges som hjælpeskema ved anmeldelse og BK-behandling. Som bilag 6 findes projektbeskrivelse for DOK-IT.

### **Opdeling af Geninfo**

Geninfo må i dag betragtes som en ”rodekasse” i arkiveringsversionen. Heri placeres al dokumentation for pågældende e-arkivalie uden nogen opmærkning el.lign. Dokumentationen omfatter imidlertid væsensforskellige forhold, og vil f.eks. kunne opdeles efter forfatterkategori (arkiv, myndighed) og indhold (administrationshistorisk eller teknisk dokumentation).<sup>71</sup> I skemaet nedenfor angives et forslag til inddeling af geninfo, men der tages ikke stilling til, hvordan dette teknisk skal gennemføres: som opmærkning, egentlig opdeling eller noget tredje.

Geninfo kan bestå af – og består af – et varieret materiale: nogle dele er forfattet af arkivet, andet af de afleverende myndigheder. Meget dokumentation fra myndighederne vil være skabt for at understøtte myndighedernes egen virksomhed, det kan f.eks. være materiale vedrørende datas oprindelse og brug. En del af dette materiale vil internt i myndighederne være, hvad man kan kalde ”internt, alment tilgængeligt”, idet det f.eks. kan blive udleveret til alle medarbejdere eller kan ligge på Intranet; det kan f.eks. være brugervejledninger. Andet materiale kan være dokumenter, som er nærmere knyttet til e-systemets oprettelse og drift, f.eks. aftaler om udveksling af data mellem myndigheder. Dokumentationen kan også rumme materiale, som allerede er alment

---

<sup>71</sup> Denne opdeling er hentet fra Lone Smith Jespersen (2005) side 47-49 og 56-62

tilgængeligt, f.eks. lovstof og publikationer. Endelig kan dokumentationen rumme materiale, som er skabt direkte til afleveringssituationen, det vil typisk være alt materiale fra arkivet samt notater m.m., som myndighederne har skrevet på arkivets foranledning.

### Forslag til opdeling af Geninfo

Dato og forfatter for alle dokumenter	Forfatter-kategori	Formål	Administrationshistorisk dokumentation (eksempler)	Teknisk dokumentation (eksempler)	Umiddelbart tilgængelig efter Arkivloven
	Arkiv	Arkivets sagsbehandling	BK-bestemmelse Afleveringsbestemmelse Oversigt over indhold af geninfo	Digital beskrivelse mm ITV-materiale?	Ja
	Myndigheder	Rettet mod arkivet	Anmeldesskema Dokumentationsskema Evt. notater i tilknytning hertil som supplement til myndighedsskabt, myndighedsrettet dokumentation	ER-diagrammer mm	Ja
	Myndigheder	Rettet mod myndighedens egen virksomhed. Internt alment tilgængeligt	F.eks. brugervejledninger og skærmdumps. Interne rapporter.		Nej
	Myndigheder	Rettet mod myndighedens egen virksomhed.	F.eks. vedr. datas oprindelse (hvor kommer data fra), datas brug (hvem har brugt data) og dokumentation for myndighedens kvalitetssikring af e-systemet	Teknisk dokumentation mm	Nej
	Myndigheder	Alment tilgængeligt	Lovstof, publikationer, blanketter og skemaer m.m.		Ja

Ifølge den gældende fortolkning af Arkivloven regnes arkivskabt dokumentation og dokumentation fra myndighederne, skabt efter arkivets anvisninger, som umiddelbart tilgængelige. Derimod er materiale skabt af myndighederne til deres eget brug først tilgængeligt efter 20 år.<sup>72</sup> Det skal imidlertid foreslås, at det overvejes, om der kan findes mulighed for tillige at gøre materiale skabt af myndighederne umiddelbart tilgængeligt, hvis det drejer sig om materiale, der hos myndighederne var ”internt, alment tilgængeligt” eller måske endda allerede er offentliggjort, før aflevering til arkivet. Brugervejledninger og lovstof kunne være eksempel på en sådanne typer af materiale.

Der bør skabes teknisk mulighed for at vise dokumentationen i tilknytning til søgeresultatet, men med mulighed for at ”lukke for” de dokumenter, som ikke er umiddelbart tilgængelige. I dokumentationen vil brugeren nemt og hurtigt kunne få værdifulde oplysninger om et e-arkivalies indhold og validitet, og dokumentationen kan således uddybe beskrivelsen af e-

<sup>72</sup> Jf. Notat: Tilgængelighedsfrister eller ej for dokumentation i arkiveringsversioner, LSJ. 7. april 2005, rev. 15. april 2005. Journalnummer:2005-550-0001

arkivalierne – måske kan en god dokumentation endda udfylde samme funktion som de arkivskabte essays, der blev foreslået af Tom Nesmith.<sup>73</sup>

Forslaget om at dokumentere flere administrationshistoriske forhold end hidtil og at gøre det mere systematisk og konsekvent rejser spørgsmålet om, i hvilken udstrækning der kan tillades dokumenter i geninfo, som tillige findes andre steder, f.eks. i en myndigheds ESDH-system (journalsagsarkiv) eller i bibliotekerne. Alternativt kunne man forestille sig, at det blev tilladt geninfo at indeholde henvisninger til relevant dokumentation. Det kunne ske som henvisninger til ESDH-systemet eller journalsags-arkivet, som en litteraturliste med fulde bibliografiske oplysninger og som en liste over relevante lovtekster. En sådan praksis ville dog muligvis give problemer med visning af dokumentationen i forbindelse med søgning.

### **Kravrevision og interne procedurer**

I bilag 8 findes en oversigt over et e-arkivalies livscyklus, både som e-system hos myndigheden, i afleveringssituationen hvor det bliver til et e-arkivalie og når det som e-arkivalie skal anvendes af arkivbrugere.

Her skal der gives nogle forslag til ændringer i Statens Arkivers interne procedurer, som kan medvirke til at sikre systematisk dokumentation af administrationshistoriske forhold:

- For at sikre at myndighederne er villige til at indgå i processen med at dokumentere brugen af e-systemer, vil det være en fordel, hvis deres forpligtelse hertil *tydeligt fremgår af bekendtgørelsen om aflevering af e-arkivalier*. Det kunne f.eks. ske som en udpegning af de områder, der bør dokumenteres. Derimod bør man undgå at udpege specifikke dokumenter, som skal indgå i dokumentationen, da de enkelte myndigheder håndterer forskellige dele af processen noget forskelligt. Der kan naturligvis i vejledningsmaterialet gives eksempler på hensigtsmæssige dokumentation.
- Tidspunktet for fastsættelse af hvilken dokumentation, der skal afleveres til det enkelte e-arkivalie, må overvejes nøje. Myndighederne er typisk i kontakt med Statens Arkiver (mindst to gange i forbindelse med hvert e-system, nemlig ved anmeldelse og ved aflevering. Ved anmeldelse af e-systemet (som skal ske før e-systemet tages i brug) må det antages, at dokumentation vedrørende e-systemets opbygning og forventede brug (herunder aftaler med andre myndigheder) vil være i frisk erindring, og det vil derfor være et godt tidspunkt til at indsamle dette materiale. Ofte vil der i løbet af e-systemets levetid ske ændringer i dets brug, og også af

---

<sup>73</sup> Nesmith (2005)

den grund kan det blive nyttigt at indsamle materiale allerede ved systemets ibrugtagning. Det bør så suppleres ved aflevering med materiale, som kan dokumentere de ændringer, der er sket i e-systemets levetid. Det foreslås, at *indsamling af dokumentation sker i forbindelse med sagsbehandling af anmeldelse/godkendelse/BK-behandling og igen i forbindelse med aflevering.*

- For at understrege vigtigheden af, at et e-systems administrative brug dokumenteres, og for at være sikker på at få afleveret så god en dokumentation som muligt, bør det allerede af *bevaringsbestemmelsen* fremgå, at en eventuel bevaring ikke blot omfatter e-systemet, men også den tilhørende dokumentation. Der kan af bestemmelsen fremgå nogle præcise dokumenter, som findes allerede på tidspunktet for anmeldelse af e-systemet, men det er vigtigt, at der også skabes mulighed for, at forhold, som er ændret i perioden fra e-systemets anmeldelse til dets aflevering bliver dokumenteret.
- Endelig må det påpeges, at med denne rapports forslag om at dokumentere, hvordan data er skaffet til veje og anvendt, kan det ikke længere anses for forsvarligt at kassere dokumentation for e-systemer, der kasseres med henvisning til, at data vil blive bevaret på et højere aggregeringsniveau. Det foreslås derfor, at der *bevares dokumentation om datas oprindelse fra de e-systemer, der kasseres med den begrundelse, at deres data bevares i andre systemer.*

### **IT-arkitektur: TGP og Daisy**

Dette forslag tager ikke stilling til IT-arkitektoniske spørgsmål om, hvorvidt registrering skal foregå i Daisy, TGP eller et tredje sted. Det skal dog præciseres, at forslaget stiller nogle krav til IT-arkitekturen:

- Registrering af *interne referenceoplysninger* skal ske i Daisy (eller i tæt tilknytning til Daisy) – som i dag
- Registrering af *informationsnetværk* skal ske i Daisy, eller i tæt tilknytning til Daisy, eller i det mindste med adgang til at knytte sig til aktører og heuristiske enheder, som findes i Daisy. Giver der ikke denne adgang, vil oplysningerne om informationsnetværk ikke kunne anvendes ved søgning. Disse oplysninger må så videregives i notefelter eller lignende.
- Registrering af *beskrivelse af e-arkivaliet* (e-arkivaliets formål og indhold, sprog, kvalitetskontrol m.m.) skal kunne tilknyttes de relevante heuristiske enheder og arkiveringsversioner i Daisy som notefelter og afkrydsningsfelter.



- Der skal være mulighed for at vise de umiddelbart tilgængelige dele af geninfo (*dokumentation*) i tilknytning til visning af beskrivelser af e-arkivalierne
- For maskinelt at kunne udpege de umiddelbart tilgængelige dele af *geninfo* må denne struktureres, så disse kan genkendes
- Visning af hele den arkivskabte beskrivelse (reference, informationsnetværk, beskrivelse og umiddelbart tilgængelig dokumentation) skal både kunne ske i forbindelse med søgning/identifikation (f.eks. i Daisy) og i forbindelse med visning af e-arkivaliet (TGP)

### **Retrospektiv brug af modellen for registrering og dokumentation**

Når der er etableret tilstrækkelige registreringsfaciliteter (i Daisy, TGP eller et tredje sted), vil rapportens forslag kunne anvendes over for e-arkivalier, efterhånden som de afleveres. For at fremtidige arkivbrugere vil kunne få fuldt udbytte af Statens Arkivers samling af e-arkivalier, må også de på dette tidspunkt allerede afleverede e-arkivalier registreres efter samme retningslinier. En sådan retrospektiv registrering vil kræve, at der fra Statens Arkivers journalsager og geninfo fremfindes de relevante oplysninger.

En retrospektiv registrering og beskrivelse af de allerede afleverede e-arkivalier vil i vidt omfang være mulig på baggrund af de oplysninger, som allerede findes i Statens Arkivers journalsager og i e-arkivaliernes geninfo. Derimod vil det næppe være muligt retrospektivt at bringe dokumentationsstandarder op på det foreslåede niveau, men der bør dog åbnes mulighed for at tilføje eventuelt materiale fra journalsagerne til geninfo.

### **Hvordan bevares e-systemernes præsentationsstruktur?**

Som nævnt i rapportens indledning kan e-systemer forstås som sammensat af data, kontekst, præsentationsstruktur og logisk struktur. Bevaring af data og logisk struktur er der taget højde for i den gældende bevaringsbestemmelse, og denne rapport behandler bevaring af kontekst. Derimod er det i dag sådan, at præsentationsstrukturen går tabt ved aflevering, om end der er mulighed for at give en (usystematisk) beskrivelse af præsentationsstrukturen i dokumentationen af e-arkivaliet. Præsentationsstrukturen går tabt som en konsekvens af det teknisk nødvendige valg af konverteringsstrategien, der netop går ud på at bevare data, men ikke funktionalitet.

Konsekvensen af ikke at bevare præsentationsstrukturen er, at det ikke vil være muligt for senere arkivbrugere at undersøge, hvordan forvaltningen har arbejdet med de bevarede data: Hvilke data har været anvendt ved behandlingen af konkrete sager? Er det e-systemets data, som er indgået i sagsbehandling, eller er det beregninger på grundlag af e-systemets data?

### **Hvordan tilpasses registrering og dokumentation fortsat virkelighedens e-systemer?**

Nøje kendskab til myndighedernes arkivdannelse er en forudsætning for Statens Arkivers bevarings- og dokumentationsstrategi og den tilknyttede sagsbehandling, idet arkiverne kun gennem indsigt i den faktiske arkivdannelse hos myndighederne kan udskille de arkivalier, som vil give eftertiden et retvisende billede af myndighedernes virksomhed og af den viden om det danske samfund, som opsamles af myndighederne.

Den stigende brug af IT-baserede processer hos myndighederne betyder, at mange kendte administrative rutiner – også arkiveringsrutiner – undergår omfattende ændringer i disse år, f.eks. betyder den øgede brug af esdh-systemer store omlægninger af myndighedernes administrative rutiner. En anden udfordring, som meget snart vil være presserende, er det forhold, at stadig mere sagsbehandling indebærer ”opslag” i e-systemer hos andre myndigheder. Statens Arkiver har en oplagt interesse i at sikre sig viden om såvel de enkelte elementer i myndighedernes besluttelsesgrundlag i forskellige dele af sagsbehandling og øvrig ressortvaretagelse som om de forskellige arbejdsprocessers slutprodukter (ikke mindst hvor de ”ender”).

Arkiverne har således fået en ny opgave med at følge myndighedernes administrative opgavevaretagelse *uden nogen tidsforskydning*. For at varetage denne opgave må arkiverne for det første sørge for at holde sig informeret om nye tiltag af både it-teknisk og administrativ art hos myndighederne, og for det andet sørge for, at arkivernes egne systemer er så fleksible, at de kan dokumentere og registrere de relevante oplysninger.

## **Bilag - oversigt**

### **Bilag 1: Casestudies – arkivskabte beskrivelser/registreringer**

Centralregistret for testamenter  
Fondsregister  
Kriminalforsorgens Arresthussystem  
Datatilsynet: Anmeldelsessystem og fortegnelse  
Rigspolitichefen: Anmeldelsesregister  
Arealanvendelse og dyrkningspraksis i udvalgte områder  
Biologisk Vandløbskvalitet  
CORINAIR  
Register over brugere af boformer  
Civil-, straffe-, foged-, auktions- og skiftesystem  
Forskningscenter Risøs ESDH-system (LQN)  
Fælles Visum System (LQN)  
Rektorkollegiets e-journal (LQN)  
DPC Projektdatabase (LQN)  
Erhvervssystemet (BVL)  
SMK- ESDH-system (BVL)  
Pensionsafkastbeskatningsloven (BVL)  
Statens Ejendomsinformation (BVL)  
Register over tvang i psykiatrien, tvangsbehandlinger (MHA)  
Maritimt Administrationssystem (MHA)

### **Bilag 2: Scenarium ”Emnesøgning”**

### **Bilag 3: Scenarium ”Søgning med kendt dataleverandør”**

### **Bilag 4: Dokumentations-skema (e-systemer)**

### **Bilag 5: Vejledning til sagsbehandleren for Dokumentations-skema (e-systemer)**

### **Bilag 6: Projektbeskrivelse for DOK-IT**

### **Bilag 7: Eksempler på registrering af e-systemers formål**

### **Bilag 8: E-arkivaliets livscyklus**