

Indhold og tilgængelighed i danske kommunale arkiveringsversioner



Maja Kromann Sørensen
Vejleder: Marianne Rostgaard

Aalborg Universitet
2025

Abstract

This thesis aims to explore the contents and accessibility of archival versions of Danish public sector database systems and lay the foundation for further research in the field. I have examined the structure, content and documentation for three archival versions of a municipal system used in the care of elderly and disabled citizens, preserved at the Aalborg City Archives. Due to the exploratory nature of the study, the results are mixed. I have found that the documentation necessary to interpret the contents of the databases is generally good, but minimal. Due to the Danish preservation strategy, the structure, or original order, of the digital material is largely preserved, save for minor compromises in the early end migrated instances due to technical limitations set by the preservation requirements. The databases also contain a wealth of other information about the municipal administration of care for the elderly and disabled that has yet to be examined. This study leaves many questions open to further research, particularly about the structure and accessibility of the material.

Dette speciale handler om at undersøge, hvad der ligger i danske i danske arkiveringsversioner, og hvor tilgængeligt det er. Jeg har taget udgangspunkt i tre arkiveringsversioner af systemet Care, der ligger hos Aalborg Stadsarkiv. Her har jeg fundet frem til, at det vha. dokumentation og arkivets præ-lavede søgninger generelt er muligt at finde de vigtigste oplysninger om personer i arkiveringsversionerne. Der ligger dog også mange andre oplysninger deri, som er knap så lette at finde og forstå. Der er endnu lang vej igen før vi har fuld forståelse for, hvordan kvalitet, tilgængeliggørelse og brug af arkiveringsversioner fungerer i praksis.

Indholdsfortegnelse

Abstract	2
Indholdsfortegnelse	3
Indledning	4
Problemformulering	4
Rammevilkår og begrænsninger	5
Dansk digital arkiveringsstrategi	6
Lovgivning	9
Arkivloven	9
Bevaring og kassation af digitalt skabte arkivalier i kommunerne	10
Bekendtgørelse om arkiveringsversioner	12
Data og dokumenter	16
Forskningsdiskussion: Det tekniske	19
Visning og søgning	26
ASDA	27
Udvælgelse af arkiveringsversioner	30
Intro til systemet	31
Dokumentation	33
Opsamling	40
Struktur og indhold i arkiveringsversionerne	43
Eksempler på klienter	46
Almindelig ældre	46
Fysisk handicappet person	49
Mentalt handicappet person/"Sindslidende"	51
Ansatte i forvaltningen	54
Andre typer oplysninger	58
Opsamling	59
Tilgængelighed	60
SQL-søgninger i ASDA	60
Søgning på dokumenter	68
Diskussion: Tilgængelighed og brug	69
Andre vinkler på emnet	75
Konklusion	77
Bibliografi	79

Indledning

Dansk digital arkivering har indtil videre handlet mest om selve bevaringen af digitalt materiale. Man har fokuseret på at sikre, at data og dokumenter fra offentlige IT-systemer er blevet konverteret og gemt i langtidsholdbare formater. Hvorledes materialet skal gøres tilgængeligt for brugere af arkiverne, har der derimod ikke været særlig stor fokus på - i hvert fald ikke offentligt nok, til at der er blevet skrevet noget om det før for relativt nyligt. Der er heller endnu ingen, der har undersøgt, hvilke konsekvenser den danske bevaringsstrategi kan have for tilgængeliggørelsen af det digitale materiale, som stort set alt dokumentation fra det offentlige består af er i dag.

Afhængig af hvem, der skal bruge de digitale arkivalier, og hvad de skal bruge dem til, kan tilgængeliggørelse og tilgængelighed defineres lidt forskelligt. Der er flere forskellige brugerperspektiver, der er værd at undersøge (f.eks. arkiverne selv, forvaltningen, personer, der ønsker egen-access, og forskere), og med dette speciale åbner jeg derfor for et meget bredt felt. Det ligger udenfor dette speciales rammer at undersøge arkiveringsversioner ud fra alle perspektiver. Da der stort set intet er skrevet om emnet i forvejen, vil det også være en udfordring kun at vælge et perspektiv fra en enkelt brugergruppe. Derfor vil jeg undersøge, hvad der ligger i arkiveringsversioner, hvordan det ligger deri, og hvordan man finder det. Det vil jeg gøre med udgangspunkt i en case med tre dele: Tre arkiveringsversioner af omsorgssystemet Care, der bruges i ældre- og handicapplejen hos Aalborg Kommune. Selv med denne afgrænsning er det svært at sætte en ramme for præcis, hvad jeg vil forsøge at nå frem til. Udgangspunktet for projektet er, at ingen endnu skrevet noget om, hvordan arkiveringsversioner fungerer i praksis, og der er derfor tale om en udforskende undersøgelse.

Problemformulering

Med udgangspunkt i tre forskellige arkiveringsversioner af omsorgssystemer Care på Aalborg Stadsarkiv vil jeg undersøge, hvilken indflydelse den danske bevaringsstrategi har på de digitale arkivaliers tilgængelighed.

Mere specifikt vil jeg undersøge, hvad der ligger i dem, og hvordan finder man det. For at kunne besvare spørgsmålet vil jeg kigge på arkiveringsversionernes struktur, indhold, dokumentation og søgemuligheder.

Hvad mener jeg med struktur, indhold, dokumentation og tilgængelighed:

- Med struktur mener jeg tabelstruktur (tabeller og relationer). Dvs. hvordan indholdet er fordelt ud på tabeller, og hvordan er de relaterede til hinanden. Det har betydning for bl.a. at genfinde og samle sager om personer, hvis data kan ligge i mange forskellige tabeller.
- Med indhold mener jeg de data, der ligger i tabellerne. I praksis vil det sige rækker og kolonner, samt tilhørende dokumenter. Det er de informationer en arkivbruger egentlig er interesseret i at finde.
- Med dokumentation mener jeg kontekstdokumentationen. Det er et krav fra Rigsarkivet, at det skal med, og forskellige typer dokumentation er afgørende for om arkiveringsversionen kan bruges i fremtiden uden både meget højt specialiserede historiske IT-færdigheder og dybdegående historisk viden om praksis og arbejdsgange i den offentlige institution, som et system blev brugt af.
- Med tilgængelighed mener jeg først og fremmest teknisk tilgængelighed. Dvs. om arkiveringsversionerne er lavet og præsenteret, på en måde, der gør en bruger i stand til at finde, hvad de søger. Tilgængelighed har også en administrativ del, der har at gøre med, hvordan der gives adgang. Herunder hører både regler om personfølsomme oplysninger og hvor mange led data skal igennem fra arkiveringsversionen og ud til den endelige bruger.

Rammevilkår og begrænsninger

De arkiveringsversioner, jeg har arbejdet med, er kommunale. Arkiveringsversionerne kommer fra Aalborg Kommune og ligger derfor hos Aalborg Stadsarkivs samling hos KOMDA. KOMDA arkiverer digitalt materiale for 11 forskellige kommuner,¹ og jeg går derfor ud fra at kommunal arkivering fra disse ligner hinanden. I dette speciale kan jeg ikke forholde mig til, hvordan Rigsarkivet eller NEA håndterer kommunale afleveringer, og jeg forholder mig heller ikke til, hvordan statslige arkiveringsversioner håndteres.

¹ KOMDAs medlemsoversigt: <https://www.aalborgstadsarkiv.dk/Komda.asp?Komda=KomdaDeltagere>

Da det er Rigsarkivet, der sætter de overordnede tekniske rammer for arkivering, må man dog gå ud fra, at de ligner hinanden. Mine valg af arkiveringsversioner vil blive uddybet nærmere i afsnittet "Udvælgelse af arkiveringsversioner." De forskellige organisationer har dog hver deres egen måde at tilgængeliggøre og præsentere arkiveringsversioner på, og jeg har kun haft adgang til KOMDAs. Derfor kan jeg kun forholde mig til, hvordan de har valgt at gøre det.

Ifølge KOMDA var der ingen nedskrevne procedurer for test og kvalitetssikring af de tidlige arkiveringsversioner. Deres gældende procedure betragtes som en forretningshemmelighed, og jeg kan derfor ikke henvise til den. Jeg kan derfor kun i mindre grad forholde mig til, hvordan arkiveringsversionerne er blevet testet og kvalitetssikret. Alle arkiveringsversioner er dog som minimum blevet testet med Rigsarkivets testprogram.

Dansk digital arkiveringsstrategi

Den overordnede arkiveringsstrategi i Danmark er bestemt af Rigsarkivet og opdateres løbende. I dette afsnit tager jeg udgangspunkt i den gældende strategi for digital bevaring.² Det er også Rigsarkivet, der har fastsat de tekniske krav til arkiveringsversioner, som også kommunerne følger. Dem vil jeg komme nærmere ind på i afsnittet "Bekendtgørelse om arkiveringsversioner".

Dansk digital arkiveringsstrategi er baseret på migrering. Dvs. at filer og data, når de arkiveres, først konverteres til forudbestemte bevaringsformater og derefter løbende migreres til nyere formater efterhånden, som de gamle udgår. På denne måde undgår man en dag at stå med digitale arkivalier, som hverken menneske eller computer rent teknisk kan læse (læs: åbne). Teknologi udvikles løbende, og det har derfor været forskelligt, hvor ofte man har vurderet det nødvendigt at migrere. Tidligere har man sagt hvert femte år, men i den gældende strategi lader man det stå som "når det vurderes, at data er i fare for teknologisk forældelse, eller hvis Rigsarkivet beslutter at anvende andre standarder i sin bevaringsmetodik"³ uden faste intervaller. I praksis har der kun været migreret én gang i de sidste ca. 20 år - i hvert fald hos KOMDA - og det

² Rigsarkivet har valgt at gemme de tidligere så godt væk, at de ikke er til at finde igen. Internet Archive bevarer uheldigvis ikke linkede PDF-filer, og der ser heller ikke ud til at have været selvstændige strategier for det digitale.

³ Rigsarkivet, "Strategi for digital bevaring 2025," s. 6.

var i 2014. Denne migrering var i kategorien "Rigsarkivet beslutter at anvende andre standarder." Det vender jeg tilbage til i specialets analysedel.

Det er ikke alt i Rigsarkivets strategi for digital bevaring, der har relevans for kommunearkiverne. En del af strategien har at gøre med Rigsarkivets egne interne projekter, som for eksempel investering i et nyt system til modtagelse og bevaring. Omvendt kommer planerne om udstedelse af nye bekendtgørelser til at påvirke alle offentlige arkiver i Danmark.⁴ Det mest interessante i den nyeste strategi er, at man lægger op til at indføre såkaldt parallelarkivering.⁵ Dvs. bevaring af både originale versioner og arkiveringsversioner af data og dokumenter.

Parallelarkivering blev forslået af Asbjørn Skødt i artiklen "That's Us with the Originals",⁶ hvor han diskuterer metoden som en måde at sikre autenticiteten for de digitale arkivalier, selvom de migrerede udgaver ændrer sig over tid. Mens Skødt primært fokuserer på at sikre autenticitet, kommer han også omkring flere andre fordele ved parallelarkivering. Disse fordele inkluderer bl.a. at man kan bevare udgaver af data (og dokumenter), med funktionaliteter, der ellers ville gå tabt for evigt, hvis man kun beholder de konverterede udgaver. Ved at bevare originalerne side om side med arkiveringsversionerne vil man også få mulighed for at skabe nye og forbedrede arkiveringsversioner, hvis en bedre metode bliver udviklet i fremtiden. Man vil også få mulighed for at udnytte potentialet for emulering, når eller hvis der udvikles en god løsning på denne ellers usikre og meget dyre løsning. Ifølge Skødt arbejder man på sagen.⁷ I sin strategi lægger Rigsarkivet udelukkende vægt på at bevare originale funktionaliteter i forskellige filformater, som går tabt ved konvertering i dag.

Hverken Rigsarkivet eller Skødt nævner bevaring af databaser eller systembrugerflader som en del af parallelarkivering, men forholder sig først og fremmest til filtyper. Han nævner således ikke, om man også vil bevare originale databaseformater - f.eks. hvis de ikke er relationelle, som Rigsarkivet ellers kræver. Der er flere konsekvenser ved migrering, som parallelarkivering vil kunne mitigere. F.eks. kan der, som vi skal se senere, laves ændringer i tabeller og relationer (altså datastrukturen) ved migrering, der betyder, at søgning i arkiveringsversionen kan blive ændret. Ved kon-

⁴ Ibid. Begge dele nævnes på s. 4 og uddybes en smule på s. 6.

⁵ Ibid. s. 7. Rigsarkivet kalder det flere "bevaringsniveauer".

⁶ I *Nordic Model of Digital Archiving*.

⁷ Skødt, "That's us with the originals," s. 183-84.

vertering i første omgang kan data, som vi også skal se senere, blive gemt som dokumenter i stedet, hvilket betyder, at de skal håndteres helt anderledes i arkiveringsversionen. Jo flere af disse ændringer, der akkumuleres over tid, og jo mere forskellig arkiveringsversionen bliver fra driftsversionen, jo mindre brugbar risikerer dokumentationen også at blive. I alle tilfælde vil parallelarkivering betyde en øget datamængde og større krav til at kunne bevare flere typer data.

Bevaringsstrategier i Danmark og udlandet

Greg Bak har diskuteret forskellene på de digitale bevaringsstrategier i Danmark og Canada. De er meget forskellige og har ifølge Bak begge fordele og ulemper. Nogle af dem har jeg allerede været inde på. Den danske bevaringsstrategi har rødder tilbage til 1970'erne, og vi har ikke ændret på grundprincipperne siden dengang. Der er som beskrevet ovenfor, at man konverterer hele systemers indhold til et systemafhængigt bevaringsformat og bevarer det sammen med relevant dokumentation, men uden systemernes brugerflade eller funktioner.⁸ Strategien er understøttet af Danmarks stærke tradition for god journaliseringspraksis.⁹ Bak argumenterer også for, at danske arkiver (mest Rigsarkivet) har haft en stor økonomisk fordel ved at lægge ansvaret for konvertering over på myndighederne frem for hos arkivet. Det kan ifølge Bak dog også have den ulempe, at arkivet mister kontrollen med såkaldt "technical appraisal", der beskrives som en blanding af at sikre digitale arkivaliers integritet, og en vurdering af, om bevaringsformatet er tilstrækkeligt til at sikre arkivalieres indhold.¹⁰ I praksis holder denne argumentation kun delvist. De kommuner, der deltager i KOMDA, har uddelegeret begge dele til KOMDAs ansatte, som tester indkomne arkiveringsversioner for netop integritet. Om bevaringsformatet er tilstrækkeligt er slet ikke op til myndighederne, men derimod til Rigsarkivet. Der hvor denne kritik for alvor bliver interessant er, at myndighederne har uddelegeret den praktiske del af konverteringen til bevaringsformat til private leverandører (af IT-systemerne). Med andre ord er langt de fleste kommunale arkivalier på et kritisk tidspunkt i arkiveringsprocessen helt udenfor for både myndighedernes og arkivernes kontrol. Det er der endnu ingen, der ser ud til at have et problem med.

⁸ Bak, "Continuities and innovations," s. 99 i *Nordic Model of Digital Archiving*.

⁹ Ibid. s. 104.

¹⁰ Ibid. s. 111.

Bak gør også opmærksom på, at tilgængeliggørelse (især for forskere) i flere årtier slet ikke var en prioritet i dansk digital arkivering. Det er modsat i Canada, hvor det var et krav helt fra starten.¹¹ For de statslige arkivalier har det, ifølge Rigsarkivet, ændret sig en del,¹² men hos kommunearkiverne er forskningsadgang til digitale arkivalier endnu ikke blevet en prioritet. Prioriteringen hos kommunearkiverne og hos KOMDA specifikt har Rasmus Grunwald Falk og Simon Stidsing Hansen skrevet om. Deres gennem af udviklingen indenfor digitalarkivering afspejler den overordnede danske bevaringsstrategi. Her var det den fysiske og tekniske bevaring, der var i fokus til at starte med. Brug af arkiveringsversionerne var faktisk i første omgang tænkt som forskning, men det var forskning for historikere engang i fremtiden.¹³ Over tid er fokus skiftet til at være arkivets nuværende brugere: Borgere og myndigheder. Mere specifikt er arkivskaberne blevet mere involverede i kvalitetssikringen og i udformningen af søgninger. Man har derudover udviklet søgninger på tværs af arkiveringsversioner, således at man hurtigt kan finde en borgers oplysninger fra mange systemer på én gang.¹⁴

Lovgivning

I det følgende vil jeg gennemgå den lovgivning, der har betydning for specialets emne. Lovgivningen sætter rammerne for både myndigheder, offentlige arkiver og leverandører af IT-systemer til det offentlige.

Arkivloven

De overordnede rammer for dansk offentlig arkivering sættes af Arkivloven. Den giver arkiverne mandat til at modtage og opbevare offentlige data og dokumenter. Arkivloven sætter også meget overordnede rammer for, hvad der skal arkiveres. Stads- og kommunearkiverne har hjemmel i arkivlovens § 7, der giver kommuner og regioner mulighed for at oprette egne arkiver. Det samme har KOMDA, der arkiverer for flere kommuner. Hvad der overordnet skal bevares hos dem er det samme som det, der bevares fra statslige myndigheder hos Rigsarkivet og nævnes i § 7 stk 1: "[A]rkivalier,

¹¹ Bak, "Continuities and innovations," s. 108.

¹² I Rigsarkivets årsrapport for 2024 (s. 10) står der f.eks., at man arbejder på en platform, hvor man kan tilgængeliggøre digitalt skabte arkivalier. Man er også begyndt at forberede sig på at kunne bruge dem til at udvikle en dansk sprogmodel (AI).

¹³ Falk og Hansen, "Digitale arkivalier: Fra sorte bokse til big data," s. 77 i *Dokumentation i en digital tid: ESDH, arkiver og god forvaltningsskik*.

¹⁴ Ibid. s. 77-81.

der har historisk værdi eller tjener til dokumentation af forhold af væsentlig administrativ eller retlig betydning for borgere og myndigheder.”

For stads- og kommunearkiverne gælder desuden § 21: “Kommunale og regionale arkivalier, der er omfattet af lov om behandling af personoplysninger, og som skal bevares, skal afleveres til et offentligt arkiv.”¹⁵ Denne paragraf betyder, at kommunen ikke må aflevere bevaringsværdige data med personoplysninger til private arkiver. Det samme er gældende for statslige myndigheder, men de kan ikke aflevere til kommunearkiverne (se § 12). Omvendt kan kommunerne godt aflevere til Rigsarkivet. Det er ifølge Arkivloven frivilligt for kommuner og regioner at oprette sit eget arkiv. Da dette speciale har at gøre med arkiveringsversioner hos Aalborg Stadsarkiv via KOMDA kan jeg dog ikke forholde mig til kommunale arkivalier, der afleveres til Rigsarkivet.

Da de fleste kommunale arkivalier hos Aalborg Stadsarkiv er under 20 år gamle er de stadig omfattet af den almindeligt gældende tilgængelighedsfrist (§ 22). Ydermere indeholder de arkiveringsversioner, jeg bruger som cases i dette speciale, personfølsomme oplysninger og er derfor omfattet af § 23, der fastsætter tilgængelighedsfristen på 75 år. Når det kommer til tilgængeliggørelse af materialet, er det derfor kapitel 7, “Adgang til ikke umiddelbart tilgængelige arkivalier”, der bliver afgørende. Her er det for de kommunale arkiveringsversioners vedkommende særligt § 32 og 34, der gør sig gældende. De fastsætter hhv. at det er kommunalbestyrelsen, der skal give adgang, og at også Datatilsynet skal give tilladelse, hvis der er tale om personfølsomme oplysninger eller hvis “oplysningerne stammer fra et edb-register, der har været ført for den offentlige forvaltning.”¹⁶

Bevaring og kassation af digitalt skabte arkivalier i kommunerne

Hvad der skal gemmes specifikt hos kommunerne uddybes nærmere i ca. 16 forskellige bekendtgørelser, der er eller har været gældende mellem 2003 og nu. Rigsarkivet skelner overordnet mellem dem, der var gældende før og efter 2007, eller mere specifikt kommunalreformen i 2007. Da mit udvalg af arkiveringsversioner er fra begge

¹⁵ Arkivloven §21

¹⁶ Arkivloven § 32 og 34. I praksis er denne opgave blevet uddelegeret til stadsarkivaren.

sider af 2007, vil jeg se på relevante regler fra begge perioder. Det er ikke alle 16 bekendtgørelser, der er eller var gældende for Aalborg Kommune. Før 2007 var bevarings- og kassationsbestemmelserne delt op i amtskommuner og primærkommuner samt specialregler før København, Frederiksberg og Bornholm.

For Aalborg Kommune var det før 2007 "Bekendtgørelse om bevaring og kassation af arkivalier i primærkommunerne,"¹⁷ der var gældende. Her nævnes IT-systemer til ældre- og handicappleje i bilag 1: "[Bevares skal] En arkiveringsversion af data fra systemer vedrørende hjemmehjælp, handicapomsorg, ældreinstitutioner m.v., der indeholder identifikationsoplysninger, visitationsoplysninger, tildelt hjælp og normeringsoplysninger for den enkelte institution."¹⁸

Generelt omtales it-systemer som regel kun som bevaringsværdige i bekendtgørelserne, hvis de indeholder personoplysninger. Efter 2007 blev bekendtgørelserne om bevaring og kassation af digitalt skabte arkivalier i kommunerne lagt sammen til én, der er gældende for alle kommuner. Den gældende bekendtgørelse er nr. 183 af 26/01/2018. I dens første opdatering (BEK nr. 1469 af 28/09/2020) gøres det klart, at bevaring og kassation hos kommunerne overordnet tager udgangspunkt i kommunens journalisering via KL-emnesystematikken. Dvs. at man fra Rigsarkivets side vurderer om hvert enkelt KLE-nummer er bevaringsværdigt eller ej. Bekendtgørelsen er derfor løbende blevet opdateret efterhånden som KLE har fået tilføjet nye saggrupper.¹⁹ I denne bekendtgørelse fastslås det også, at kommunale digitale arkivalier skal afleveres i form af arkiveringsversioner på lige fod med de statslige arkivalier.²⁰

Når man bevarer hele systemers indhold, bevarer man også eventuelle ekstra data, der måtte ligge der, som man i papirtiden ikke ville have anset for bevaringsværdige. I de enkelte systemer kasseres både tabeller og kolonner, og man ville teknisk godt kunne kassere ud fra om indhold er bevaringsværdigt på traditionel vis, men man gør det ikke. I praksis kasseres som regel kun tabeller og kolonner, der indeholder rent teknisk systeminformation, som ikke giver mening i en arkiveringsversion, f.eks. passwords. Årsagerne til, at det gøres således, er ikke skrevet ned, men har formentlig at gøre med, at det tager meget lang tid at gennemgå et helt system for at vurdere,

¹⁷ BEK nr 1000 af 28/09/2004.

¹⁸ BEK nr 1000 af 28/09/2004, bilag 1.

¹⁹ Man tilføjede et bilag til "Bekendtgørelse om bevaring og kassation af digitalt skabte data og dokumenter fra kommunerne" med hele emnelisten med bevarings- og kassationsvurderinger. Flere af de efterfølgende ændringer er af samme karakter for at inkorporere nye emner i KLE.

²⁰ BEK nr 183 af 26/01/2018 1.2.

hvilke data, der er bevaringsværdigt. Derudover kan det have ganske store konsekvenser for både mening med data og søgbarhed, hvis man begynder at slette ting i en relationel database, hvor samtlige tabeller hænger sammen med andre tabeller. Denne vurdering bakkes op af praktiske erfaringer hos Aalborg Stadsarkiv, der brugte samme rationale ved testafleveringen af Care tilbage i 2002-2003.²¹ Hvad der anses som bevaringsværdigt har altså ikke ændret sig ret meget. Den store forskel ligger i, at man ikke længere kasserer ret meget af det som *ikke* er bevaringsværdigt. Det kan blive en fordel for fremtidens historikere, der får en omvendt oplevelse af, hvad nutidens historikere bøvler med (at man har kasseret for meget i det 20. århundrede). Det tilføjer dog også en del rod, som kan gøre det svært at finde det, man leder efter.

Bekendtgørelse om arkiveringsversioner

For digital arkivering er det næstvigtigste sæt regler efter arkivloven bekendtgørelsen om arkiveringsversioner. Der har været tre udgaver af bekendtgørelsen, hvor den første er fra 2004. Den blev opdateret med en ny i 2010, og den gældende er fra 2020. Da dette speciales undersøgelse har at gøre med sammenligning af arkiveringsversioner over tid, vil jeg gennemgå hovedtræk og udvalgte detaljer i alle tre. De udvalgte detaljer vil specifikt være kravene til format og standarder for data, dokumenter og dokumentation. Bekendtgørelserne har også krav til lyd, video og geodata. Jeg vil undlade at gennemgå detaljer, der har at gøre med forskningsdata og geodata, da mine valgte arkiveringsversioner ikke indeholder nogen af delene. Aalborg Stadsarkiv ligger slet ikke inde med digital forskningsdata, mens geodata snarere findes i byggesagssystemer eller ESDH-systemer.

Bekendtgørelsernes overordnede ordlyd har ændret sig meget lidt siden 2004. De fastslår alle tre, at:

- Data fra bevaringsværdige offentlige IT-systemer skal bevares i form af arkiveringsversioner
- Arkiveringsversionerne skal følge anvisningerne i bekendtgørelsens bilag
- Arkiveringsversionerne skal afleveres til og godkendes af det modtagende arkiv (og hvis det er fra staten, skal det være Rigsarkivet)

²¹ Aalborg Stadsarkiv, Kortlægning, aflevering og genrejsning af IT-arkivalier,” s. 8.

- Myndighederne ikke må slette data før en arkiveringsversion er blevet godkendt

Forskellene i denne indledende del af bekendtgørelserne ligger hovedsageligt i, at sproget er ændret en smule for at følge med tiden. F.eks. er man gået fra specifikt at tale om “data fra elektroniske arkivsystemer, herunder databaser eller registre, elektroniske journalsystemer, elektroniske sags- og dokumenthåndteringssystemer (ESDH-systemer) samt elektroniske dokumenthåndteringssystemer (EDH-systemer)”²² til “data fra it-systemer, lyd og video”²³ til i den seneste det meget brede “digitale data”²⁴, når der står, hvilke arkivalier bekendtgørelserne er gældende for. Det vidner om, at det digitale er gået fra at være noget specifikt og specialiseret i den offentlige forvaltning, til at være noget mere generelt og almindeligt, der ikke kun findes i IT-systemer. Som det kan tolkes ud af bekendtgørelsens efterfølgende afsnit (produktion og aflevering) og bilag samt Aalborg Stadsarkivs digitale samling, er IT-systemer dog stadig den mest almindelige oprindelse for digitale arkivalier.

Udover arkiveringsversioner af IT-systemer, kan der være arkiveringsversioner af offentlige data, der har ligget udenfor et IT-system (enten gamle data eller fra personer i forvaltningen, der ikke har fulgt journaliseringsreglerne), eller forskningsdata fra f.eks. offentlige universiteter. Sidst men ikke mindst kan der være tale om private arkivalier såsom foreningsarkiver. Når man taler om arkiveringsversioner vil det sige, at data og dokumenter er blevet konverteret fra originalformater til arkiveringsformater. Disse formater er beskrevet i bekendtgørelsernes bilag.

De største forskelle på kravene til arkiveringsversioner findes i bekendtgørelsernes bilag, hvor man finder alle de tekniske detaljer. Det er ikke alle tekniske detaljer, der er relevante for denne undersøgelse. Derfor vil jeg fokusere på de vigtigste, og i særlig grad på, hvad man har ændret på, da det har relevans for migrering.

Bekendtgørelse nr. 342 af 11/03/2004

Første bekendtgørelse etablerede, at arkiveringsversioner af IT-systemer så vidt muligt skal have en relationel databasestruktur. Tabeller skal have hver sin fil, og på dette tidspunkt skulle det være ARK-filer.²⁵ Tabellernes metadatafiler skulle til gengæld være XML.²⁶ I denne bekendtgørelse var der ingen specifikke regler for navngivning

²² BEK nr. 342 af 11/03/2004, §2

²³ BEK nr. 1007 af 20/08/2010, §2

²⁴ BEK nr. 128 af 12/02/2020, §2

²⁵ BEK nr. 342 af 11/03/2004 s. 20.

²⁶ BEK nr. 342 af 11/03/2004 s. 10.

af tabeller. Det kan myndigheden selv bestemme,²⁷ men i praksis er det formentlig leverandøren, der gør det.

Dokumenter skal ifølge denne bekendtgørelse bevares som TIFF-filer og, hvis myndigheden ligger inde med en computerlæsbar tekst-udgave, som en tekstfil. Lyd-filer bevares som MP3 og video som MPEG-2.²⁸ Alle dokumenter skal desuden have en "entydig identifikation", og at denne identifikation skal være et nummer på 8 cifre.²⁹ På dette tidspunkt var geodata endnu ikke kommet med.

I denne bekendtgørelse forstås metadata specifikt som de XML-filer, der beskriver arkiveringsversionens tabeller.³⁰

Kontekstdokumentation kaldes i denne bekendtgørelse "generelle informationer" eller GENINFO. Sammenlignet med de nyere bekendtgørelser, er kravene til dokumentation meget overordnede og uspecifikke. Det er et krav, at arkiveringsversioner "skal indeholde dokumenter med information om det elektroniske arkivsystems administrative funktion og systemets struktur og funktionalitet."³¹ Dokumenterne skal være enten TIFF eller tekst, og det er op til det modtagende arkiv, hvilke der skal med.³²

Denne bekendtgørelse omtaler også såkaldte hjælpetabeller. De har filtypen TAB. I hjælpetabellerne står metadata om arkiveringsversionen, arkivskaber, generelle informationer (kontekstdokumentationen), dokumenter og filer (inkl. tabeller).

Bekendtgørelse nr. 1007 af 20/08/2010

Der kom relativt omfattende ændringer i denne bekendtgørelse om arkiveringsversioner. Stort set alle ændringer er stadig gældende. Man gik væk fra TAB-filerne (hjelpe-tabellerne) og skiftede dem ud med indeksfiler i XML-format. Det samme er tilfældet med tabeller, der nu skulle være XML i stedet for ARK. De kommer også til at hedde noget andet. Hjælpetabellerne ARKVER.TAB og SKABER.TAB blev lagt sammen til én XML-fil, der kaldes Arkivbeskrivelsesfilen. Arkivbeskrivelsesfilen indeholder metadata for hele arkiveringsversionen som en enhed og er obligatorisk. De nye navne

²⁷ Ibid. s. 20.

²⁸ Ibid.

²⁹ Ibid. s. 8.

³⁰ Ibid. s. 21

³¹ Ibid. s. 10.

³² Ibid.

blev ikke ændret med den nyeste bekendtgørelse, så når jeg senere i analysen henviser til denne type metadata, er det disse navne, jeg bruger, selvom arkiveringsversion er fra før 2010.

Kontekstdokumentationen blev også præciseret en hel del. Det er her begrebet kontekstdokumentation bliver introduceret, og her man introducerer kategorier til det. Også kontekstdokumentationen får sin egen indeksfil, *contextDocumentationIndex*. Det er dog stadig op til det modtagende arkiv at vurdere, hvilken dokumentation, der skal med så længe den overholder kravet om at "dokumentere it-systemets administrative funktion samt struktur og funktionalitet."³³

For tekstdokumenters vedkommende skulle man nu ikke længere gemme en computerlæsbar tekstudgave sammen med billedfilen. Som arkiveringsversionerne er sat op lige nu, har det ikke nogen praktisk betydning for søgning- Det vender jeg tilbage til i afsnittet "søgning på dokumenter". Det er også i denne bekendtgørelse, at geodata kommer til. Det bevares som GML-filer efter en dansk udgave af ISO 19136.

Bekendtgørelse nr. 128 af 12/02/2020

Kravene i den gældende bekendtgørelse er stort set identiske med kravene i den forrige. De få tekniske ændringer, som findes, er ikke synlige, når man kigger i arkiveringsversionerne som bruger. En af ændringer er f.eks., at det ikke længere er obligatorisk at have et XML-skema til hver enkelt tabel. Antallet af datatyper, der må bruges til decimaltal, er også blevet udvidet en smule, hvilket en bruger formentlig heller ikke vil bemærke. Den største ændring er, at forskningsdata er blevet tilføjet som en kategori af indhold, en arkiveringsversion kan have.

Bekendtgørelserne har den overordnede konsekvens at alle danske arkiveringsversioner, og i praksis alle danske digitale arkivalier fra det offentlige, skal leve op til de samme tekniske krav og have de samme formater uanset, hvilket arkiv, der håndterer dem.

³³ BEK nr. 1007 af 20/08/2010, bilag 6.B.

Data og dokumenter

I danske arkiveringsversioner er der overordnet to slags indhold: Data og dokumenter. Det er nødvendigt at forstå forskellen af mindst to årsager: For det første håndteres de forskelligt i arkiveringsprocessen. For det andet er det lettere at forstå undersøgelsens sammenhæng med eksisterende forskning, hvis man forstår, at de håndteres anderledes og hvordan. Næsten alt forskning om digital arkivering tager udgangspunkt i dokumenter, som jeg har defineret dem nedenfor.

Data

Data kan forstås *meget* bredt afhængig af sammenhæng. Her forstås det som tabelindhold. Langt de fleste arkiveringsversioner indeholder data i form af tabelindhold, mens færre har dokumenter (se nedenfor). Kun undtagelsesvis findes der arkiveringsversioner, der kun består af dokumenter.³⁴

Tabelindhold lagres i XML-format. Til XML-tabellerne hører XML-skemaer, der bruges til validering af tabelindholdet. I tabellerne er følgende datatyper tilladte ifølge den gældende bekendtgørelse: String, integer (hele tal), decimal, float, double, boolean, hexBinary, date, time, datetime og duration.³⁵ De ovenstående skal derudover følge standarden SQL:1999, som på visse punkter er lidt mere begrænset. Hvis data i originalformat afviger herfra, skal de derfor laves om, f.eks. ved at afkorte meget lange decimaltal.³⁶ I princippet kan der ved denne konvertering derfor gå indhold tabt.

Hvis data slet ikke kan leve op til kravene om datatyper, skal det konverteres til et dokument i stedet. I arkiveringsversioner, der er arkiverede efter den første bekendtgørelse konverteres der, hvis et felt indeholder mere end 2048 tegn.³⁷ I de to senere bekendtgørelser er der ikke en maksimal feltlængde, og her handler det i stedet om at tage højde for systemer, hvor "dokumenter, lyd, video og/eller geodata m.v. [lagres] i binært format f.eks. som blobfelter i datatabellerne," eller systemer, hvor der anvendes opmærkningssprog i tekstfelter (f.eks. html), og hvor det ikke er muligt at konvertere til string.³⁸

³⁴ Arkiveringsversioner, der kun består af dokumenter, har dog stadig XML-baseret metadata.

³⁵ Rigsarkivet, Rigsarkivet, Vejledning til bekendtgørelse 128, figur 5.1, s. 45.

³⁶ Ibid. s. 46.

³⁷ BEK nr. 342 af 11/03/2004, s. 8 (bilag 2.E).

³⁸ Rigsarkivet, Vejledning til bekendtgørelse 128, bilag 5.E, s. 49.

Dokumenter

Dokumenter skal forstås meget bredt i denne sammenhæng. I bekendtgørelserne omtales “elektroniske dokumenter” eller “digitale dokumenter”, og her er der tale om tekstdokumenter. I praksis er det også oftest tekstdokumenter, man tænker på, når man taler om dokumenter. Lyd, billeder, video og geodata håndteres dog på samme måde i en arkiveringsversion. Altså forskelligt fra tabelindhold. Med andre ord er dokumenter alt andet end tabelindhold (data). Som nævnt ovenfor kan visse typer tabelindhold dog også arkiveres som dokumenter, hvis det ikke kan overholde kravene til datatyper, som de er opstillet i bekendtgørelsen.

Alle tekstdokumenter og billeder skal arkiveres som enten TIFF eller JPEG-2000. I KOMDAs regi er TIFF det mest almindelige. Begge dele er billedformater. Der er en række forskellige muligheder for, hvilke versioner af TIFF eller JPEG-2000, man kan bruge, deres bitdybde (antal farver) og hvordan de kan komprimeres alt efter om der er tale om relativt simple sort/hvid tekstdokumenter, fotos, kort eller tegninger samt, hvor store de originale versioner er, og hvor stor betydning detaljegraden har. Generelt sigter man efter at bevare dokumenter i samme kvalitet som originalformaterne mht. til opløsning og farver. Samtidig ser Rigsarkivet også, at man ikke vælger et bevaringsformat, der er *bedre* end originalformaterne, da de for det første fylder unødigt, og for det andet kan være besværlige at åbne.³⁹

Offentlige IT-systemer kan være systemer med eller uden dokumenter, som bruges til sagsbehandling, eller det kan være et register. De vigtigste registre er f.eks. CPR-registret, der indeholder stamoplysninger om alle borgere i Danmark. Andre systemer kan have en snitflade⁴⁰ til CPR, så de kan bruge disse informationer uden at skulle indtaste dem manuelt. På denne måde kan ændringer i f.eks. civilstand eller navn i princippet opdateres automatisk, og man nedbringer mulighederne for menneskelige fejl ved indtastning. Aalborg Kommune har desuden et decentralt register over borgere tilknyttet Aalborg Kommune, der indeholder yderligere informationer end hvad, man kan få fra CPR.

³⁹ Rigsarkivet, Vejledning til bekendtgørelse 128, bilag 5.E, s. 50-53.

⁴⁰ En forbindelse til et andet system, som der kan udveksles data med.

Metadata i arkiveringsversioner

I litteraturen om digital arkivering omtales metadata ofte som det vigtigste element i digital arkivering udover selve arkivalierne. Derfor vil jeg her forklare, hvordan metadata fungerer i arkiveringsversioner. André Pacheco, et al. har gennemgået en del af den relevante litteratur på området med det formål at kortlægge, hvilke krav 10 forskellige metadatastandards, der er i brug rundt omkring, har, og ud fra dem komme med et forslag til en generel model, der samler de vigtigste elementer.⁴¹ Deres artikel forsøger at inkludere både data (som relationelle databaser), dokumenter og hele samlinger, men det lykkes dem ikke helt at gøre det klart, hvordan deres model kan bruges på data i en relationel database. For dokumenter og hele samlinger er der nogle elementer, der overlapper med dem, man finder i danske arkiveringsversioners metadatafiler. Stort set alle elementer (dokumenter, tabeller, kontekstdokumenter, hele arkiveringsversionen) i en arkiveringsversion har f.eks. et nummer (*identifier* i Pacheco, et al.s model). Men det er f.eks. kun arkivbeskrivelsesfilen og *contextDocumentationIndex*, der har datoer og skaber/forfatter.

Metadata for samtlige dokumenter i en arkiveringsversion registreres først og fremmest i XML-filen *docIndex*. Metadata her består af dokumentets ID, eventuelt ID for et såkaldt parent-dokument (et overordnet relateret dokument), ID på det medie, dokumentet ligger på, navn på den mappe, dokumentet ligger i, dokumentets navn og oprindelige filformat (som extension, dvs. som f.eks. .docx eller .pdf), dokumentets bevaringsformat og eventuelt det relevante skema, hvis der er tale om GML-filer (geodata).⁴² Metadata som dette registreres udenfor selve dokumentet, fordi metadata, der normalt ville være *embedded* i filen, går tabt ved konvertering. Alle øvrige metadata som f.eks. skaber, datering, administrativ kontekst og historik eller tilhørsforhold til en sag eller person skal i stedet findes i arkiveringsversionens tabelindhold.

For tabellerne består metadata af et *tableIndex*, hvori databasens struktur angives.⁴³ Her står oplysninger om hver enkelt tabel og hver enkelt tabels kolonner. Oplysningerne inkluderer tabelnavn, kolonnenavn, beskrivelser af samme, og original datatype, som de elementer, der går igen fra metadatamodeller lavet til dokumenter. Derudover er der oplysninger om bl.a. antal rækker, NULL-værdier, primærnøgler,

⁴¹ Pacheco et al. "A metadata model for authenticity in digital archival descriptions," s. 643-651.

⁴² Rigsarkivet, Vejledning til bekendtgørelse 128, s. 37

⁴³ BEK nr. 128 af 12/02/2020, bilag 4.C.

fremmednøgler,⁴⁴ som ikke har noget til fælles med dokumenternes metadata. Det er i stedet metadata, der eksisterer for, at tabellerne rent teknisk kan fungere som en relationel database.

Dokumenters og tabellers metadata kobles sammen via *functionalDescriptions* i *tableIndex*. Her identificeres de kolonner, hvor dokumenters metadata findes. Her er man interesseret i sagsidentifikation og titel, dokumenttitel og dato, afsender og/eller modtager, digitale signaturer, FORM-reference og oplysninger om bevaring og kassation. Hvis disse informationer ikke findes i systemet, er det ikke et krav, at de skabes, men blot at man kan sikre "identifikation af sagligt sammenhørende dokumenter."⁴⁵

Grundlæggende metadata om arkiveringsversionen i det hele taget, inklusive docIndex, skal ligge i en samlet mappe (*Indices*), hvor der udover indeks for dokumenter også skal ligge indeksfiler for alle filer, tabeller og kontekstdokumenter. Her findes også arkivbeskrivelsesfilen, som indeholder basisinformation om hele arkiveringsversionen.⁴⁶

Alle indeksfilerne er XML-filer, og det er generelt ikke noget, som en arkivbruger kommer i direkte berøring med. I praksis er det først og fremmest tabelnavne, kolonnenavne og beskrivelser af samme, der ser ud til at have relevans for fremsøgning via SQL. Har man først fundet de data eller dokumenter, man søgte, kan de øvrige metadata blive brugbare. Den danske model for arkiveringsversioner ser ikke ud til at følge nogen af de internationale standarder for, hvilke elementer metadata skal indeholde. Man går snarere ud fra, at det meste relevante metadata er blevet tilføjet af myndighederne, mens et system var i drift. Påførsel af metadata på har i en dansk sammenhæng mest at gøre med journalisering, mens påførsel af metadata på tabeller er mere arkivets (og leverandørens) ansvar end myndighedens.

Forskningsdiskussion: Det tekniske

En hel del arkivteori har at gøre med *arrangement*, eller hvordan arkivalierne er fysisk organiseret. Her er tale om det grundlæggende arkivprincip *respect des fonds*⁴⁷ og hvor meget, man tør gå på kompromis med det for at gøre arkiverne mere brugbare.

⁴⁴ BEK nr. 128 af 12/02/2020, bilag 6.C.

⁴⁵ BEK nr. 128 af 12/02/2020, bilag 6.C og figur 6.6.

⁴⁶ Rigsarkivet, Vejledning til bekendtgørelse 128 s. 31-32.

⁴⁷ MacNeil, "Archival Theory and Practice: Between Two Paradigms," 8.

Blandt andre Heather MacNeil har argumenteret for, at dette ikke længere er gældende for digitalt skabte arkivalier, da de er organiserede via metadata fremfor en fysisk struktur.⁴⁸ Data og digitale filer adskiller sig på flere måder væsentligt fra papir - inklusiv i deres struktur og måder at være organiseret på. De kan for eksempel have former, der er fysisk umulige på papir (eller andre analoge medier), de kan være interaktive eller forbundne på måder, der gør dem vanskelige at arkivere på en måde, der ligner papir. Samtidig kan data og filer omorganiseres og sammensættes på uendelig mange måder afhængig af, hvad man vil bruge dem til. Relationelle databaser er helt fra bunden sat op til at kunne bruges på denne måde, da søgning i dem foregår med SQL. SQL-søgning gør det f.eks. muligt at hente data fra de tabeller man ønsker og sætte dem sammen i en virtuel tabel.

Et andet stort emne indenfor digitalarkivering er metadata. Herunder hører også arkivbeskrivelse. Devon Mordell diskuterer i "Critical Questions for Archives as (Big) Data" et potentielt nyt arkivparadigme, hvor digitale metoder, herunder machine learning, kan bruges til arkivbeskrivelser og *arrangement*.⁴⁹ Mordell diskuterer også en definition af data, hvor data kan være alt digitalt materiale, som af nogle betragtes som et neutralt råmateriale til brug i kvantitative studier.⁵⁰ Det forholder hun sig, med rette, meget kritisk til. Det adskiller sig fra den definition af data, som jeg har beskrevet i afsnittet om data og dokumenter ovenfor, hvor data specifikt betyder tabelindhold i databaser i danske arkiveringsversioner. Uanset hvilken definition af data, man bruger, er det dog aldrig neutralt. En af Mordells pointer er, at vi ikke skal glemme de menneskelige dele af arkiverne og arkivfaget, som vi er nået så langt med i Cooks forrige fire paradigmer.⁵¹ Som vi senere skal se, er de menneskelige aspekter stadig til stede i danske arkiveringsversioner. Mordell diskuterer også brugen af automatiske påførte arkivbeskrivelser og særligt brug af natural language processing til at gøre det. Hun gør bl.a. opmærksom på, at denne proces på ingen måde er neutral og ofte er meget uigennemsigtig. Tværtimod risikerer man at få et resultat, der afspejler sprogmodellens træningsdata, og man kan ende med at fastholde de udfordringer med repræsentation af kvinder og minoriteter i arkiverne, som vi ellers var begyndt at arbejde på.⁵² Mordell foreslår ikke, at man helt dropper disse nye teknologier, da de kan være

⁴⁸ MacNeil, "Archival Theory and Practice: Between Two Paradigms," 10.

⁴⁹ Mordel, "Critical Questions for Archives as (Big) Data," s. 144.

⁵⁰ Mordel, "Critical Questions for Archives as (Big) Data," s. s. 147-150.

⁵¹ Ibid. s. 146-149.

⁵² Ibid. s. 152-155.

et redskab til at åbne arkiverne for nye brugere.⁵³ Hun argumenterer for, at man kan komme udenom mange af udfordringerne med bedre gennemsigthed i den måde, arkiverne håndterer data på og ved at inkludere dataproveniensen.⁵⁴ Sidstnævnte harmonerer med Baks digitale proveniens, som jeg vil komme ind på senere i dette afsnit.

I en dansk kontekst er vi ikke nået til at lave automatiske beskrivelser. Ganske vist er meget metadata genereret automatisk, men tabel- og feltbeskrivelser, arkivbeskrivelsefilen, *contextDocumentationIndex* og alt kontekstdokumentationen er manuelt arbejde.

Om digitale arkivalier rent faktisk kan bruges som Big Data i praksis er jeg ikke kvalificeret til at udtale mig om. Om det kan bruges kvantitativt i mere traditionel forstand, har at gøre med, hvordan det tilgængeliggøres. Det skal have en computerlæsbar form, og man skal kunne udvælge data på en måde, der giver mening for det formål, man har valgt. Det er ikke nødvendigvis så ligetil. Som beskrevet i de forrige afsnit ligger dokumenter i danske arkiveringsversioner f.eks. som TIFF, der - for nu - ikke er umiddelbart computerlæsbare, fordi de er billeder. Det kan ændre sig i fremtiden, hvis arkiversnes genrejningsprogrammer f.eks. for indbygget Transkribus-lignende modeller til automatisk transskribering. Det kan være et eksempel på et at de værktøjer, Mordell diskuterer. Tabelindhold er til gengæld computerlæsbart og er i princippet skabt således, at det kan bruges kvantitativt relativt nemt. I hvert fald sammenlignet med papir, TIFF-dokumenter og filer i andet end tekstformat. Det kræver dog, at man kan finde rundt i databasernes tabeller og relationer. Der kan også være tabeller, kolonner og indhold, der eksisterer i en form, der skulle understøtte systemfunktioner, men som gør kvantitativ brug af data lidt mere besværlig, end det kunne have været. F.eks. ved at have data spredt ud på *mange* tabeller eller omvendt at have tabeller med to vidt forskellige typer oplysninger. Sidstnævnte vil der komme et eksempel på senere.

I dette speciale er der først og fremmest tale om relationelle databaser, som ikke kan eksistere i papirform og hvor data heller ikke ligger i nogen bestemt rækkefølge og heller ikke har en hierarkisk orden. Som vi senere skal se, betyder det dog ikke, at de ikke har en orden, der skal respekteres for, at de giver mening.

⁵³ Ibid. s. 156.

⁵⁴ Ibid. s. 159.

Jane Zhang har undersøgt original orden i digitale arkivalier. Hun har sammenlignet praksis hos et bibliotek og arkiv (Wellcome Library), der indsamler fra privatpersoner og foreninger, og et digitalarkiv (LAC TDR, det canadiske nationalarkivs digitale afdeling), der indsamler fra det offentlige samt et fælles digitalt arkivprojekt mellem en række amerikanske delstater (PeDALS), der også indsamler fra det offentlige. Førstnævnte får ofte materiale ind, der kan være så kaotisk, at arkivaren er nødt til at skabe eller genskabe en orden efter aflevering til arkivet. Her er der som regel tale om materiale, der har ligget i det filsystem, der har været en del af en computers styresystem⁵⁵ - det har med andre ord været opbevaret på samme måde som dine personlige dokumenter.

På det canadiske nationalarkiv foregår arkiveringen på en måde, der minder en smule mere om Danmark. Nærmere bestemt bruger den canadiske stat også elektroniske dokumenthåndteringssystemer, som der arkiveres fra, og de organiseres ved hjælp af klassifikationskoder på hvert enkelt dokument.⁵⁶ Til sammenligning bruger vi i Danmark sagsnumre fremfor numre på enkelte dokumenter⁵⁷ samt KLE eller FORM-numre til emnesortering af sager.⁵⁸ Strukturen, som i denne sammenhæng vil sige dokumenternes orden, og hvor meget man går op i den, er til gengæld ganske anderledes. På det canadiske arkiv foregår meget af processen ifølge Zhang via automatisk overførsel fra statens dokumenthåndteringssystem helt udenom eventuelle leverandører. Man håndterer arkivalierne som individuelle filer med tilhørende metadata, der kommer ind fra statens forskellige afdelingers biblioteker, som de selv har oprettet og organiseret i systemet,⁵⁹ og der er derfor ikke nogen fast struktur, som vi kender det i Danmark.⁶⁰ Afhængig af hvor frit et system man har haft, kan man dog også i Danmark have kommuner, der håndterer sager eller bruger systemer forskelligt fra hinanden.

PeDALS er det eksempel, der kommer tættest på at ligne dansk praksis. Man håndterer tilsyneladende stadig kun arkivalier i form af dokumenter, men det gøres i en SQL-database, hvor dokumenter og metadata kobles sammen. Dokumenternes

⁵⁵ Zhang, "Original order in digital archives," s. 175.

⁵⁶ Ibid. s. 178-79.

⁵⁷ Teknisk set har dokumenter også numre i danske arkiveringsversioner, men det er deres ID i databasen fremfor en klassifikation som et sagsnummer.

⁵⁸ KLE bruges af kommunerne og FORM af staten. Det fremgår af kontekstdokumentationen, hvilken systematik en arkiveringsversion bruger, hvis man har husket at spørge kommunen ifm. afleveringen.

⁵⁹ Zhang, "Original order in digital archives," s. 178.

⁶⁰ Her mener jeg strukturen i en arkiveringsversion, som Rigsarkivet har lavet meget faste rammer for.

metadata genereres automatisk, mens arkivarerne står for metadata på serie- og samlingsniveau. Her beskrives proveniens og administrativ historik⁶¹ - sammenligneligt med den kontekstdokumentation, man i Danmark indsamler til arkiveringsversioner. Her er der blot tale om, at arkivet skaber det, hvor man i Danmark hellere vil have dokumentation skabt i forbindelse med brug af et system.⁶²

Zhang kommer bl.a. ind på formålet med at bevare den originale orden, nemlig bevaring af arkivaliernes kontekst og at gøre arkivalierne lettere tilgængelige.⁶³ *Original order* findes ifølge Zhang stadig i de digitale arkiver, og den har stadig samme relevans som i papirarkiverne. Den har blot fået flere former. Nogle digitale arkiver er ikke så forskellige fra papirarkiverne, da de er organiserede i en hierarkisk mappestruktur. Det er typisk arkiver, der kommer fra private eller fællesdrev, der ikke har været en del af et dokumenthåndteringssystem⁶⁴ (eller ESDH-system i en dansk sammenhæng). Andre, og her er særligt tale om store offentlige arkiver, bruger netop dokumenthåndteringssystemer, som er organiserede ud fra metadata på dokumentniveau. Zhang argumenterer for, at den originale orden i arkiverne ikke kun handler om fysisk orden, men at et i lige så høj grad refererer til dokumenternes metadata og relationer til andre dokumenter i et system.⁶⁵ Hun nævner ikke relationelle databaser, men denne måde at organisere dokumenter på minder meget om dem. Der er dog forskelle, og hvis man tager udgangspunkt i Zhangs undersøgelse, har de først og fremmest at gøre med, hvordan arkivets struktur muliggør søgning.

Afhængig af arkivets struktur kan den originale orden enten hindre eller støtte fremsøgning af arkivalier. I Zhangs cases er arkivsystemerne alle sat hierarkisk op efter arkivskaber, serie og dokument som man kender det fra papirarkiverne, hvor man har forsøgt at bevare den orden, filerne lå i, der hvor de kom fra. Dette faciliterer tilgængelighed ud fra kontekst, igen som man kender det fra papirarkiverne. Det er derimod metadataene og det faktum, at dokumenterne alle har individuelle metadata, der for alvor gør de digitale arkiver søgbare på nye måder.⁶⁶

⁶¹ Zhang, "Original order in digital archives," s. 181-82.

⁶² Rigsarkivet, Vejledning til bekendtgørelse 128, s. 97.

⁶³ Zhang, "Original order in digital archives," s. 169.

⁶⁴ Ibid. s. 185.

⁶⁵ Ibid. s. 186.

⁶⁶ Ibid. s. 187-88.

Danske arkiveringsversioner er ikke sat op på denne måde, men derimod som relationelle databaser, hvor metadata kan ligge forskelligt og have forskellige relationer til data og dokumenter. I teorien er tilgangen ikke så forskellig. Da arkiveringsversionerne ikke ligger hierarkisk, har man slet ikke mulighed for at genfinde på traditionel vis, og man er derfor meget afhængig af metadata. Typiske metadata, der bruges til fremsøgning er CPR-numre, sagsnumre eller adresser.

En af Zhangs centrale pointer er, at de digitale arkivaliers struktur i form af relationer i et system er, hvad der i de digitalt skabte arkiver, karakteriserer den originale orden, og at disse relationer, når de bevares i form af metadata er nøglen til genfinding. Samtidig er den struktur arkivalierne ender med at blive bevaret i ofte dikteret af det system eller den software man bruger, og hvordan man bruger den.⁶⁷

Sidstnævnte er et aspekt af digital arkivteori, som Greg Bak har behandlet nærmere. Det kommer til at have relevans for denne undersøgelse, da dansk digital arkivering hovedsageligt bevarer systemdatabaser, hvis struktur forvaltningen ikke nødvendigvis selv har valgt.⁶⁸ Bak introducerer konceptet digital proveniens. Det adskiller sig fra traditionel proveniens ved at inkludere den hardware og software, som arkivalierne er skabt i og med. Bak deler den digitale proveniens op i tre dimensioner: Teknisk, social-teknisk og social.⁶⁹ Det er den tekniske dimension, der har mest at gøre med filformater og databasestruktur, mens det er den social-tekniske, der har at gøre med, hvordan information er kommet fra virkeligheden og ind i systemet - altså hvordan hardware og software er blevet brugt. Sammenhængen mellem hardware og software i dansk offentlig arkivskabelse og arkivering ville hurtigt blive en hel undersøgelse for sig, og jeg vil derfor se bort fra hardware i denne omgang.

Softwaresiden af den tekniske dimension har at gøre med, hvorfor software ser ud og fungerer, som det gør. Ifølge Bak er softwares udformning både betinget af hardware og den kultur det er skabt i,⁷⁰ hvilket nok ikke kommer som nogen overraskelse for historikere. I denne undersøgelses kontekst har softwaresiden at gøre med, hvorfor databaserne fra de offentlige IT-systemer ser ud som de gør. Bak nævner selv

⁶⁷ Ibid. s. 191-92.

⁶⁸ Man har selv valgt, at tingene skal organiseres i sager eller i et register. Man har også selv valgt, hvad der skal være i en sag eller et register. Men det er så vidt jeg ved leverandøren, der har valgt, hvordan data og dokumenter organiseres i rækker og tabeller.

⁶⁹ Bak, "Digital provenance," s. 852.

⁷⁰ Ibid. 852-54.

de første Apple-computere som eksempel på de kulturforskelle, der har haft indflydelse på udvikling af computerteknologi. Nærmere bestemt var computerteknologi i starten af den personlige computers æra et domæne for dedikerede entusiaster, der satte pris på åbenhed og brugerkontrol, her med Steve Wozniak som eksempel. Med Steve Jobs kom derimod et fokus på at sælge til masserne og på leverandørens kontrol over teknologien.⁷¹

Hvem der har kontrol over softwares udformning og funktioner overlapper med den social-tekniske dimension. Her kommer vi også for alvor nærmere den konkrete praksis i dansk offentlig forvaltning og arkivering, som dette speciale handler om. Den social-tekniske dimension handler om, hvordan software bruges, hvordan brug påvirkes af systemdesign, og hvordan brug påvirker systemdesign. Baks pointe er her, at det er kulturelt og teknisk bestemt, hvordan computere og software skal bruges, og hvad man bruger dem til.⁷² I en dansk sammenhæng f.eks. at man i den periode, jeg arbejder med, har brugt enten stationære eller bærbare PC'er med forskellige udgaver af Windows til offentlig sagsbehandling og økonomistyring. Da den offentlige forvaltning er en kunde med relativt meget magt sammenlignet med f.eks. privatpersoner eller mindre virksomheder, har den stor indflydelse på systemernes udformning og funktioner.⁷³ Der er derfor indflydelse på systemernes design og brug fra to meget forskellige dele af vestlig kultur: Offentlig bureaukrati og tech-sektoren på det private marked. Dertil kommer så, hvordan enkelte afdelinger og personer bruger systemerne.

Her kan man måske også tilføje et aspekt, som Bak ikke har nævnt: Den social-tekniske dimension kan måske også indeholde en vigtig skelnen mellem, hvilke data, der er indtastet, uploadet eller på anden måde skabt direkte af mennesker, og hvilke der er genereret af en computer. Enten som den lidt usynlige (for brugere) del af en database (fremmednøgler) eller som automatiske metadata. Særligt for datoer er det stadig blandet, om det er automatisk eller manuelt, men det kan navne også være via fremmednøgler, som vi skal se senere.

Man kan ikke bruge computere eller systemer på måder, der ikke i forvejen er gjort teknisk mulige. Det betyder dog ikke, at brugerne følger udviklernes intentioner. Ifølge Bak følger der en kultur med brug af computersystemer, der har at gøre med,

⁷¹ Ibid. 853-54.

⁷² Bak, "Digital provenance," 854-55.

⁷³ Kjellberg og Hall-Andersen, "Kvalitetssikring - i hvilken grad bør arkiverne styre myndighedernes digitale sagsdannelse," s. 47: I Danmark indkøber man typisk lokalt udviklede ESDH-systemer fremfor at bruge f.eks. amerikanske produkter, der kan være svære at tilpasse danske administrative forhold.

hvordan man har lært at bruge dem. Som eksempel bruger han overgangen fra DOS-systemer til GUI (graphical user interface, dvs. programmer med vinduer, hvis funktioner, man mest bruger en mus til at bruge) til smartphones, hvor alt foregår med touch. Med disse tre eksempler illustrerer han også udviklingen i, hvor meget teknisk kunnen, der skal til for at bruge en computer over tid. DOS krævende relativt meget af brugeren, mens smartphones er designede til at være så intuitive som muligt. De færdigheder, der skal til for at bruge et system kalder Bak for *tacit knowledge*.⁷⁴ De går tabt hos brugerne over tid, og det er derfor en del af den digitale proveniens at vide (og gerne dokumenterer) den *tacit knowledge*, som et systems brugere havde. Det dokumenterer man ikke direkte, som arkiveringen foregår i Danmark lige nu, men arkivbrugerne kan hjælpes til at gætte det via de skærbilleder og vejledninger, der bliver bevaret i kontekstdokumentationen. Forståelse for, hvordan et system blev brugt fra denne vinkel, vil også kunne gøre det lettere at forstå, hvorfor arkiverne ser ud som de gør.

Visning og søgning

Dokumenter i form af tekst, billeder, lyd og video er i princippet lige så enkle at åbne og læse (eller se/lytte til) i en arkiveringsversion, som de ville være i originalformatet. Dvs. deres bevaringsformater (med undtagelse af meget store billeder⁷⁵) kan åbnes i de fleste standardprogrammer til visning af lyd og billede. I praksis er det dog ikke helt så enkelt. Dokumenter eksisterer side om side med data og skal i første omgang tilgås i samme program, og først senere hentes ud derfra.

Data er en lidt mere kompliceret sag. Som nævnt lagres tabeller som XML-filer. De kan åbnes i ethvert program til tekstfiler (f.eks. Notepad eller en browser), men her ligner de ikke tabeller, som en almindelig arkivbruger nok vil forvente at se. For at vise tabeller som tabeller skal filerne åbnes i et såkaldt genrejsningsprogram, der kan omsætte XML-formatet til et tabelformat, der minder mere om det, man finder i f.eks. excel. Hos KOMDA åbnes arkiveringsversioner i programmet ASDA, der også bruges

⁷⁴ Bak, "Digital provenance," 855-56.

⁷⁵ Ibid. s. 52.

til søgning. Som arkiveringsversioner fungerer nu, er arkivernes genrejsningsprogrammer centrale for arkiveringsversioners tilgængelighed. For Aalborg Stadsarkivs vedkommende kan de slet ikke tilgås udenfor ASDA.⁷⁶

ASDA

Hos KOMDA bruger man programmet ASDA til visning af og søgning i arkiveringsversioner. ASDAs forskellige funktioner, fordele og ulemper vil dukke op i løbet af denne undersøgelse, og det er derfor værd at give programmet en kort introduktion her.

Når man åbner en arkiveringsversion i ASDA, er der op til fem faner: Søgninger, tabeller, kontekstdokumenter, datamodel og dokumenter (hvis der er nogen). Til følgende beskrivelse tager jeg udgangspunkt i de arkiveringsversioner, jeg har brugt til specialets undersøgelse, men beskrivelserne er i princippet gældende for alle arkiverede databaser.

Under *søgninger* findes SQL'en til søgepakkerne, som ikke kan bruges ad denne vej. Søgepakkerne har sin egen side andetsteds i ASDA (se nedenfor). Derudover findes også leverandørens SQL-forespørgsler, som bekendtgørelsen kræver,⁷⁷ og de øvrige søgninger, som arkivet har oprettet. Dem vender jeg tilbage til.

Under *tabeller* finder man samtlige tabeller i arkiveringsversionen. De kan ordnes enten i numerisk (efter tabelnummer) eller alfabetisk rækkefølge (efter tabelnavn). Dobbeltklikker man på en tabel dukker den op til højre for listen. Her kan man vise relationer både ind og ud af tabellen. Fra relationerne kan man vælge en eller flere tabeller, der kan kobles på. Dette kan gøres i flere led. Til højre for tabellerne findes søgefelter til kolonnerne i den aktive tabel (den man sidst har klikket på). Klikker man søg uden at indtaste noget får man de første 1000 rækker (poster) i tabellen inklusiv i de relaterede tabeller, man har valgt. Søgeresultatet findes nederst i vinduet og har form af en tabel som man kender den i excel og lignende.

I fanen *kontekstdokumenter* finder man kontekstdokumenterne i nummereret rækkefølge. Den tilgængelige metadata er dokumentID, titel og beskrivelse.

⁷⁶ Man kan godt lave udtræk derfra, som kan sendes til f.eks. en borger eller sagsbehandler, som så ikke er i kontakt med ASDA, men arkivarerne kan som udgangspunkt kun bruge ASDA til at søge i arkiveringsversionerne.

⁷⁷ BEK nr. 128 af 12/02/2020, bilag 6.D.

I fanen *datamodel* findes også en liste med samtlige tabeller i arkiveringsversionen, men uden at man kan se data deri. Her står i stedet metadata om tabellerne: Tabelnummer, tabelnavn og -beskrivelse, antal poster (rækker), antal felter (kolonner) og tabelstørrelse. For hver enkelt tabel står kolonnernes nummer, navn, datatype, original datatype, beskrivelse, om der er tale om en primærnøgle eller fremmednøgle, om der er tale om en kolonne med dokumenter eller sager, og om det er tilladt at have NULL-værdier (tomme felter) i kolonnen. Til hver kolonne står der også, hvis man klikker på dem, hvilke relationer de har både ind og ud ad tabellen. Med andre ord præsenteres her den information, som findes i *tableIndex*.

I fanen *dokumenter* findes samtlige dokumenter i arkiveringsversionen i nummeret rækkefølge bortset fra kontekstdokumenterne. Til dokumenterne følger basal metadata i form af nummer (DokID), eventuelt nummer på et "parent"-dokument⁷⁸ (ParentID), nummer på det medie dokumentet ligger på (MedieID), nummer på dokument-samling,⁷⁹ det oprindelige filnavn og filtypen i arkiveringsversionen (som regel TIFF). Basal metadata, der kunne være relevant for en arkivbruger er dog ikke med her: Dato, forfatter/ophav, relationer til personer eller hvilke tabeller de kommer fra. Den er i mine cases ingen let tilgængelig oversigt over, hvilke tabeller der indeholder dokumenter. Hvis man ellers kan finde den tabel, hvor dokumenterne står i, og hvor man således også kan se deres kontekst, er dokumentfanen ikke særlig brugbar. Som jeg senere vil komme ind på, er søgning på dokumenter heller ikke et særlig godt udgangspunkt for søgning i den type system, jeg har brugt som case. Det kan måske forholde sig anderledes i små dokumentbaserede arkiveringsversioner.

SQL-søgninger

ASDA er først og fremmest et søgeværktøj. Som udgangspunkt bruges ASDA til at finde enkelte sager eller mindre grupper af relaterede sager. F.eks. sager på en enkelt person ad gangen. I de arkiveringsversioner, jeg bruger som case, har KOMDA dog også tilføjet præ-lavede søgninger på lister (indexes). Dem skal vi se nærmere på senere. Til at finde enkelte sager er der også sat præ-lavede søgninger op, således at den enkelte arkivar, der betjener arkivets brugere ikke manuelt skal skrive SQL-

⁷⁸ Et relateret højere rangerende dokument. Et eksempel kan være en email med tilhørende vedhæftede filer.

⁷⁹ Ifølge BEK nr. 128 af 12/02/2020 bilag 4.G må der være op til 10.000 dokumenter per samling. Et dokument kan dog godt bestå af flere filer i denne sammenhæng.

statements hver gang. Det er i det hele taget ikke muligt for en bruger af ASDA udenfor KOMDA at skrive sine egne søgning. Ifølge KOMDA har det at gøre med, at man også kan bruge SQL til at ændre i data og tabeller, hvilket selvfølgelig ikke går an i en arkiveringsversion.

Søgepakker

Søgepakker er SQL-søgninger på tværs af arkiveringsversioner. Søgepakkerne er for Aalborg Stadsarkiv delt op i forskellige sagstyper. Til dette speciales undersøgelse vil jeg kun bruge søgepakken til "omsorgssager", da det er denne søgepakke, der inkluderer det system, jeg bruger som case. Man kan dog også søge på f.eks. et CPR-nummer på tværs af alle arkiveringsversioner, man har adgang til i ASDA. Søgepakkesøgningerne oprettes efter ønske fra de deltagende arkiver⁸⁰ og afspejler således, hvordan arkiveringsversionerne bruges i dag.

Kategorier

I ASDA har KOMDA kategoriseret samtlige arkiveringsversioner ud fra basal metadata: Tid, arkivskaber og kategori.

Tid er arkiveringsversionernes slutår, og her måler man på slutåret for data. Der kan således komme en gammel database eller filsamling ind i 2025, som kommer til at stå til f.eks. 2018. Det betyder at f.eks. kontekstdokumenter godt kan være fra senere end arkiveringsversionens slutår.

Arkivskaber er meget overordnet. Dvs. for kommunale afleveringer er det kommunen, der er arkivskaber og ikke f.eks. den specifikke afdeling i kommunen, der har brugt systemet. I Aalborg Stadsarkivs samling er der afleveringer fra både Aalborg Kommune og alle de mindre kommuner, der blev en del af Aalborg efter 2007. Derudover kan arkivskaber også være foreninger eller andre private aktører, der har afleveret til arkivet.

Kategori vil sige systemets overordnede formål. I Aalborg Stadsarkiv er følgende kategorier repræsenterede:

- Arbejdsmarked
- Byggesager
- Fritid

⁸⁰ Oplyst via email fra KOMDA.

- Institution
- Økonomi
- Omsorg
- Øvrige
- Sager og dokumenter
- Social
- Sundhed
- Uddannelse

Kategorierne har KOMDA selv fundet på og eksisterer dermed uafhængigt af Rigsarkivets bekendtgørelser, og er heller ikke nævnt i arkivbeskrivelsesfilerne til de enkelte arkiveringsversioner. De skaber dog et ganske godt overblik over arkiveringsversionerne. For de mindre kategorier som Omsorg, hvor der i Aalborg kun er Care-systemet, gør kategoriseringen det også meget nemt og hurtigt at finde de arkiveringsversioner, man leder efter uden at skulle indtaste navnet (eller huske hvad det hedder, hvis man har glemt det). I arkivbeskrivelsesfilerne skal alle arkiveringsversioner have en FORM-klassifikation.⁸¹ Det er sammenligneligt med et KLE-nummer, og henviser således til et emne i offentlig forvaltning. KOMDAs kategorisering svarer ikke til hverken FORM eller KLE, men jeg nævner dem her, da det fungerer som en meget forsimplet udgave af en emneinddeling af arkiveringsversioner. For den størrelse samling Aalborg Stadsarkiv har, giver denne simple inddeling mere praktisk mening end de mere komplicerede emnesystematikker.

Udvælgelse af arkiveringsversioner

Der ligger på skrivende stund 151 arkiveringsversioner i Aalborg Stadsarkivs samlinger hos KOMDA. Man kan naturligvis ikke undersøge dem alle på én gang, og jeg har derfor udvalgt et begrænset antal. Der er overordnet set to muligheder for begrænsning: Man kan vælge et repræsentativt udsnit på tværs af fagområder,⁸² eller man kan vælge et enkelt fagområde. Jeg har valgt sidstnævnte, og mit valgte fagområde er ældre- og handicapområdet. I Aalborg har man kun brugt ét system til det igennem de

⁸¹ BEK nr. 128 af 12/02/2020, figur 6.1.

⁸² I praksis vil det sige på tværs af kategorier i ASDA.

sidste snart 30 år. Det samme er tilfældet i de kommuner, der blev en del af Aalborg Kommune efter 2007. Det begrænser mit perspektiv, men eliminerer til gengæld faktorer der har mere at gøre med forskelle i systemdesign og sagsbehandling end med arkivering. Systemet har været arkiveret fem gange. To er fra de tidligere Hals og Sejlfjord kommuner, som jeg ikke har inkluderet. Fra Aalborg er systemet arkiveret i 2003, 2009 og 2023. Arkiveringsversionernes officielle årstal for arkivering både i metadata (arkivbeskrivelsesfilen) og ASDA (som bruger metadataen) er bestemt ud fra indholdets datering. Kontekstdokumentationen kan dog sagtens være fra senere, og det kan også tage flere år at sætte en arkiveringsversion i stand. Jeg har derfor valgt at forholde mig til den praktiske virkelighed og inkludere de reelle arkiveringsår (i parentes).

Foruden de logistiske årsager har jeg valgt KMD Care, da arkiveringsversionerne er begrænsede i antal og samtidig strækker sig over lang tid. To er derudover eksempler på migreringsstrategien i praksis, og jeg har derfor mulighed for at undersøge, hvordan ændringer i forbindelse med migrering til nyt format har påvirket struktur, indhold og dokumentation.

Inden selve analysen vil jeg starte med kort at introducere systemets formål og de tre arkiveringsversioner. Her vil jeg nævne grundlæggende metadata som antal tabeller og poster, samt hvilken bekendtgørelse, de er arkiveret efter.

Intro til systemet

KMD Care har igennem hele perioden haft samme formål, nemlig håndtering af ydelser, hjælpemidler, hjemmehjælp og pleje, boligstyring af plejeboliger (plejehjem), funktionsvurderinger og medicinering til ældre og handicappede.⁸³ Det har også været brugt af den samme afdeling i kommunen (ældre- og handicapforvaltningen) igennem hele perioden.⁸⁴ Systemet bruges af en bred skare af kommunale ansatte i ældre- og handicapforvaltningen, herunder sagsbehandlere, sygeplejersker, hjemmehjælpere, SOSU'er, fysioterapeuter, mv. Systemets dokumentation indeholder ingen lister over

⁸³ Jf. Arkivbeskrivelsen i ASDA. Jeg går ud fra, at det meste information i denne beskrivelse kommer fra arkivbeskrivelsesfilen, men det er ikke oplyst. Det er ikke alt information fra denne fil, som er synlig i ASDA.

⁸⁴ I første arkiveringsversion er det også afdelingen, der står som arkivskaber i afleveringsbestemmelsen. I de nyere udgaver er det Aalborg Kommune.

typer af ansatte, der har brugt det, og ej heller antal, men systemet indeholder en tabel med stillingsbetegnelser, som kommunen sporadisk har brugt. Den kan give et fingerpeg om typen af ansatte, men desværre ikke nogen præcis information om antal, da det kun er et fåtal af de ansatte, der er blevet tilknyttet denne tabel. Care-systemet har ikke sager, som der er tradition for i ESDH-systemer, men organiserer derimod persondata alene omkring CPR-nummer for klienter (borgere) og personalenummer for ansatte.

KMD Care 2002 (2003)

Første arkiveringsversion af KMD Care er fra 2003. Den indeholder data fra 1996 til 2002, men aflevering fandt sted i 2003 ifølge afleveringsbestemmelsen. Denne arkiveringsversion har 260 tabeller med lige under 60 millioner poster, men ingen dokumenter. Det betyder ikke nødvendigvis, at man ikke brugte dokumenter i denne del af forvaltningen. Da den første bekendtgørelse først kom i 2004 er denne arkiveringsversion lavet ud fra Cirkulære nr. 25 af 08/03/2002 (IT-afleveringscirkulæret),⁸⁵ som er stort set identisk med bekendtgørelse 342.

KMD Care 2006 (2009)

Anden arkiveringsversion er ifølge afleveringsbestemmelsen fra 2009 og indeholder data fra 1996-2006. Arkiveringsversionen følger bekendtgørelse 342 (2004) og har 288 tabeller med lige over 22 millioner poster og 135.936 dokumenter. Det står ikke helt klart om, det blev muligt at gemme dokumenter i systemet, eller om dokumenterne er data, der blev konverteret i tilfælde, hvor felter overskred grænsen på 2048 tegn. Ifølge Falk og Hansen begyndte KOMDA dog først at modtage arkiveringsversioner med dokumenter, som ikke blot var konverteret tabelindhold, efter 2010.⁸⁶ Arkivbeskrivelsesfilen indeholder et felt til at vise, om en arkiveringsversion indeholder dokumenter. Det kunne måske være en fordel, hvis der også var et felt til at sige, om dokumenter er konverterede fra data, eller om der kunne gemmes dokumenter i driftsversionen.

Både KMD Care 2002 og KMD Care 2006 blev migreret fra bekendtgørelse 342s format til bekendtgørelse 1007s format i 2014.⁸⁷ Ingen af de to er migreret til 2020-format,

⁸⁵ Oplyst via email fra KOMDA.

⁸⁶ Falk og Hansen, "Digitale arkivalier: Fra sorte bokse til big data," s. 75.

⁸⁷ Dokumentationen, som blevet tilføjet i forbindelse med migreringen, er dateret til 2014.

da det gældende format, som beskrevet tidligere, blot er en mindre opdatering af bekendtgørelse 1007s format.

KMD Care 2020 (2023)

Tredje og nyeste arkiveringsversion er fra 2023 og indeholder data fra 1996-2020. Den er arkiveret efter den nuværende bekendtgørelse 128 (2020). Den har 448 tabeller med 240 millioner (næsten 241) poster og 612.743 dokumenter. Denne arkiveringsversion har selvsagt ikke været migreret til nyt bevaringsformat endnu.

Dokumentation

Som nævnt i alle tre bekendtgørelser, skal arkiveringsversioner indeholde dokumentation, der beskriver systemernes "administrative funktion samt struktur og funktionalitet" og dokumentation af arkiveringsversionens tabeller og relationer.⁸⁸ Førstnævnte findes kun i kontekstdokumenterne, mens sidstnævnte både findes i kontekstdokumenterne og i en metadatafil (tabelindeks). Derudover indeholder alle arkiveringsversioner også en arkivbeskrivelsesfil, der også er en type dokumentation. I dette afsnit er det kun kontekstdokumenterne, det handler om, da den øvrige dokumentation ikke er tilgængelige som selvstændige filer i ASDA, men derimod er indbygget i måden, arkiveringsversionerne præsenteres på. Der er således et minimumskrav for dokumentation. Det er tilgængeligt nogle ret uspecifikke krav, men de er, som vi skal se, blevet tolket relativt ens over tid.

KMD Care 2002 (2003)

Der er syv kontekstdokumenter til denne arkiveringsversion. Som i alle andre arkiveringsversioner er et af dokumenterne afleveringsbestemmelsen. I denne arkiveringsversion er det her, man finder listen over kasserede tabeller, hjælpetabeller, som blev oprettet ved arkivering, samt listen over kontekstdokumenter ("generelle informationer"). Det er også heri, at man får den egentlige datering på arkiveringsversionen. Det sidste dokument i rækken er tilføjet arkiveringsversionen senere. Det er et dokument

⁸⁸ Stort set alle arkiveringsversioner indeholder et kontekstdokument med tabel- og feltbeskrivelser. Her menes faktisk tabel- og kolonnebeskrivelser og der henvises nærmest aldrig til enkelte felter. Rækker i tabeller kaldes som regel poster.

med dokumentation af migrering, og her nævnes det, at der er tale om migrering fra 342-format til 1007-format.

Afleveringsbestemmelsen er underskrevet af repræsentanter fra Aalborg Stadsarkiv og fra arkivskaberen, der i dette tilfælde er Ældre- og Handicapforvaltningen i Aalborg Kommune. Leverandøren nævnes slet ikke, hverken som part i afleveringen eller som en del af systemets navn. Det omtales blot med det mere uformelle "Care-systemet". Jeg finder denne detalje relevant, fordi de øvrige kontekstdokumenter afslører, at systemet slet ikke havde noget at gøre med KMD i 2003. Leverandøren var Rambøll Informatik og systemets egentlige navn på dette tidspunkt var Rambøll Care. Rambøll Informatik blev opkøbt af KMD i 2011.⁸⁹ Rambøll Care fremgår af Rigsarkivets vejledende kassationsliste for primærkommuner (dvs. før kommunalreformen i 2007), hvor det stod til bevaring, mens KMD Care ikke fandtes på dette tidspunkt. Det gør et andet KMD-system til gengæld, nemlig KMD Omsorg. Det samme gør Rambøll Care. Hvis man således søger efter arkiveringsversioner ud fra navne på bevarede kommunale systemer i Rigsarkivets (vejledende) kassationsliste,⁹⁰ eksempelvis som forsker, vil man ikke kunne finde systemet fra Aalborg Kommune. Man skal i stedet på forhånd vide, at man i Aalborg har ændret navnet på arkiveringsversionerne.⁹¹

Rambøll er forfatter på alle øvrige kontekstdokumenter. Det første er en "administrativ beskrivelse af CARE-systemet i Aalborg Kommune," hvori systemets forskellige moduler meget kort beskrives. Følgende moduler er nævnt:

- Klientmodul
- Sygeplejemodul
- Bistand/plejemodul
- Personalemodul
- Rengøringsmodul
- Hjælpemiddelsmodul

⁸⁹ Breinstrup, "KMD køber Rambøll Informatik," juli 2011.

⁹⁰ Rigsarkivet, Vejledende kassationsliste for primærkommunernes IT-systemer. Systemer til ældre- og handicapplejen står på s. 4.

⁹¹ Denne problematik vil dog ikke blive et rigtigt problem, før der er nok arkiveringsversioner til, at man kan miste overblikket. Der går nok nogle årtier endnu, men det er heller ikke længe på et arkiv. Som det er nu skal man på forhånd vide, hvilket arkiv en arkiveringsversion ligger hos for at kunne søge efter den. Ifølge KOMDA er der kun tale om en ændring i det viste navn og ikke arkiveringsversionens rigtige navn, men det rigtige navn fremgår ikke, og det står heller ingen steder, at der overhovedet findes to slags navne. Det kan også blive en udfordring, hvis man på et tidspunkt får mulighed for at søge på tværs af kommuner, der har brugt de samme systemer, som er arkiveret forskellige steder.

- Boligstyringsmodul

Modulerne går igen i Rambølls vejledning, som er næste dokument. Her udfoldes systemets funktioner og indhold nærmere. Sammen med ER-diagrammet⁹² og dokumentet med tabel- og feltbeskrivelser er det der, hvor man får mest information om, hvad man kan finde i arkiveringsversionen, og hvordan man skal finde det. Foruden den generelle brugervejledning er der også en vejledning til at generere rapporter baseret på systemets data. Den er dog meget tydeligt ikke beregnet til nye brugere; der er ingen forklaringer på, hvad vejledningen overhovedet drejer sig om. At der er tale om generede rapporter baseret på systemets data, må man gætte sig til ud fra den generelle vejledning og rapportvejledningens opsætning.

Overordnet er dokumentationen i denne tidlige arkiveringsversion meget begrænset, men dog brugbar til at finde og forstå systemets indhold. Særligt tabel- og feltbeskrivelserne og ER-diagrammet, der kan give et visuelt overblik over databasen. Den tekniske dokumentation kan give umiddelbare praktiske informationer om tabellernes indhold og relationer, mens systemets vejledning gør det lettere at forstå eller gætte (afhængig af hvor god vejledningen er), hvad indholdet har været brugt til.

Der er ingen dokumenter i denne arkiveringsversion. Ifølge brugervejledningen var der i systemet mulighed for i stedet at henvise til dokumenter udenfor systemet. Der står dog ikke, hvor disse henvisninger ligger, eller hvor dokumenterne skal findes. Systemet har dog en tabel, LINKS, hvor dokumenthenvisninger til andre systemer ifølge tabelbeskrivelsen findes, men den er tom. Som bruger kan man ledes til at tro, at man slet ikke brugte dokumenter til arbejdet med ældre- og handicappede mellem 1996 og 2002. Det virker dog usandsynligt i dansk bureaukrati. Det er langt mere sandsynligt, at man ikke har kunne oprette en snitflade til kommunens ESDH-system og derfor har journaliseret parallelt, at man har arkiveret parallelt på papir, eller at man har kasseret dokumenter. Det står ingen steder i kontekstdokumentationen, hvilken situation der er tale om, og hvor man kan finde eventuelle dokumenter. I arkiveringsversioner med dokumenter er det et krav, at "sagligt sammenhørende" dokumenter skal forblive sammenhørende,⁹³ men det var og er ikke et krav, at sagligt sammenhørende dokumenter fra IT-system og papir skal forblive forbundne. Det er et krav, at papirdokumenter, der

⁹² Entity-relations diagram. Det er et visuelt overblik over tabellerne og deres relationer.

⁹³ BEK nr. 342 af 11/03/2004, s. 9.

er registrerede i et IT-system, skal have deres registrering markeret med en *functionalDescription*, der inkluderer oplysning af, at der er tale om papirdokumenter.⁹⁴ Hvis de ikke er registrerede, er der dog ikke meget at gøre. Det er også et krav, at sagligt sammenhørende dokumenter fra forskellige IT-systemer skal forblive forbundne. Det er bare ikke særlig tydeligt i bekendtgørelserne eller det forudgående cirkulære. Havde LINKS-tabellen f.eks. været brugt til at henvise til et bevaringsværdigt ESDH-system, skulle disse links være bevarede. Havde der været links til et ikke-bevaringsværdigt system, skulle dokumenterne overføres til Cares arkiveringsversion. Det blev bare først relevant med KMD Care 2006 (2009).

KMD Care 2006 (2009)

Til denne arkiveringsversion er der 17 kontekstdokumenter, hvilket altså vil sige mere end en fordobling siden 2003. Af afleveringsbestemmelsen fremgår bekendtgørelse 342. Udover denne tilføjelse er informationen den samme som i første aflevering: Tilføjede hjælpetabeller og listen over kontekstdokumenter. Kasserede tabeller er blevet flyttet til et andet dokument.⁹⁵ Rambøll Informatik er også blevet tilføjet som leverandør i afleveringsbestemmelsen i denne udgave. Listen over anvendte moduler er også rykket til dette dokument. Moduler er de samme som i den tidligere aflevering, der er tilføjet tre nye: Forebyggende [hjemmebesøg], Specialcenter for voksenhandicappede og Center for sindslidende.

Det næste dokument er en "standardafleveringsmodel" udviklet af Rambøll Informatik og Statens Arkiver (Rigsarkivet). Systemet havde, som man kan se, stadig ikke noget med KMD at gøre ved denne aflevering. Afleveringsmodellen er udviklet til at blive brugt til alle afleveringer af systemet, så her kan man gå ud fra, at arkiveringsversioner af Care er nogenlunde ens uanset hvilket arkiv, der er afleveret til. Dokumentet beskriver en række af de mest indgribende ændringer, der er foretaget ved konvertering fra drift til arkiveringsversion. Her findes også listen over de kontekstdokumenter, som Rambøll stod for (de fleste af dem). Det samme gør listen over kasserede tabeller og felter. Dokumentet beskriver også, hvordan felter, der har indeholdt meget lange tekststykker,⁹⁶ bevares separat som dokumenter udenfor tabellen. Her

⁹⁴ Rigsarkivet, Vejledning til bekendtgørelse 128, s. 109.

⁹⁵ Kontekstdokument 6

⁹⁶ Med datatypen LONG, som ikke er tilladt ifølge BEK nr. 342 af 11/03/2004.

er altså tale om, at indhold ikke blot skifter datatype eller format, men også bliver fjernet fra sin oprindelige placering i databasen. Der ændres således på databasens oprindelige orden og det kan potentielt blive sværere at finde disse oplysninger. På dette tidspunkt bevarede man sådanne dokumenter både som TIFF og TXT. Andre felter med lange tekster over 2048 tegn, men med en tilladt datatype (VARCHAR), deles op i to felter. Denne information går igen i det kontekstdokument, der beskriver forskelle mellem driftsversionen og arkiveringsversionen,⁹⁷ hvor de specifikke tabeller, det drejer sig om står oplistet.

I systemet kunne man også linke til dokumenter, der lå i andre systemer, ligesom i den første udgave. Links skal i denne sammenhæng forstås som en snitflade til et andet system. Standardafleveringsmodellen henviser her til ESDH-systemer som eksempel. Overstående dokument beskriver, hvordan disse links skal håndteres: Hvis de ligger i et andet bevaringsværdigt system skal en passende henvisning dertil (f.eks. sagsnummer) indsættes i LINK-feltet i LINK-tabellen. Hvis ikke skal links bevares og dokumenterne hentes ind og bevares i arkiveringsversionen for Care. Hvis kommunen ikke har brugt funktionen, tages den tomme tabel med som verifikation,⁹⁸ hvilket er tilfældet med denne arkiveringsversion.

I standardafleveringsmodellen står der også, at "arkiveringsfrekvensen" er "et til enhver tid opdateret system – ikke et periodeopdelt system." Det er derfor alle versioner har data, der går tilbage til 1996.

De øvrige kontekstdokumenter fra Rambøll er tabel- og feltbeskrivelser, ER-diagram og dokumenter med beskrivelser af forskelle på den generelle udgave af systemet, og den specifikke udgave Aalborg Kommune har brugt, samt forskelle på driftsversionen og arkiveringsversionen. Med andre ord, er der her tale om den tekniske dokumentation. Kontekstdokumenterne nr. 7 (Forskelle på den generelle driftsversion og kommunens faktiske driftsversion) og 8 (Forskelle på den generelle arkiveringsversion og kommunens faktiske arkiveringsversion) beskriver kommunespecifikke forskelle fra den generelle udgave af systemet. Der står ikke andet deri end hvilke moduler, kommunen har brugt, og hvor man derfor kan forvente at finde indhold fra. Eventuelle kommunespecifikke forskelle i brug af systemet er der ingen dokumentation for.

⁹⁷ Kontekstdokument nr. 6.

⁹⁸ Kontekstdokument nr. 2. Rambøll Informatik og Statens Arkiver, Standardafleveringmodel.

Tabel- og feltbeskrivelserne og ER-diagrammet er klart det mest brugbare dokumentation, hvis man skal finde noget uden hjælp fra arkivets præ-lavede søgninger. Dem vil jeg vende tilbage til senere. Til den tekniske dokumentation hører også en liste over leverandørens præ-lavede SQL-søgninger og en beskrivelse af, hvordan data automatisk kan indsættes i Word-dokumenter.

Til denne arkiveringsversion følger der også tre vejledninger med. Vejledningen til at generere rapporter er med igen, og denne gang med information om rapporternes funktion og en smule om hvilke data, der bruges i hver enkelt. Brugervejledningen er også med igen. Derudover er der en "kommunespecifik" vejledning til modulet "forebyggende hjemmebesøg", der blev tilføjet siden forrige aflevering. Denne vejledning er med to gange som to forskellige (men ens) dokumenter med hhv. kommunespecifik vejledning og "kommunens supplerende administrative beskrivelse af CARE". Havde der været arkiveret efter en af de to nyere bekendtgørelser ville man formentlig kun have haft den med én gang og givet den to kontekst kategorier.

KMD Care 2006 blev migreret mere eller mindre samtidig med KMD Care 2002, og der følger tre dokumenter med om denne proces. Udover den information, der også fandtes til 2002-udgaven, beskriver de også fejl i relationer, der er fundet i både denne arkiveringsversion og i afleveringer fra de tidligere Hals og Sejlflod kommuner, der var blevet en del af Aalborg Kommune på tidspunktet for migreringen (2014).

KMD Care 2020 (2023)

Der er 15 kontekstdokumenter til denne arkiveringsversion. Afleveringsbestemmelsen er naturligvis med igen, og denne gang er der arkiveret efter den gældende bekendtgørelse (nr. 128/2020). Ved denne aflevering var systemet overgået til KMD, der derfor har leveret en del af dokumentationen. Dokumentationens forfatter fremgår af bilag A i afleveringsbestemmelsen. Heri står en liste med 17 dokumenter. De manglende dokumenter fra listen er: Konverteringsbeskrivelse og KOMDAs testsvar og bemærkninger, som begge skulle have været leveret af Aalborg Stadsarkiv eller KOMDA.

Til denne aflevering er der kommet mere information med om den organisatoriske og administrative kontekst, end der var i de to tidligere afleveringer - eller det ser i hvert fald ud til at have været intentionen! Gengangerne udover afleveringsbestemmelsen er tabel- og feltbeskrivelse for driftsversionen og et ER-diagram, der viser både

drift- og arkiveringsversion vha. farverne på tabellerne⁹⁹ samt et notat med en beskrivelse af ændringer foretaget ved konverteringen til arkiveringsversion. Derudover var det intentionen også at inkludere en oversigt over relevante love for systemet, et dokument med snitfladebeskrivelser, et selvstændigt ER-diagram for arkiveringsversionen og brugervejledningen til systemet. Der eksisterer dokumenter i ASDAs liste for alle fire, men to af dem er erstattet af dummy-dokumenter, mens et tredje giver en fejlmeddelelse. Brugervejledningen er i stedet en vejledning til systemets app, Smart-CARE, der ser ud til at have fungeret som en lite-udgave af systemet på mobile enheder. Når der kun er en vejledning til mobile enheder med, kan man få den opfattelse, at systemet er overgået til *kun* at være til mobile enheder, hvilket ikke er tilfældet.¹⁰⁰ Den "rigtige" brugervejledning er ikke med. Det er til gengæld et andet sæt vejledninger blandt de nytilkomne dokumenter: En samling meget detaljerede vejledninger til sagsbehandlingen indenfor en lang række opgaver, som systemet har været brugt til. Uden brugervejledning eller skærmbilleder er der dog ingen umiddelbar forbindelse mellem disse beskrivelser og arkiveringsversionens indhold.

En anden tilkommer er en oversigt over Aalborg Kommunes ældre- og handicapforvaltning mellem 1998 og 2014. Går man ud fra, at 2014-udgaven stadig er gældende, vil det sige stort set hele systemets levetid. Også kommunens kvalitetsstandarder for hjemmepleje og træning med udgangspunkt i relevante love er ny. Dette dokument er et godt eksempel på, at bekendtgørelsens punkt om dokumentation af kvalitetssikring er svært at forstå, da det kan forstås både som kvalitetssikring af kommunens arbejde og som kvalitetssikring af journaliseringen.¹⁰¹ Dokumentet er sammen med kommunens vejledninger til sagsbehandling også relevant for at forstå lovgrundlaget for systemet, da oversigten over relevante love og paragraffer, som nævn ovenfor, mangler.

⁹⁹ Kasserede tabeller er markeret med rødt.

¹⁰⁰ Hverken sagsbehandling eller data har ændret sig nok til, at det giver mening at alt foregår i en app. Derudover er der i dokumentet med tekniske bemærkninger beskrevet, at de fleste tabeller, der har med appen at gøre, er kasserede. Det lægger meget tydeligt op til, at appen kun var et supplement til systemet.

¹⁰¹ BEK nr 128 af 12/02/2020, bilag 6, figur 6.2, kategori 1.i. I Rigsarkivets vejledning til bekendtgørelsen (s. 100) er det forklaret, at man er ude efter kvalitetskontrol af data og indtastning i systemet og ikke nødvendigvis den opgave, der udføres uden for computeren.

Opsamling

Generelt er der kommet mere dokumentation over tid, både administrativ og teknisk. Det ser ud til, at man har lavet en liste over den dokumentation, man ønsker at tage med fra starten (i afleveringsbestemmelsen) og så forsøgt at få fat i det. Det er dog ikke lykkedes for dem alle, og 2020-udgaven mangler der dokumenter, som man havde intentioner om at tage med. Ifølge KOMDA er denne observation korrekt. Indsamling af kontekstdokumentation foregår således, at man udarbejder listen til afleveringsbestemmelsen først, og så ændrer man ikke på den senere, da afleveringsbestemmelsen betragtes som et historisk dokument.¹⁰²

Meget af informationen i kontekstdokumentationen den samme, selvom det har ændret form og antal dokumenter over tid. De mest basale informationer, der er nødvendige for at kunne bruge data, er til stede i alle tre versioner af KMD Care: Tabel- og feltbeskrivelser, ER-diagram og overordnet information om datas administrative kontekst. Der er til gengæld meget *lidt* information om arkivskaber, og der er kun begrænset sammenhæng mellem den tekniske og administrative dokumentation.

Der er generelt meget lidt dokumentation af systemernes brugerflade, og dermed hvordan systemets brugere interagerede med databasen. Der findes skærmbil- leder i vejledningerne, men kun i begrænset omfang. Det kan lede til situationer som med den seneste udgave, hvor manglen på en brugervejledning kan så tvivl om, hvilken platform systemet har været brugt på (den tekniske proveniens). I dette tilfælde kan man gætte det ud fra de tidligere afleveringer, hvor systemet meget tydeligt er til Windows, og ud fra hvordan brugen af systemet omtales i den øvrige dokumentation. F.eks. nævnes KMD Care og SmartCARE som to forskellige systemer et enkelt sted i vejledningen til SmartCARE. Sagsbehandlingen som den beskrives i kommunens vej- ledninger dertil, lader også til at være for kompleks til noget en typisk app vil kunne bruges til. For at kunne gætte rigtigt kræver det dog kendskab til de tekniske aspekter af systemernes digitale proveniens, som ikke er dokumenteret, men som er relativt almindelig viden for personer i dag. Det er det dog ikke nødvendigvis for personer om 100 år. God dokumentation er en forudsætning for god tilgængeliggørelse i teknisk forstand. Jo bedre dokumentationen er, jo mindre ekspertviden indenfor IT kræver det at kunne tilgå og bruge digitale arkivalier.

¹⁰² Oplyst via email fra KOMDA.

Kontekstdokumentation og metadata er nøglen til tilgængelighed for digitalt skabte arkivalier, da data ofte ligger meget spredt og i en struktur, der kan være svær at finde rundt i og få overblik over på egen hånd. For kort at sammenligne med papirarkivalierne kan vi som historikere huske, at tilgangen til dem er relativt enkelt: Man åbner den relevante kasse, eventuelt finder det relevante læg, og så starter man fra toppen og læser indtil, man har fundet det, man ledte efter. Det kan man ikke med en arkiveringsversion. De har ikke et fast start- og slutpunkt og de har som regel heller ikke en fast rækkefølge, man skal læse i. Der kan også være flere millioner poster fordelt på flere hundrede tabeller, der kan være alt fra et enkelt tal til en halv sides ren tekst.

Sammenhængen mellem administrativ praksis og databasen

Rigsarkivet anbefaler, at man med skærbilleder forklarer sammenhængen mellem et systems oprindelige brugerflade og tabeller,¹⁰³ men det ser der ikke ud til at være nogen, der gør.¹⁰⁴ Rigsarkivets anbefaling burde måske gøres obligatorisk, således, at tabellers sammenhæng med sagsbehandling eller systemfunktioner bliver en del af tabelbeskrivelsen, fremfor at denne blot beskriver, hvad tabellerne indeholder. Konsekvensen ville så være, at man kommer til at stille lige så strenge krav til dokumentationen hos en privat leverandør, som man gør til den offentlige journalisering. Det er ikke nødvendigvis dårligt fra et arkivperspektiv, og slet ikke et brugerperspektiv, men det kan nok ende med at blive dyrt. Hvis man er optimist, kan man dog håbe på, at det fører til bedre dokumentationsvaner hos leverandørerne generelt.

Pointen er tofoldig: For det første vil denne type dokumentation gøre det lettere at forstå kildernes ophav, altså historikerens kildekritiske behov. For det andet forholder det sig formentlig sådan både nu og langt ud i fremtiden, at brugere af arkiveringsversioner ikke selv hverken udvikler eller vedligeholder databaser. En tilgang til fremfinding og forklaring af data, der tager udgangspunkt i den brugerflade de blev skabt i, kunne derfor gøre data lettere tilgængeligt. Særligt for systemer som Care, hvor der er flere hundrede tabeller. Omvendt bør dette heller ikke være den *eneste* tilgang, da man i så fald gør det sværere at genbruge data til f.eks. kvantitative analyser.

¹⁰³ Rigsarkivet, Vejledning til bekendtgørelse 128, s. 99. Figur 6.2, kategori 1.e.

¹⁰⁴ Foruden kontekstdokumentationen til mine tre cases i dette speciale, har jeg også gennemgået kontekstdokumentationen fra mange af arkiveringsversionerne i Hjørring Kommunearkivs samling hos KOMDA ifm. et tidligere projekt. Jeg har dog ikke haft mulighed for at se om Rigsarkivet eventuelt følger sin egen vejledning på dette punkt.

Udfordringer med dokumentationen

Arkiveringsversioner er ikke lavet således, at der kan tilføjes dokumentation, efter de er blevet godkendt. Undtagelsen er dokumentation i forbindelse med migrering. Årsagerne har formentlig at gøre med mindst to ting: For det første skal arkiveringsversionernes to backups gemmes på "kolde medier" (magnetbånd og blu-ray),¹⁰⁵ som er besværlige at opdatere, hvis man vil ændre på noget. For det andet strider det imod almindelig arkivpraksis at ændre på arkiver, når først de er afleverede, registreret og godkendt. Kontekstdokumentationen betragtes som en del af arkivet, selvom de ikke er en del af arkivets data som sådan. Det har den konsekvens, at man ikke kan tilføje ny dokumentation, hvis man under senere brug af arkivet opdager, at der mangler information, der viser sig at have enten praktisk eller kildekritisk relevans, eller hvis dokumenter, kommunen ikke kunne finde, skulle ske at dukke op.

Når systemer går fra drift til arkivering, ændres de, nogle gange endda meget, og det samme gør de, når de migreres. Denne fleksibilitet i forhold til, hvad vi kan acceptere som originale arkivalier hører til det digitale, og er meget svært at komme udenom, hvis vi er interesserede i, at arkivalierne forbliver brugbare. Ved migrering tilføjes ny dokumentation. Både i form af kontekstdokumentation og, som vi skal se senere, nye tabeller, hvis det er nødvendigt. Princippet om ikke at ændre på noget holder kun indtil arkiveringsversionen bliver ubrugelig, hvis man ikke gør det. Hvis man er interesseret i at gøre arkiveringsversioner brugbare for andre end IT-professionelle, kan det måske være en fordel, hvis man udvider denne fleksibilitet en smule mere. En måde at gøre det på kunne f.eks. være at opsamle relevant dokumentation (udover det, der allerede er med) efterhånden som man finder ud af, hvad der er behov for, og så tilføjer det ved næste migrering eller næste aflevering. Som jeg også selv tidligere har undersøgt, kan det være meget svært at gætte, hvilken information, der er relevant at have med i kontekstdokumentation både fra arkivets side og myndighedens. Arkivet har ikke brugt systemet, og har ikke nødvendigvis tid til at sætte sig ind i det ifm. arkivering, og de bliver afhængige af, at kommunen er i stand til at forstå, hvad man beder om.¹⁰⁶ Kommunen er omvendt så vant til at bruge systemet, at de ikke nødvendigvis

¹⁰⁵ Oplyst af KOMDA.

¹⁰⁶ Arkivet her er det modtagende arkiv, som i dette speciales tilfælde er Aalborg Stadsarkiv. KOMDA, der tester arkiveringsversionen og laver søgninger for arkivet, kan måske bedre vurdere, hvad der er behov for særligt mht. den tekniske dokumentation.

kan vurdere, hvad en fremtidig bruger ikke ved, og slet ikke hvad en bruger om 100 år ikke ved.¹⁰⁷ Særligt, hvis computerteknologi bliver ved med at udvikle sig i højt tempo. Her er Baks digitale proveniens igen relevant. Brugerskaren for arkiveringsversioner er i øjeblikket også så snæver, at det ikke nødvendigvis er alle potentielle brugerbehov, der bliver tilgodeset. Det er i praksis heller ikke muligt, men det er værd at tænke over om den nuværende dokumentation f.eks. er tilstrækkelig for fremtidige slægtsforskere, som kan have meget blandede forudsætninger for brug af arkivet. En mere fleksibel tilgang til dokumentation strider imod Rigsarkivets ambition om at have dokumentation, der er "skabt i forbindelse med anskaffelse og brug af systemet,"¹⁰⁸ men hvis dokumentationen til en arkiveringsversion er utilstrækkelig og samtidig ikke kan ændres, risikerer man at data over tid bliver endnu sværere tilgængeligt, end det var i forvejen, fordi for meget viden er tavs eller helt bliver glemt. Disse udfordringer kender vi godt fra papirarkivalier, hvor der ofte slet ingen dokumentation eller forklaringer er, men der er ingen grund til ikke at gøre arbejdet bedre nu, end man gjorde tidligere. Det er et særligt relevant fokuspunkt for de digitale arkivalier, fordi teknologi forældes og udvikles langt hurtigere end papir.

Struktur og indhold i arkiveringsversionerne

I det følgende vil jeg undersøge, hvad der ligger i arkiveringsversionerne af KMD Care, hvordan det ligger, og hvordan man finder det med de søgemetoder, som ASDA stiller til rådighed. Det vil jeg gøre ved hjælp af fire eksempler på indhold: Tre klienter (borgere) i systemet og en mere overordnet indgang til personale i systemet. Hvorfor sidstnævnte er mere overordnet har at gøre med søgemulighederne, som jeg vil beskrive senere. Eksemplerne lyder mere specifikt således:

1-3: Forskellige typer klienter. En ældre, en person med fysisk handicap og en person med mentalt handicap eller psykisk sygdom (kaldet "sindslidende" i systemet).

4: Ansatte i kommunens ældre- og handicapforvaltning med udgangspunkt i personalets stamdata.

¹⁰⁷ Se eventuelt Sørensen, "Indsamling af kontekstdokumentation til kommunale arkiveringsversioner" for en grundigere diskussion.

¹⁰⁸ Rigsarkivet, Vejledning til bekendtgørelse 128, s. 97.

Alle personer havde aktive sager i alle tre arkiveringsversioner af KMD Care. Eksemplerne bruges til mere specifikt at sige, hvilke typer information man bevare om de forskellige klienttyper, og om det ændrer sig over tid. Beskrivelser vil blive holdt på et mere overordnet niveau (typer af information fremfor specifikke oplysninger) for at holde personerne helt anonyme.

Udvælgelse af eksempler

Udvælgelse af eksempler er i sig selv en øvelse i at søge i arkiveringsversionerne, mere specifikt at vælge et sted at starte. Databaser har som sagt ikke et start- og slutpunkt. Her er det til at starte med uden hjælp fra arkivets SQL-søgninger. Til udvælgelse af klienter har jeg i første omgang taget udgangspunkt i tabellen CLIENT, hvori der findes stamdata for alle klienter i systemet. Det er en del af klientmodulet, der ifølge leverandøren er det centrale modul i systemet. Her startede jeg med at udvælge efter alder i arkiveringsversionen fra 2003. Stamdata i tabellen er dog ret begrænset og består mest af fremmednøgler. Derfor har jeg også taget udgangspunkt i nogle af arkivets indekssøgninger, herunder indeks over personer, der har fået tildelt hjemmehjælp, der sammen med rengøring ligger i den milde ende af, hvad kommunen kan hjælpe med.¹⁰⁹

For den ældre person valgte jeg en, der ville være over 75 år på arkiverings-tidspunktet (2003) og som minimum modtog hjemmehjælp. For at udvælge personer med andre fysiske eller mentale handicaps, end dem, der oftest kommer med alderen, gik jeg her efter yngre personer. I denne sammenhæng har jeg bestemt det til under 50 år ved første arkivering.

For fysisk handicappede er det relativt ligetil at finde personer ud fra bevilligede hjælpemidler såsom kørestole, rollatorer, mv. Det betyder dog ikke, at man får alle med fysiske handicaps med som muligheder, da ikke alle fysiske handicap har at gøre med gangbesvær. Det kan også være blinde, døve, personer med astma eller personer med progressive uhelbredelige sygdomme som parkinson eller osteoporose (knogleskørhed), hvis tilstand endnu ikke er nået et stadie, der kræver f.eks. kørestol. Med dette søgekriterie kan man dog sortere alle fra, som *ikke* er fysisk handicappede, hvilket var den mest oplagte metode, når alle diagnoser, henvisninger og bevillinger

¹⁰⁹ I hvert fald ud fra de moduler til forskellige former for hjælp, der findes i Care.

til alle typer klienter står i de samme tabeller. Det er heller ikke alle med fysiske handicaps, der har medfødte eller på anden vis permanente handicaps. Her kan man tage indekset over diagnoser i brug, hvor det kan være lettere at gætte. Derudover er denne oplysning som regel nævnt i første journal eller henvisning på de personer, hvis data, jeg har kigget på.

Indekssøgningen med diagnoser kan også bruges til at udvælge sidste eksempel. Heri står alle diagnoser, som de ansatte har brugt, og i den tidligste udgave af KMD Care var der lidt over 2000 poster, hvorfra jeg valgte en tilfældig person med en psykisk diagnose.

Herefter bruges ASDAs søgepakker til at finde de oplysninger, der måtte være om personerne ud fra CPR, som er den mest almindelige måde at fremsøge personer på i både drift- og arkiveringsversioner. Ifølge KOMDA skal søgninger også gerne afspejle, hvad sagsbehandlere ville forvente at finde i systemet.¹¹⁰

Det er ikke muligt at vælge helt tilfældigt, når man skal gøre det manuelt, men med "tilfældigt" mener jeg snarere, at jeg ikke har ledt efter et specifikt resultat udover de kriterier, jeg har opstillet ovenfor. Da udgangspunktet for undersøgelsen er, at det er den første, har jeg heller ikke haft noget specifikt at lede efter. Jeg har derfor ikke valgt ud fra f.eks. antal poster eller tekniske fejl i en persons sag. Jeg har valgt ud fra en formodning om, at forskellige klienter har haft at gøre med forskellige dele af systemet, hvorfor deres arkiver ser forskellige ud, men det har jeg først kunne be- eller afkræfte efter udvælgelsen.

Det kan af lignende årsager være svært at vurdere om mine udvalgte klienter er repræsentative for klienter i systemet generelt mht., hvad der er bevaret om dem. Som vi senere skal se, er typen af informationer dog relativt ens for de forskellige klienter, og det virker derfor sandsynligt, at de største forskelle på, hvad der er blevet gemt om forskellige klientgrupper findes indenfor hver enkelt type information (f.eks. i journaler, henvisninger og tildelte hjælpemidler). Af hensyn til anonymisering er det dog ikke noget, jeg kan komme nærmere ind på.

Alle eksempler på klienter er valgt i den tidligste udgave af KMD Care. Herefter har jeg søgt på dem i den relevante søgepakke for at se om deres sag stadig var aktiv i den seneste udgave.

¹¹⁰ Falk og Hansen, "Digitale arkivalier: Fra sorte bokse til big data," s. 80.

Eksempler på klienter

Almindelig ældre

Personen her optræder i alle tre arkiveringsversioner og første gang i 2001. Personen var på dette tidspunkt i 70'erne. Jeg kalder vedkommende en "almindelig" ældre, da der ikke er tale om en person med medfødt fysisk eller mentalt handicap eller en person med demens. Af vedkommendes journal kan man også læse, at der er tale om hjælp til typiske problemer, der opstår med alderen, inklusiv mere seriøse problemer, hvis risici også stiger med alderen. Ved søgning med ASDAs søgepakke dukker der lidt over 2000 poster op. Poster skal her forstås som dele af personens sag. Mens poster i tabellerne er rækker, kan det her sagtens være data hentet fra flere tabeller, men ASDAs søgninger viser ikke umiddelbart, hvilke tabeller det er fra. Her skal man ind og kigge i den tilhørende SQL-sætning. I personens sag findes følgende typer oplysninger:

For KMD Care 2002:

- Henvisning til hjælpemidler
- Henvisning til rengøring
- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Journaler
- Tildelte hjælpemidler
- Notitser

For KMD Care 2006:

- Henvisning til hjælpemidler
- Journaler
- Henvisning til rengøring
- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Tildelte hjælpemidler
- Notitser

For KMD Care 2020:

- Tilknytning til plejebolig

- Visiteret hjælp
- Visitation af hjælpemidler
- Henvendelser
- Visitation af bolig
- Klientjournal
- Henvisninger
- Funktionsvurdering

Her har vi at gøre med en person, der har haft at gøre med stort set alle moduler, som de ældre klienter kan have med at gøre uanset om de har kurs mod plejehjem eller ej. Dvs. dem der vedrører hjælp, pleje og rengøring i eget hjem samt hjælpemidler og til sidst overflytning til plejebolig (det er det samme som et plejehjem). Ud af listen her er det tydeligt, at den største ændring i databasens struktur skete mellem 2006-udgaven og 2020-udgaven. Det er også her, at der er det største spring i tid, og her hvor systemet var blevet opkøbt af KMD.

Af hensyn til anonymisering har jeg undladt at nævne præcise antal af hver enkelt type post. Der er dog en tydelig udvikling over tid, som kan skyldes både ændringer i systemet, ændringer i sagsbehandlingen, ændringer i kommunens praksis for hjemmehjælp/pleje/mv. og ændringer i personens situation. Sidstnævnte vil for en ældre person ikke være overraskende, men det tyder også på, at nogle af de ovenstående gør sig gældende. Meget vil kunne læses ud af personens journaler/klientjournal, som er der, hvor klart det meste tekst er blevet indtastet af systemets brugere.

Langt de fleste poster for denne person findes i KMD Care 2020 (2023). Der er færre i den første udgave, hvor de fleste poster er journaler. Journaler og hjælpemidler fra tidligere går igen pga. den måde der arkiveres på (øjebliksbillede), så det er forventeligt, at mængden af data om en person stiger over flere arkiveringsversioner. Men over 1500 af posterne er henvisninger til forskellige former for hjælp. Alle henvisninger ligger her samlet, præcis som de gør i den tabel, de kommer fra, men det gør dem uoverskuelige for en bruger, der måske har brug for oplysninger om henvisninger til noget specifikt. I langt de fleste henvisninger er det også meget sparsomt med information om, hvad henvisningen handler om. Henvisninger ordnet boolsk, hvor der er svaret JA på den type hjælp, der er henvist til og NEJ på de øvrige muligheder. Der

står også en dato, og hvem der har henvist. Der står til gengæld ikke, hvorfor henvisningen er lavet. Der refereres heller ikke til en journal, hvor denne information måske kunne findes.

I den ældste udgave af systemet ligger søgninger på henvisninger, som det kan ses af listen ovenfor, hver for sig under hver enkelt type hjælp. Her er også et felt til "årsag", som ikke findes i søgeresultaterne fra den nyeste version. I tabellerne kommer data dog stadig fra en tabel, der er sat op på samme måde som den nyeste udgave. Årsagerne kommer til gengæld fra hver sin tabel, som findes for hver sin type henvisning (hjemmehjælp, hjemmepleje, hjælpemidler, rengøring). Dette findes tilsyneladende ikke i den nyeste udgave. Her er således et eksempel på, hvor stor en forskel det kan gøre, at arkivet vurderer, hvordan data præsenteres for brugere, og hvad de i så fald vurderer til at være den rette præsentation.

Jeg har indtil videre fokuseret meget på henvisninger, fordi omkring $\frac{3}{4}$ af alle personens poster er henvisninger. Mens journaler og tildelte hjælpemidler ser ud til at være dateret korrekt, ser en meget stor del af henvisningerne¹¹¹ ud til at være dateret til samme dato tidligt i 2001. Da der er kommet langt over 1000 henvisninger til med samme dato over tid, også efter 2001, tyder det på en systemfejl, som hverken er blevet noteret af leverandøren i forbindelse med arkivering. Det kan også være, at systemet blot kom til at fungerer anderledes og at de mange henvisninger ikke er en fejl. I begge tilfælde ville information om det være relevant at have med i kontekstdokumentationen så senere brugere kan forstå, hvad der foregår. Hvis fænomenet ikke er enkeltstående, betyder det også, at et højt antal poster ikke nødvendigvis betyder, at man har at gøre med en omfattende eller kompliceret sag. Uden kontekst kan vi heller ikke vide, om det rent faktisk er reelle nye henvisninger med forkert dato. Da datoen passer med de arkiverede datoer fra 2001 virker det dog usandsynligt. Konsekvensen af antallet af henvisninger og de meget sparsomme oplysninger deri er dog, at de bliver svære at bruge til noget. De mange henvisninger, der tilsyneladende er uden reelt indhold, skaber derudover en hel del støj i søgningen, der gør data om henvisninger besværlig at håndtere.

Den midterste version med data op til 2006 indeholder færre poster om personen end de andre to. Data er desuden overlappende med den tidligste udgave, mens data, der burde eksistere i perioden mellem de to første udgaver (2002-2006) først

¹¹¹ Jeg havde desværre ikke tid til at gennemgå dem alle.

findes i den nyeste version. At der er tale om arkiveringsversionerne som en "til enhver tid opdateret system"¹¹² holder derfor ikke. Helt generelt sår det tvivl om indholdet i hele KMD Care 2006 - medmindre fejlen er opstået som følge af en forkert SQL-sætning, hvis journalerne ligger et andet sted i arkiveringsversionen. Det er dog ikke lykkedes mig at finde dem manuelt.

Det er derfor en manuel undersøgelse at finde ud af, hvor længe personen var tilknyttet systemet. Poster er ordnet kronologisk under hver kategori (i ovenfor nævnte rækkefølge) i hver enkelt arkiveringsversion, så det er en smule mere tidskrævende end blot at scrolle til bunden. Et godt bud på et relevant sted at kigge må dog være klientjournalen, og her er sidste registrering foretaget i 2010. Alene ud fra søgeresultaterne ser det ikke ud til, at systemet har haft en funktion til at lukke sager ved dødsfald. Det blev dog, som nævnt, registreret i stamdata. Der er heller ikke registreret om personen eventuelt skulle være flyttet fra kommunen, hvilket kunne være en årsag til, at data ophører uden en slutning. Sådanne oplysninger skal findes andetsteds. Ved manuel søgning i klient-tabellen med stamdata i KMD Care 2020 (2023) dukker personen imidlertid ikke op, så i dette tilfælde har det kun mindre betydning, at det ikke er med i søgeresultaterne. Hvis man er interesseret i at undersøge dødsdatoer eller andre afslutninger i systemet kan det måske i stedet give mening, at lede i adviser, altså sagsbehandlernes interne kommunikation.

Fysisk handicappet person

Denne person er født med et fysisk handicap og er kørestolsbruger. Personen optræder i alle tre arkiveringsversioner, hvor første gang var i 1997. Her var personen relativt ung, og tæller derfor ikke som en ældre i mine udvælgelseskriterier. Hvis man vil have hele denne persons historie, ville det være oplagt også at søge i papirarkivalierne. Ved søgning i ASDA dukker der ca. 4700 poster op med følgende typer oplysninger:

For KMD Care 2002:

- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Journaler
- Tildelte hjælpemidler

¹¹² KMD Care 2006 (2009), kontekstdokument nr. 2: Rigsarkivet, Standardafleveringsmodel, s. 3.

- Diagnoser
- Notitser

For KMD Care 2006:

- Diagnoser
- Journaler
- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Henvisning til bopæl
- Tildelte hjælpemidler
- Notitser

For KMD Care 2020:

- Tilknytning til plejebolig
- Visiteret hjælp
- Visitation af hjælpemidler
- Henvendelser
- Visitation af bolig
- Klientjournal
- Målinger
- Henvisninger
- Funktionsvurdering
- Visiterede behandlingsydelser
- Service modtaget
- Dokumenter

Over 3000 poster er henvisninger fra KMD Care 2020 (2023). Ligesom med forrige eksempel er de fleste af dem gamle henvisninger, der ser ud til at være blevet duplikeret. Problemet med henvisninger er således ikke enkeltstående. Henvisningsdatoerne er alle fra 1997, men for en del af dem er der også en dato for en deadline,¹¹³ som varierer med datoer i en blandet orden (ikke kronologisk) fra hele perioden. Det er inklusiv i årene, der gerne skulle være dækket af KMD Care 2006 (2009), hvor de til gengæld ikke er at finde. Det er journalerne heller ikke. Ligesom i ovenstående er

¹¹³ Det er ikke forklaret, hvad der menes med det. Hvis der er tale om legitime henvisninger i stedet for tekniske fejl, kan det måske være den reelle henvisningsdato.

der dog ingen uoverenstemmelser med tildelte hjælpemidler.¹¹⁴ Manglende data fra 2006-udgaven er således heller ikke enkeltstående. Som et eksempel på, hvorfor både måden at vise henvisninger på og det overvældende antal i en blandet rækkefølge er meget upraktisk, kan personens journal afsløre, at vedkommende modtog hjælp til rengøring hver uge i perioden, som kun er dækket af 2020-udgaven. Det kunne tage mange timer at scrolle alle henvisninger igennem for at finde dokumentation for, om der er henvist til rengøringshjælp (hjælpen kunne også være foretaget af en hjemmehjælper, hvilket er fra et helt andet modul). I ASDA kan alle poster gemmes som enten TXT eller PDF. Afhængig af om PDF'en er billede eller tekst kan begge dele bruges til nærmere søgning, når der er tale om flere tusind poster. Det kræver dog, at man bruger ASDA på en computer med Windows for at kunne hente noget ud derfra.¹¹⁵

Ser man bort fra henvisningerne har denne person stadig omkring dobbelt så mange poster som første eksempel. Det hænger sammen med, at de har været i systemet i længere tid, men også, at de har haft mere omfattende behov for hjælp og særligt hjælpemidler.

I 2020-udgaven af KMD Care ligger de fleste poster for denne person, med undtagelse af dokumenter, ikke i kronologisk rækkefølge. Ikke engang journalerne. Det kan derfor være svært at se, hvornår personen sidst optræder i systemet. Det seneste dokument er fra 2019, hvor der blev bevilget et nyt hjælpemiddel.

Mentalt handicappet person/"Sindslidende"

Denne person er udvalgt via indekset over diagnoser og stod til at have en psykisk sygdom, der typisk opstår relativt tidligt i livet (sammenlignet med f.eks. demens). Personen optræder i alle tre arkiveringsversioner og første gang i 1999. Her var vedkommende midaldrende. Sidste journal, som for denne person heldigvis står i kronologisk rækkefølge, er fra 2008. Ved søgning i ASDAs søgepakke dukker der ca. 1800 poster op med følgende typer oplysninger:

For KMD Care 2002:

- Henvisning til hjælpemidler

¹¹⁴ Det tyder på, at fejlen måske er modulbestemt, og at hjælpemiddelmodulet er gået fri.

¹¹⁵ Min tildelte computer til ASDA på Aalborg Stadsarkiv var med Linux (Mint), så jeg kunne desværre ikke teste denne søgemetode.

- Henvisning til rengøring
- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Journaler
- Tildelte hjælpemidler
- Diagnoser
- Notitser

For KMD Care 2006:

- Diagnoser
- Henvisning til hjælpemidler
- Journaler
- Henvisning til rengøring
- Henvisning til hjemmehjælp
- Henvisning til hjemmepleje
- Tildelte hjælpemidler

For KMD Care 2020:

- Tilknytning til plejebolig
- Visitation af hjælpemidler
- Visitation af bolig
- Klientjournal
- Henvisninger
- Funktionsvurdering
- Service modtaget

Begge udfordringer fra de forrige eksempler går igen for denne person, og ca. 1600 poster er henvisninger, mens data fra KMD Care 2006 (2009) er mere eller mindre fraværende. Læser man personens journal kommer det også frem, at vedkommende slet ikke lider af den sygdom, de var registreret med, men et helt andet problem, der dog stadig er af psykiatrisk karakter. Her har vi således at gøre med et eksempel, hvor søgemetoden via diagnose ikke var så brugbar, som man skulle tro. Personens egentlige diagnose findes ikke i tabellen med diagnoser, som de ansatte havde at vælge imellem i KMD Care 2002 (2003) eller KMD Care 2006 (2009), så det virker sandsynligt, at man har valgt det næst-bedste, som man først fik øje på (sygdommen, som personen blev registreret med, står relativt højt i listen). Systemets brugergrænseflade

er meget dårligt dokumenteret, men måden diagnoser er opstillet på, og måden de er blevet brugt, tyder på, at der har været en drop-downmenu uden mulighed for at vælge en brugerdefineret diagnose. Hvis det var tilfældet, er denne praksis ikke dokumenteret i kontekstdokumentationen ("generelle informationer" på dette tidspunkt). Her er måske et eksempel på kommunespecifik praksis, som man ikke har dokumenteret i første omgang, men som ville være hjælpsomt at tilføje. I praksis vil man derfor ikke kunne søge efter personer med denne diagnose, næsten uanset hvor gode søgemulighederne er. Alene ud fra søgningerne ser det ud til, at man helt havde droppet at registrere diagnoser i KMD Care 2020 (2023). Både tabellen med diagnoser (den blev aldrig opdateret med nye diagnoser) og listen over klienter med deres diagnoser findes dog endnu - samtlige datoer deri er bare den samme (6. Juni 2006 klokken 15:42). Det samme er tilfældet i den tidligere aflevering. Går man helt tilbage til den første aflevering var der slet ikke dato på, så den senere tilføjede dato kan være en dato for enten en systemopdatering eller tilføjet i forbindelse med afleveringen af KMD Care 2006 (2009). Antallet af klienter i listen er næsten identisk med listen fra KMD Care 2006 (2009), og det virker derfor sandsynligt, at man ikke længere bruger diagnoser på samme måde som tidligere. Da vejledningen til systemet ikke er kommet med i kontekstdokumentationen til KMD Care 2020 (2023), er det ikke muligt at tjekke, om det er en officielt udfasningen af funktionen (så ville den ikke optræde i vejledningen) eller blot Aalborg Kommune, der ikke længere bruger den.

Til fremsøgning og udvælgelse af alle tre eksempler er jeg, som sagt, startet med KMD Care 2002 (2003), og været heldig at finde personer, der havde aktive sager i alle tre arkiveringsversioner. Søgemulighederne for mentalt handicappede er dog bedre i både KMD Care 2006 (2009) og 2020 (2023), hvor der var blevet muligt at henvise til "modul for sindslidende" og "modul for udviklingshæmmede". Kontekstdokumentationen for KMD Care 2020 (2023) indeholder et dokument med organisationsændringer i Aalborg Kommunes Ældre- og Handicapforvaltning,¹¹⁶ der kan afsløre, at de bedre søgemuligheder formentlig hænger sammen med, at man i 2007 havde fået opdelt forvaltningen i flere specialiserede afdelinger, hvoraf tre var "autisme og sindslidelse", "udviklingshæmmede og døve" og "fysisk handicappede og særlige tilbud". Med bedre søgemuligheder, mener jeg her ikke kun, hvad arkivet har valgt at gøre,

¹¹⁶ KMD Care 2020 (2023), kontekstdokument nr. 2, "Beskrivelse af organisation og organisationsændringer i driftsperioden." fra Aalborg Kommune.

men også hvad arkiveringsversionens indhold og struktur giver mulighed for. I dette konkrete tilfælde afhænger det dog også af, om de nye moduler var kommet til inden, personen udgik af systemet. Hvis der findes en henvisning til en af disse afdelinger for personen i eksemplet her, ligger den dog godt gemt blandt de hundredevis af auto-generede henvisninger i søgningen.

Ansatte i forvaltningen

Dette eksempel er med for at undersøge, hvordan man kan finde oplysninger om personer i systemet, som ikke er systemets "subjekter" om man vil. Eller med andre ord, hvordan man kan finde oplysninger om personer uden at have søgninger sat op til det. Ansatte i forvaltningen kan man ikke søge på via søgepakken. Mens søgepakken er sat op til at finde alt relevant data og dokumenter om klienter, er der ikke samme mulighed for at finde alt (registreret) arbejde, som en ansat har udført. Den første SQL-søgning på ansatte kommer først i KMD Care 2020 (2023), og her er der kun én. Tilgangen til fremfindning er derfor en del anderledes og mere manuel end for klienter. I første omgang skal man håbe, at tabellerne med ansatte ikke er blevet kasserede. Som en konsekvens af, at man udelukkende kasserer systemtabeller, er de bevarede i alle tre udgaver af KMD Care.

I KMD Care 2002 (2003) hedder personaletabellen PERSON_L. Informationer om ansatte er i KMD Care 2002 (2003) spredt ud på forskellige tabeller med basis-tabellen i centrum.

Basis-tabellens kolonner indeholder følgende oplysninger:

- Personnelnummer
- Stillingsnummer (henviser til separat tabel med liste over stillingstyper)
- Postnummer
- CPR-nummer
- Et ekstra CPR-nummer (et-cifret). Ifølge datamodellens feltbeskrivelse bruges det til at identificere, hvor personen arbejder.
- Fornavn
- Efternavn
- Telefonnummer

- Adresse
- By
- Startdato for ansættelse
- Slutdato for ansættelse
- Om personen er aktiv
- Om personen faktisk er en "rute"
- Om personens data automatisk skal overføres til "Kommunedatas scheduling system" (KMD Vagtplan, som ikke er bevaret)
- Om personens data er blevet overført eller ej
- Personens initialer
- Om personen har hemmeligt nummer
- Om personen har hemmelig adresse
- Om personen er "contractor" (leverandør?)
- Arbejdstelefonnummer
- Opsigelsesårsag via nummer (fremmednøgle) fra en anden tabel (pers_disch_reason)
- "Substitute" initialer. Vikar?

Derudover er der en del tomme kolonner. Helt tomme kolonner betyder formentlig, at det er en funktion, kommunen ikke har brugt. De tomme kolonner kan indeholde oplysninger om:

- Jobtitel (occupation)
- Opsigelsesårsag
- Noter
- Transportform til og fra arbejde
- Dato for sidste opdatering af arbejdsvilkår
- Om personen ønsker overarbejde
- Vagt (dag, aften, nat)
- Om personen kan repræsentere sine kolleger (tillidsrepræsentant?)
- Om personen får tildelt en uniform
- Dato for første tildelte uniform
- Mødetider
- Kontortider

Oplysningerne i personaletabellen PERSONNEL i KMD Care 2006 (2009) er mere eller mindre identiske med den forrige version. Der er tilføjet 11 kolonner (der var 37 i den første) til ny information, som generelt ikke er blevet brugt. Det samme er tilfældet med KMD Care 2020 (2023). De nye kolonner inkluderer kommunenummer, en ekstra kolonne med personens navn, email, om personen kun bruger ikke-motoriseret transport (cykel) og gennemsnitshastighed på cykel (20 som standard). Mit eksempel havde ikke fået en email på dette tidspunkt, men alle ansatte har fået kommunenummeret tilføjet (Aalborg er nr. 851).

Personaletabellen i KMD Care 2020 (2023), PERSONNEL er igen stort set de samme som i den første udgave, men med flere muligheder for at tilføje ekstra oplysninger. Nu er der hele 68 kolonner. Langt de fleste nye kolonner er igen tomme. Nogle kolonner indeholder også data, som ikke kan bruges til ret meget. F.eks. er der en kolonne (MANUAL_COOR) til at identificere, hvorfra koordinater for ansattes adresser kommer, som er udfyldt med enkelte bogstaver, men kolonnerne til disse koordinater er tomme. En anden kolonne er tekster til disse koordinatkilder, som består af fejlmeddelelser. Det ser ud til, at blot fordi der bliver mulighed for at gemme flere informationer, betyder det ikke, at det rent faktisk bliver gjort. Sidder man med en detaljeret beskrivelse af et systems funktioner, er det derfor ikke ensbetydende en detaljeret beskrivelse af, hvad der ligger deri.

Nogle af de tilføjede kolonner har slet ikke med egentlige ansatte at gøre. Som den ovenstående liste viser, er personaletabellen nemlig ikke kun blevet brugt til personale, men også til ikke nærmere forklarede "ruter". Det virker sandsynligt, at der kan være tale om kørselsruter for f.eks. hjemmepleje eller udbringning af mad. For "ruter" er der i den nyeste udgave f.eks. tilføjet dage og tidspunkter, hvor en "rute" er tilgængelig. Fra et brug-af-data-perspektiv giver det ingen mening at have disse meget forskellige typer oplysninger liggende i den samme tabel, men man må gå ud fra, at det har været praktisk fra et brug-af-systemet-perspektiv at gøre det på denne måde i stedet for at give "ruter" sin egen tabel. I alle tre versioner findes også en robot, som var "ansat" mellem ca. 2002 og 2018.¹¹⁷ Man har fra Rigsarkivets side fokuseret lidt på at dokumentere kommunespecifik brug af systemer (ikke at det nødvendigvis har

¹¹⁷ I første udgave havde den ikke nogen ansættelsesdato, så den har man senere estimeret til sidst i 2003, hvilket ikke er korrekt, da den optræder i KMD Care 2002 (2003). Det interessante er dog, at man overhovedet har indtastet den slags data for en robot. Robottens rolle er ikke forklaret, men da den har eksisteret længere end det har været almindeligt at bruge fysiske robotter i ældreplejen, er der formentlig tale om et stykke ren software.

ført til bedre dokumentation),¹¹⁸ men dokumentation af, at leverandøren har sat et system op til at blive brugt på en måde, der ikke umiddelbart stemmer overens med f.eks. tabellers primære formål (menneskeligt personale i dette eksempel) er ikke noget, man har tænkt over. Det kunne f.eks. blive en udfordring for en bruger at gætte, hvor oplysninger om "ruter" ligger, hvis det kun var dem, man var interesseret i at undersøge, når søgning skal foregå manuelt.

Det er i den forbindelse også værd at bemærke at personaletabellerne har tre forskellige navne. De har til gengæld det samme navn ("Personnel") i dokumentationen til alle tre arkiveringsversioner. Hvis det er tilfældet for denne tabel, er det formentlig også for andre, hvor man har valgt eller været nødt til at forkorte navne på den ene eller den anden måde. Det er ikke oplyst, at der kan være forskel på tabelnavne i dokumentationen og de reelle tabelnavne i arkiveringsversionerne. Det har den konsekvens, at dokumentationen er sværere at bruge, hvis man skal lede efter oplysninger manuelt - særligt i systemer med flere hundrede tabeller.

Har man først fundet en ansat, man er interesseret i at undersøge, burde det være relativt ligetil, at finde alt, hvad en ansat er registreret til at have foretaget sig, ved at koble relaterede tabeller på basistabellen vha. ASDAs "vis relationer"-funktion. På denne måde kan man (i teorien) finde alle dokumenter, som en ansat står som skaber af, eller alle journaler, som en ansat har skrevet, såfremt man ved, hvilken tabel det står i. Hvis man specifikt vil undersøge ansatte, der har skrevet f.eks. journaler, kan man åbne tabellen med journaler CLIENT_JOURNAL og søge i kolonnen med de ansattes initialer/nummer (INI, fremmednøgle til PERSONNEL).

I KMD Care 2020 (2023) finder man den første oprettede søgning på ansatte. Den kan finde en liste over ansatte (kun med deres personalenummer, ikke med navn eller CPR) og den afdeling, de er koblet til samt en masse ikke nærmere forklarede datoer. Det er ofte nogle få eller kun én dag ad gangen, og der er derfor over 100.000 poster i listen, selvom der kun står omkring 40.000 ansatte i PERSONNEL, hvoraf det ikke

¹¹⁸ Det fremgår af en præsentation, Rigsarkivet lavede for ODA i 2019. Derudover var Rigsarkivet (Statens Arkiver) med til at udvikle den tidligere nævnte standardafleveringsmodel til Care, og det er derfor sandsynligt, at de også har været blandet ind i beslutningen om at lave dokumenter til kommunespecifikke forskelle og kommunespecifikke vejledninger.

er alle, der er aktive,¹¹⁹ og mange der er “ruter” (ikke personer). Der kan måske være tale om dage, hvor en ansat har arbejdet i en bestemt afdeling.

Andre typer oplysninger

Denne undersøgelse har kun været det første spadestik i et meget dybt hul. Udover klienter og personale indeholder systemet også mange oplysninger, der bruges til personsager, men som ikke i sig selv har noget med personer at gøre. Et eksempel er data om hjælpemidler, der kan tildeles personer. Det har et selvstændigt modul i alle tre arkiveringsversioner, som er relativt omfattende. I denne sammenhæng vil det sige, at der er mange tabeller. Det er et eksempel på at man gemmer en del ekstraoplysninger, fordi relationelle databaser fungerer, som de gør. Oplysninger om hjælpemidler til borgere i deres respektive tabeller giver således kun mening, hvis man gemmer alle andre tabeller med oplysninger om hjælpemidler. En nærmere undersøgelse vil kunne gøre det klart om der er mange, fordi der er meget data, eller fordi det skal understøtte bestemte funktioner i systemet. Umiddelbart virker sidstnævnte mulighed dog mest sandsynlig.

I hjælpemiddelmodul i KMD Care 2006 (2009) findes f.eks. tabellen SUBJINDEX, der er en liste over “subjectindex for technical aids”¹²⁰ eller på dansk: Emneliste for hjælpemidler. I realiteten en liste over hjælpemidler. Den har en relation til en tabel, hvor hjælpemidlernes overordnede kategori står. Der er alt fra “etiket” til “katamaran”. Indenfor klassificering af hjælpemidler er der fire forbundne tabeller med forskellige informationer med relationer til den samlede liste over hjælpemidler i Danmark og en liste med leverandører. Men man skal relativt langt ud for at finde en forbindelse til en klient.

SUBJINDEX kan bruges som eksempel på, om man via dokumentationen kan finde ud af, hvor data i tabeller kommer fra, hvilken sammenhæng det er skabt i, og hvilken det er brugt i. Ud fra vejledningens afsnit om hjælpemiddelmodul ser det ikke ud til at være tilfældet. Ifølge brugervejledningen¹²¹ kunne man i systemet tilgå en liste over tilgængelige hjælpemidler, men det er uklart hvilken tabel, der henvises til. Der

¹¹⁹ Aalborg Kommune har i dette system ikke haft for vane at slette ansattes data selv længe efter opsigelse. Derfor står der f.eks. relativt mange ansatte, hvis ansættelse sluttede tilbage i 00'erne, og hvis slutdato derfor også er at finde i begge de tidligere afleveringer.

¹²⁰ Tabelbeskrivelse.

¹²¹ KMD Care 2006 (2009). Rambøll Informatik, Kontekstdokument nr. 10, Brugervejledning, s. 223.

står "lagervarer", så her er der muligvis tale om tabellen TECHSTOCK, som henviser til listen over alle hjælpemidler på lager. Der er ifølge vejledningen også en liste over hjælpemidler, som er udlånt til klienter, men jeg kan ikke finde en tilsvarende tabel med denne information. Der er også en liste over alle hjælpemidler i *kommunens katalog*. Hvilken tabel står det i? Det kan måske godt være SUBJINDEX, men det er ikke klart.

Opsamling

For klienter ligner indholdet meget hinanden. Det er mere eller mindre de samme typer oplysninger, der gemmes om alle tre personer. Systemet har ændret sig en del over tid, hvilket bl.a. kan ses i det stigende antal tabeller, og tilføjelsen af dokumenter mellem første og anden udgave. Samtidig er det dog også tydeligt, både i søgeresultater og, når man manuelt kigger i tabellerne, at man har forsøgt at holde systemets vigtigste strukturer nogenlunde stabile. Både klienter og ansattes stamdata er der kun blevet ændret ganske lidt på. Det er dog ikke lykkedes hele vejen. Det er tydeligst ved KMD Care 2006 (2009), hvor så meget data har ændret placering i databasen, at selv ikke KOMDA fandt det til deres søgning. Det er let at antage, at det slet ikke er der, men det kræver en grundigere undersøgelse at kunne be- eller afkræfte det. I andre tilfælde er der kun ændret en smule, men det var lige præcis nok, til at påvirke frem-søgning af oplysninger, f.eks. henvisningerne, der blev så mange, at de er svære at håndtere.

Det ville måske være passende, hvis leverandøren havde dokumentation for de største ændringer over tid, som kan være brugbart, hvis man skal finde den samme slags information igennem flere arkiveringsversioner. F.eks. henvisningsårsager, der forsvandt fra den nyeste udgaves søgepakke, men som måske stadig ligger i arkiveringsversionen et sted.

En anden ting, der ville være brugbart at vide, men som ikke er beskrevet, er hvilke data, der er indtastet eller valgt af et menneske og hvilke der er genereret automatisk. F.eks. datoer på journaler, ansættelser, opsigelser, henvisninger, osv. I mange tilfælde kan man enten tydeligt se det i dataene, mens det i andre tilfælde ikke er tydeligt. Mange af de tusindvis af henvisning står også til at være genereret automatisk ("henvist af" står til "automatisk"). Fremmednøgler er som regel også noget systemet tager sig af uden særlig meget indblanding fra brugerne, da denne data kun

fungerer som systemets relationer. Meget data, der fungerer som metadata er formentlig også påført automatisk. F.eks. forfatter på journaler, der står som en ansats initialer. Her ville det nok være mere praktisk med et navn, men det indikerer, at denne oplysning formentlig er bestemt af, hvem der er logget ind fremfor, at man f.eks. har indtastet et navn manuelt. Datoer kan til gengæld være sværere at gætte, og hvis ens mål er at undersøge et forløb over tid, kan det have relativt stor betydning. Er journal-datoer f.eks. valgt manuelt eller er det systemets eller computerens dato? Var det muligt at vælge mellem disse to muligheder? De digitale systemer, og særligt i et, der har en app til brug udenfor et kontor som det er tilfældet med Care, gør det muligt at skabe data meget tæt på begivenhederne. Som jeg vil komme ind på med et eksempel senere, betyder det ikke, at man rent faktisk gør det. Andre datoer, som i diagnoselisten nævnt ovenfor, er lettere at identificere som computergenererede. Der er således rigtig meget data i denne arkiveringsversion, som kun indirekte er skabt af mennesker. Meget af det har brugerne, altså de ansatte i forvaltningen, måske ikke engang haft indflydelse på, og at kunne forstå intentionen med dem, kræver således, at man forstår arkiveringsversionernes tekniske proveniens.

Tilgængelighed

I det følgende vil jeg først gennemgå de SQL-søgninger, KOMDA har lavet til Aalborg Stadsarkiv og derefter diskutere tilgængeligheden af arkiveringsversioner i bredere forstand. Herefter vil jeg diskutere, hvilke andre oplagte vinkler på emnet, som dette projekt lægger op til.

SQL-søgninger i ASDA

I løbet af undersøgelsen i de forrige afsnit er det blevet tydeligt, at arkivets søgemuligheder har stor indflydelse på, hvad man kan finde i arkiveringsversioner. Man kan formentlig altid finde, hvad der ligger deri, hvis man bruger nok tid på det, og har færdighederne til at forstå en database. Det er det bare langt fra alle brugere af arkivet, der har. Med hundredvis af tabeller er det meget let at overse, hvor relevant information ligger (læs: i hvilken tabel). Det kan man naturligvis slå op i tabelbeskrivelsen. I de forrige afsnit, og særligt i eksemplet om at finde oplysninger om ansatte, har jeg diskuteret manuel søgning en del. Det er dog lettere for en bruger, der ikke er vant til

at have med databaser at gøre, at bruge de præ-lavede SQL-søgninger, som KOMDA har oprettet. Det er, som tidligere nævnt, heller ikke muligt for en bruger at oprette sine egne SQL-søgninger. Derfor vil jeg også gennemgå de søgninger, som i skrivende stund er oprettet for arkiveringsversionerne af KMD Care. Med undtagelse af de SQL-søgninger, som leverandøren kommer med, kan søgninger oprettes, slettes og redigeres efter behov. Derfor afspejler de søgninger, der fremgår af dette afsnit først og fremmest, hvad der er behov for blandt arkivets nuværende brugere. I dette tilfælde vil det sige Aalborg Stadsarkiv, der håndterer forespørgsler fra borgere og forvaltningen.

Søgninger til de tre arkiveringsversioner af KMD Care, inklusive titler og rækkefølge, ser således ud i fanen "Søgninger" i ASDA (med eventuelle beskrivelser i parentes):

KMD Care 2002 (2003)

Kriteriesøgninger:

- Helbredsoplysninger

SQL-søgninger

- KOMDA_INDEX_PERSON_HJÆLPEMIDLER
- KOMDA_INDEX_PERSON_JOURNALER
- KOMDA_INDEX_PERSON_NOTITS
- KOMDA_INDEX_PERSON_DIAGNOSE
- KOMDA_INDEX_PERSON_PROBLEM-HANDLING
- KOMDA_INDEX_PERSON_HJEMMEHJÆLP
- KOMDA_INDEX_PERSON_HJEMMEPLEJE
- KOMDA_INDEX_PERSON_HENVIS-HJÆLP
- KOMDA_INDEX_PERSON_RENGØRING
- KOMDA_INDEX_PERSON_BOPÆL
- KOMDA_INDEX_PERSON_FUNKTION

Leverandørsøgninger

- View1 ("Klient telefonliste (33).")
- View2 ("Journal (7). Denne rapport henter oplysninger om borgerens journal.")
- View3 ("Mellemkommunal refusion (95).")

- View4 ("Nødkaldsliste (53). Select'en søger et øjebliksbillede frem for borgere som har udfyldt nødkaldsnummer, nøglenummer, eller står markeret som de har nødkald. Bruger ingen parametre.")
- View5 ("Køreliste - del A - Personaleoplysninger (6). Select delt op i to dele, een som henter personaleoplysninger, og een som henter clientoplysninger.")
- View6 ("Køreliste - del B - Personaleoplysninger (6). Select delt op i to dele, een som henter personaleoplysninger, og een som henter clientoplysninger.")

Søgepakkesøgninger

- Omsorgssager (DIP.AAS.2)
 - Notitser
 - Tildelte hjælpemidler
 - Diagnoser
 - Problem og handling
 - Journaler
 - Henvi sning til hjemmehjælp
 - Henvi sning til hjemmepleje
 - Henvi sning til hjælpemidler
 - Henvi sning til rengøring
 - Henvi sning til bopæl
 - Funktionsvurdering

KMD Care 2006 (2009)

Kriteriesøgninger:

- Forebyggende besøg
- Helbredsoplysninger

SQL-søgninger

- KOMDA_INDEX_PERSON_DIAGNOSER_NY
- KOMDA_INDEX_PERSON_JOURNALER
- KOMDA_INDEX_PERSON_BOPÆL_NY
- KOMDA_INDEX_PERSON_HJÆLPEMIDDEL
- KOMDA_INDEX_PERSON_RENGØRING
- KOMDA_INDEX_PERSON_HJEMMEHJÆLP
- KOMDA_INDEX_PERSON_HJEMMEPLEJE
- KOMDA_INDEX_PERSON_SINDSLIDELSE_NY

- KOMDA_INDEX_PERSON_UDVIKLINGSHÆMMET_NY
- KOMDA_INDEX_PERSON_TILDELT-HJÆLP_NY
- KOMDA_INDEX_PERSON_FUNKTIONSVURDERING
- KOMDA_INDEX_PERSON_NOTITSER
- KOMDA_INDEX_PERSON_PROBLEM-HANDLING

Leverandørsøgninger¹²²

- View1
- View2
- View3
- View4
- View5
- View6

Søgepakkesøgninger

- Omsorgssager (DIP.AAS.2)
 - Notitser
 - Funktionsvurdering
 - Tildelte hjælpemidler
 - Journaler
 - Diagnoser
 - Henviſning til hjemmepleje
 - Henviſning til bopæl
 - Henviſning til modul for sindslidende
 - Henviſning til hjemmehjælp
 - Henviſning til rengøring
 - Henviſning til hjælpemidler
 - Problem og handling

KMD Care 2020 (2023)

Der er en del flere søgninger til denne arkiveringsversion.

Kriteriesøgninger:

- Visiteret hjælpe til borger

¹²² Beskrivelserne til leverandørsøgningerne i denne version er identiske med de forrige, men hvor æ, ø og å er udskiftet med række andre tegn, som gør det både forvirrende at læse og grimt at se på. De er derfor udeladt her.

- (tre test-søgninger)
- Borger tilknytning til plejebolig
- Plejeplaner
- Funktionsvurdering
- Dokumenter
- Klientjournal
- Rehabiliteringsplaner
- Service ydet til borger
- Målinger
- Noter
- Henvisninger
- Visitation af bolig
- Visiterede hjælpemidler
- Medicinering af borger
- Visiterede behandlingsydelser
- Henvendelse på cpr
- Boligændringer

SQL-søgninger

- Personnel koblet til afdeling
- test_area admin
- test vis
- Filtyper
- DOKTEST
- test jour med afdeling
- Henvisningstyper
- antal visiterede behandlingsydelser pr år
- DOKTEST_Ramme for Demens(F)
- DOKTEST_Korrespondance-A (F)
- Plejelplan
- Client_Journal fordelt år
- Test nevisning samt distrikt og område
- år og måned
- test_område
- Test af relationer mellem moduler og områder.

- testesbjerg
- DOKTEST_Ramme for tryksår(F)
- cpr længde_test
- testområde og distrikt

Leverandørsøgninger

- AV_VisiteredeBoliger
- AV_Visiterede_ydelser
- AV_Funktionsvurderinger
- AV_klientjournaler
- AV_Plejeplaner
- AV_Dokumenter
- AV_Medicineringer
- AV_Hjælpemidler
- AV_Målinger

Søgepakkesøgninger

- Omsorgssager (DIP.AAS.2)
 - Funktionsvurdering
 - Målinger
 - Notater
 - Tilknytning til plejebolig
 - Visiterede behandlingsydelser
 - Henvendelser
 - Boligændringer
 - Visiteret hjælp
 - Klientjournal
 - Dokumenter
 - Service modtaget
 - Rehabiliteringsplaner
 - Visiterede behandlingsydelser
 - Henvísninger
 - Visitation af bolig
 - Visitation af hjælpemidler
 - Vurdering af borger

Som man kan se, kom det helt store spring med søgemulighederne med den nyeste version. Det giver også mest mening at bruge denne udgave til søgning for arkivets nuværende brugere, da det er denne arkiveringsversion, der indeholder alle oplysninger om aktive og nyligt inaktive klienter. Ifølge en rapport om IT-afleveringer, udarbejdet af Aalborg Stadsarkiv i forbindelse med den første aflevering af Care i 2003, var kommunen interesseret i at få arkiveret Care så man kunne slette inaktive klienter.¹²³ Hvis man rent faktisk begyndte at slette inaktive klienter, er det i de ældre arkiveringsversioner, man finde personer, der enten døde eller på anden vis udgik af systemet før 2006. Ifølge samme rapport fra 2003 forklarede Aalborg Kommunes Ældre- og Handicapforvaltning deres mest almindelige søgeindgange i systemet, og kom til at ligge til grund for arkiveringsversionens søgninger i den første aflevering.¹²⁴ Som man kan se, gik de samme søgninger med nogle få opdateringer igen i den følgende aflevering (KMD Care 2006). Langt de fleste søgninger er på klientoplysninger. I de første to udgaver er det samtlige søgninger med undtagelse af leverandørens View3 (mellemkommunal refusion¹²⁵) og View5 (Køreliste - del A - Personaleoplysninger).

Der har været krav om SQL-forespørgsler fra leverandørerne i alle tre bekendtgørelser. Det har også hele tiden været et krav, at de skulle være tilpasset arkiveringsversionerne og ikke f.eks. taget fra driftsversionen af systemet.¹²⁶ Til de første to arkiveringsversioner af KMD Care virker de dog ikke, og man får i stedet en fejlmeddelelse. Leverandørsøgninger fra KMD Care 2020 (2023) er der til gengæld ingen problemer med. Der er formentlig tale om en fejl i enten migreringen eller i en opdatering af ASDA. At det ikke er blevet opdaget tyder på, at de ikke bliver brugt. Hvis fejlmeddelelsen skyldes, at SQL-en ikke virker på de migrerede arkiveringsversioner er problemet dog ikke større end, at arkivet kan oprette opdaterede kopier.

I den nyeste arkiveringsversion kom der langt flere søgninger til. De fleste i kategorien "SQL-søgninger" er dog testsøgninger, som formentlig ikke er tiltænkt egentlig brug udenfor KOMDA. Mulighederne indenfor kriteriesøgning er også blevet udvidet. Nogle af dem lægger sig tæt op ad leverandørsøgningerne, mens andre er nye. Det afspejler, at arkiveringsversionernes nuværende brugere (forvaltningen) har

¹²³ Aalborg Stadsarkiv, *Kortlægning, aflevering og genrejsning af IT-arkivalier*, s. 8.

¹²⁴ Ibid.

¹²⁵ Teknisk set handler det også om klienter, nærmere bestemt udgifter til dem, men det er kommunale økonomiske anliggender, som klienterne selv formentlig aldrig hører om.

¹²⁶ BEK nr. 342 af 11/03/2004, bilag 3, punkt C. BEK nr. 1007 af 20/08/2010, bilag 6, punkt 6.C og 6.D. BEK nr. 128 af 12/02/2020, bilag 6, punkt 6.C og 6.D.

både flere og mere specifikke behov end tidligere. Måske endda at arkiveringsversionerne også bliver brugt mere end tidligere.

Søgepakkesøgningerne er ligesom de øvrige typer søgninger nogenlunde ens i de to tidligste arkiveringsversioner og lidt anderledes i den nyeste. De typer service, kommunen kan tilbyde borgere, har dog ikke nødvendigvis ændret sig særlig meget. Om nogen er det faktisk blevet mindre tydeligt, hvilke ydelser og services, man får oplysninger om. I praksis er oplysningerne dog mere eller mindre de samme, som tidligere. I hvert fald i mine eksempler. Ændringerne har snarere at gøre med, at man bruger systemet og dets arkiveringsversioner anderledes.

Målet med søgning i arkiveringsversioner af Care er i øjeblikket at kunne indtaste et CPR-nr og få alt relevant om personen ud. KOMDA har sørget for, at man kun behøver at skrive CPR-nummer på én måde (med bindestreg) selvom det kan stå på forskellige måder i tabellerne. De har i mange tilfælde også givet kolonner mere brugervenlige navne, så man ikke behøver at læse dokumentationen for at forstå, hvad man har fået ud. En undersøgelse af, hvordan arkiveringsversionerne bruges i øjeblikket og af hvem ville kunne afdække om det kun er forvaltningen, der har forespørgsler eller om der f.eks. også er egen-access-sager til dette system. Målgruppen er dog i første omgang forvaltningen, eller rettere at finde det forvaltningen, spørger efter, og det er deres behov man har tilgodeset i udviklingen af søgninger. Der er en vis grad af indforståethed i søgningerne og deres resultater, der bl.a. kan ses ved at ingen af dem har beskrivelser (bortset fra leverandørsøgningerne). Man må forvente, at Ældre- og Handicapforvaltningen og Aalborg Stadsarkivs arkivarer er i stand til at forstå, hvad man søger efter, og det materiale, man finder.

Når der kun kan søges på oplysninger om klienter, betyder det at der er meget data i mange tabeller, som slet ikke involveres og som en almindelig bruger af ASDA ikke kan bruge. F.eks. oplysninger om tilgængelige hjælpemidler i kommunen på et givent arkiveringstidspunkt, oplysninger om ansatte eller oplysninger fra de ansattes interne kommunikation (adviser). Der er således ganske meget information om personer i systemet, som arkivet ikke er sat op til at kunne bruges endnu. Det kan dog sagtens lade sig gøre. For nu er det måske vigtigere at spørge om man rent faktisk får det hele med. Her er det ikke kun en overvejelse om at finde det, forvaltningen er vant til at finde i driftsversionen af systemet. For egen-accesssager kan det også være værd at vurdere om en borger har enten ret til eller brug for noget af al den ekstra

information, man har undladt at kassere. F.eks. intern kommunikation om en borgers sag eller den basisinfo, der findes i stamdatatabellen, som f.eks. ikke er med i søgepakken. Denne information har en sagsbehandler måske ikke brug for, men ville potentielt kunne give en borger et bedre overblik over sin egen sag. Eller en sag for en af deres nærmeste, hvis der f.eks. gives adgang for pårørende til svært handicappede eller demente personer, der fungerer som værge.

Søgning på dokumenter

Det meste litteratur om digital arkivering handler om, hvad der i denne sammenhæng kaldes dokumenter. Det inkluderer, som tidligere nævnt, også lyd, video, billeder osv. I arkiveringsversioner som KMD Care giver det kun meget lidt mening at søge udelukkende på dokumenter. Det er teknisk set muligt, hvis man kan finde den tabel, hvor et dokument oprindelige titel står. Dokumenter har ikke selvstændige metadatafiler, der kan bruges til fremsøgning ud fra f.eks. titel, forfatter, datering eller sagsmæssige tilhørsforhold, selvom denne information er registreret.

Dokumenter i arkiveringsversioner kan findes på lidt forskellige måder. For KMD Care 2020 (2023) er den letteste måde at finde dem på at bruge søgepakken og scrolle til bunden, hvor dokumenter ligger i en selvstændig kategori, hvis en person har dokumenter tilknyttet. Her bør alle dokumenter fra de tidligere udgaver også ligge, hvis der er nogen. Hvis der er tale om en person, der ikke findes i den nyeste udgave, er det en anden sag.

For de to arkiveringsversioner, der indeholder dokumenter, er der ingen oversigt over hvilke tabeller, de ligger i. Her skal man i stedet scrolle igennem tabellerne og kigge efter et dokumentikon. Det man egentlig leder efter, er tabeller med kolonner, der er opmærket med *functionalDescriptions* af dokumenterne. Hvorfor er det vigtigt? Dokumenternes placering i arkiveringsversionen er en del af deres kontekst, deres oprindelige orden og deres digitale proveniens. Det er derudover her, hvor dokumenternes metadata står. I litteraturen betragtes metadata for dokumenter som ekstremt vigtigt. I danske arkiveringsversioner er det knap så vigtigt, men skal alligevel markeres via *functionalDescriptions*, som tidligere forklaret. Dokumenternes metadata er derfor en del af arkiveringsversionernes tabelindhold. I KMD Care 2020 (2023) er der kun fire tabeller med dokumenter: ACTION, LINKS, NOTICES og OBJECTIVE. Ingen

af dem indeholder *alle* dokumenter, men LINKS indeholder links til alle dokumenter i systemet, der er tilknyttet klienter. I NOTICES findes adviser, hvor man kan finde detaljer om f.eks. manuelt oprettede henvisninger, som ikke er med i søgepakkens resultater.

I KMD Care 2006 (2009) er den eneste tabel med et dokumentikon Dokument-MellemTabel, som blev tilføjet af arkivet i forbindelse med migrering. Det betyder, at alle dokumenterne ligger samlet ét sted. Det sted er bare ikke det oprindelige. For at kunne se dokumenternes metadata skal man igennem “vis relationer” og finde den tabel, dokumentet oprindeligt kom fra. I mellemtabellen er det ikke indikeret, hvilken relation, der er relevant for hvert enkelt dokument, så det kan måske være en udfordring at finde dokumenter ad denne vej. Det giver langt mere mening i stedet at lede efter de oplysninger, som dokumentet er knyttet til. Der er ingen SQL-søgninger til dokumenter i denne arkiveringsversion, heller ikke i søgepakken. Hvis man klikker på en række tilfældigt udvalgte dokumenter i fanen med dokumenter, er de alle meget korte tekster, der ligner tabelindhold fra de to andre udgaver. Det er måske her meget af det indhold, der lader til at mangle i søgepakkens resultater, er forsvundet hen - det er bare ikke dokumenteret nogen steder. Som et eksempel kan vi tage den tredje klient. Personen havde 34 journalposter i KMD Care 2002 (2003), men kun 4 i KMD Care 2006 (2009). Søger man via mellemtabellen og journaltabellen dukker der fire poster op. Tre matcher dem fra søgepakkens resultater. Hvor de øvrige 30 er blevet af, er stadig et mysterium.

Hvis man vil søge ud fra noget andet end en person, bliver det lidt mere kompliceret. I KMDs tekniske bemærkninger til KMD Care 2020 (2023),¹²⁷ nævnes der lydfiler, og hvis man vil søge ud fra filtype, er det gennem fanen med dokumenter, hvor både oprindelig filnavn (der ofte indeholder oprindelig filtype) og arkiveret filtype står.

Diskussion: Tilgængelighed og brug

Det meste forskning indtil videre har haft med adgangstilladelse af forskelligt omfang til forskellige brugergrupper at gøre. Afhængig af om man har tænkt over adgang, da man arkiverede en database, kan adgangstilladelse af forskelligt omfang dog også

¹²⁷ KMD Care 2020, kontekstdokument nr. 13, KMD, Notat med tekniske bemærkninger vdr. arkiveringsversionen, s. 16.

hurtigt blive en udfordring, hvis databasens struktur ikke tilgodeser en passende måde at facilitere det på uden meget tidskrævende indblanding fra arkivet.

Brug af digitale arkivalier er ikke belyst på dansk eller i nogen dansk sammenhæng generelt. Derfor må vi til udlandet for at kunne drage fordel af andres erfaringer. Her er tilgængeliggørelse af data fra det offentlige blevet undersøgt i England. Den første udfordring er at få tilladelse til overhovedet at få adgang til at bruge data og dokumenter til f.eks. forskning eller personlig brug.¹²⁸

Forskernes perspektiv på genbrug af administrative data fra det offentlige er blevet belyst af Anna Sexton, et al., der har undersøgt, hvordan tilladelser til genbrug af data fungerer i praksis for fire forskellige typer administrative data. Her er det gennemgående tema at indhente samtykke fra dem, der ejer dataene, og om det er nødvendigt at konsultere dem, data omhandler.¹²⁹ De fire typer data, som Sexton, et al. har kigget på er indenfor uddannelse, transport, energi (forbrug) og sundhed. Mens der generelt ikke er mulighed for at takke nej, til at data bliver indsamlet i første omgang, er der i de forskellige kategorier lidt forskellige tilgange til, om der kan takkes nej til at data bruges til sekundære formål såsom forskning. På uddannelsesområdet skal der i England kun indhentes tilladelse fra datasubjekter (dem data omhandler), hvis de administrative data skal kobles sammen med data, der ikke kommer fra myndigheder (f.eks. indsamlet af et universitet ifm. forskning).¹³⁰ Data på transportområdet er for det meste "åbne data" uden personfølsomme oplysninger, der er umiddelbart offentligt tilgængelige. Dem skal der således heller ikke indhentes tilladelse til, hverken fra datasubjekter eller fra myndighederne.¹³¹ I energisektoren kommer data ikke kun fra myndighederne, men også fra private virksomheder, der leverer strøm og gas til borgere. Her skal der indhentes tilladelse fra selskaberne, der tilsyneladende er glade for at sige nej. Her er der dog forandringer på vej, som gør at det i fremtiden bliver datasubjekterne (kunderne), der hver især skal give tilladelse.¹³² Ud af de fire typer data fra denne undersøgelse er sundhedsdata den mest sammenlignelige med data fra KMD Care. I England har de for sundhedsdata en opt-out-model med to niveauer

¹²⁸ En ting er at få adgang til at kigge i data og dokumenter på et arkiv, noget andet er at få lov at tage et udtræk af det med hjem, hvor man selv eller ens institution skal stå for datasikkerheden.

¹²⁹ Sexton, et al, "The role and nature of consent in government administrative data."

¹³⁰ Ibid. s. 9.

¹³¹ Ibid. s. 10.

¹³² Ibid. s. 10-11.

(et for egen læge og et for NHS Digital, den britiske sundhedsstyrelses dataservice). Systemet fungerer ifølge Sexton, et al.'s interviewpersoner ikke optimalt, og der skal heller ikke indhentes tilladelse, hvis data er blevet pseudonymiseret. Myndighederne er dog samtidig en del mere påpasselig med at give adgang til data i det hele taget, end man er indenfor de øvrige områder.¹³³

Om der kan takkes nej til brug af ens data på et individuelt niveau, har meget stor praktisk betydning for hvor besværligt det er at give adgang til data. Hvis individer skal udskilles fra et datasæt, kan det hurtigt føre til meget lang og manuel rensning af data. Denne udfordring havde NHS Digital ifølge Sexton et al. også, og deres løsning var et fjerne muligheden for opt-out i et par år¹³⁴ fremfor f.eks. at stoppe adgang til data indtil problemet var løst.

Hvis vi tager KMD Care som eksempel vil det være relativt ligetil at sortere individer fra, hvis man f.eks. kun ønsker et datasæt bestående af journaler. Men hvis man ønsker et datasæt bestående af data fra mange tabeller fra flere forskellige moduler eller hele systemet, kræver det en del mere arbejde at forberede. Man ville formentlig kunne bruge de samme metoder, som når man skal finde information om et individ, men man ville også have alle de samme udfordringer med, om man har fundet alt, hvad der er.

Sexton et al. har kigget på data, der i en dansk kontekst både kan sammenlignes med data, der ligger hos en myndighed og data, der er blevet arkiveret, men endnu er underlagt tilgængelighedsfrist. Ejerskabet kan derfor være meget blandet, særligt i datasæt med mange datakilder, hvilket hurtigt kan gøre forskningsadgang meget kompliceret. Sexton et al. konkluderer, at der i øjeblikket ikke er én model for at indhente samtykke og tilladelser til genbrug af data, som alle er enige om. Derfor foreslår de, at man udvikler en model, der har mindre fokus på direkte samtykke og mere har form af en social kontrakt med tillid og gennemsigtighed som udgangspunkt for brug af administrative data.¹³⁵ Det Sexton et al. efterspørger er således ikke kun bedre lovgivning, men også et bedre forhold mellem borgere, myndigheder og forskere.

Lovmæssigt har vi ikke denne udfordring i Danmark. I hvert fald ikke med arkiverede data. Her er det myndigheden, der ejer data, og involverede personer har ifølge arkivloven ikke noget at skulle have sagt. På de kommunearkiver, jeg har haft

¹³³ Ibid. s. 12.

¹³⁴ Ibid. s. 12.

¹³⁵ Ibid. s. 15.

med at gøre, er det i praksis arkivet, der giver adgang uden, at man behøver konsultere arkivskaber.

I den britiske undersøgelse er der i ingen af tilfældene tale om, at forskere har fået lov til selv at hente data direkte ud af den database, de kommer fra. Det er myndigheden selv, der har gjort det og overført et udtræk til forskning. For at sådan en manøvre kan lykkedes kræver det både, at forskerne ved, hvad for noget data de skal spørge efter, og at myndigheden er i stand til at finde det, der bliver spurgt efter. Det er så vidt jeg ved endnu ikke blevet undersøgt, hvor godt denne dynamik fungerer. Uden direkte adgang til databaser eller let tilgængelige arkivregistre, bliver forskerne afhængige af det datakatalog myndighederne gør tilgængeligt via deres egne services. Vi har også noget lignende i Danmark.¹³⁶ Udover byggesager, ser kommunale digitale arkivalier dog ikke ud til at være tilgængelig via en dataservice endnu. Her er der ingen datasæt, der er forberedte til brug på forhånd. Det er, igen, relevant at sammenligne med den britiske undersøgelse af myndighedsdata, da digitalt skabt materiale arkiveres både oftere og efter meget kortere tid, end det var tilfældet i papirtiden. Med GDPR er det også meningen, at myndighederne skal slette data, der ikke længere bruges, hvilket gør arkivet til det eneste sted, man kan finde data. Derudover har arkiverne den fordel, at de er langt bedre til at tydeliggøre, hvor eller hvem data kommer fra end flere af de danske dataservices.

Borgenes brug af arkiverne er også blevet undersøgt i både Storbritannien og Australien. Her er der faktisk skrevet ganske meget, og det handler særligt om børnesager, da adgang til dem er både kompliceret og følsomt. Fokus er her på kvaliteten af dokumenter og den måde, hvorpå de skabes, og samtidig den måde de tilgængeliggøres på. Victoria Hoyle, et al. har undersøgt, hvordan egen-access opleves af personer, der har været i statens varetægt som børn (fjernet fra deres forældre) i Storbritannien. Her har man en egen-accesslov, der minder meget om Danmarks. Dvs. at man har ret til at få udleverede personoplysninger om sig selv fra det offentlige. Ligesom i den danske lov er der undtagelser, hvis materialet indeholder oplysninger om en tredjepart. Når det gælder personoplysninger fra det offentlige, er der med sidst-

¹³⁶ F.eks. <https://datafordeler.dk/> for store statslige datasæt og <https://weblager.dk/> for kommunale bygge- og miløsager. Danmarks Statistik har også en liste over forskellige dataservices: <https://www.dst.dk/da/TilSalg/data-til-forskning/generelt-om-data/data-fra-andre-kilder>

nævnte typisk tale om andre personer, hvis rettigheder for beskyttelse af personfølsomme oplysninger skal overholdes.¹³⁷ Når sager involvere flere personer, løses denne udfordring som regel ved at censurere de dele, der har direkte med andre personer at gøre, f.eks. steder hvor de nævnes med navn. I den britiske undersøgelse hænger udfordringerne for egen-access i særlig grad sammen med mangelfuld journalisering i det offentlige, der ikke inkluderer børnene i tilstrækkelig grad, mangel på arkivalier, der afspejler personernes egne oplevelser, samt meget varierende og til tider overdreven censur i et forsøg på at overholde ovenfor nævnte undtagelse til loven om egen-access.¹³⁸ Det er sidstnævnte, der har at gøre med arkivernes arbejde. Hoyle, et al. skelner ikke mellem, om materialet ligger på arkiv eller hos arkivskaber, men derimod på kvalifikationer hos de ansatte, der skal vurdere, hvor meget der skal censureres. De finder, at det typisk er ansatte med ekspertviden indenfor GDPR, der censurerer *mindst*. Suellen Murray stod tidligere bag en lignende undersøgelse i den australske delstat Victoria og nåede også frem til lignende konklusioner, igen med særlig fokus på censureringen af data om en tredjepart.¹³⁹

De britiske og australske undersøgelser af egen-access fokuserer ikke specifikt på det digitale. De er mindre optagede af, hvilken form eller medie dokumenterne har, og mere af hvilke oplysninger, der gives adgang til, hvordan det håndteres af de offentlige institutioners ansatte, og hvordan det opleves af de personer, dokumenterne handler om. Der er dog både fordele og ulemper, som er specifikke for digitale arkivalier, som er værd at overveje. En fordel ved det digitale er f.eks., at der kan gives adgang uanset, hvor i verden en person befinder sig. Rent teknisk står der intet i vejen for, at man kan få tilsendt arkivalier om sig selv fra Danmark, selv hvis man er flyttet til Japan, uden at arkivet først skal indscanne potentielt flere hundrede dokumenter.

En af udfordringerne med det digitale er at finde alt, hvad en borger har bedt om at få udleveret. At finde ting er også en udfordring ved papir, hvor man tidligere havde for vane at organisere personsager efter faderen i en familie i stedet for enkeltvis. I arkiveringsversioner som Care er der en omvendt udfordring. Grundet den måde databaser fungerer på ligger sager slet ikke samlet. Man kan derfor ikke bare finde en sagsmappe med en person hos en bestemt arkivskaber og forvente, at det hele ligger

¹³⁷ Hoyle, et al., "Recordkeeping and the life-long memory and identity needs of care-experienced children and young people," s. 936.

¹³⁸ Ibid. s. 940-942.

¹³⁹ Murray, "Compassion and Compliance: Releasing Records to Care-Leavers under Privacy and Freedom of Information Legislation."

samlet. Oplysninger ligger i stedet meget spredt og det er op til arkivet (eller KOMDA i dette tilfælde) at samle dem. I princippet er systemer som Care sat op så alle oplysninger om en borger (klient) er påført CPR-nummer. Søger man ud fra CPR, burde man derfor kun finde oplysninger om dén person. Samtidig bør der ikke ligge information om en borger, der *ikke* har påført CPR-nummer, da det ellers bliver næsten umuligt at finde. Det kræver dog, at myndigheden har brugt systemet korrekt, journaliseret korrekt og undgået fejl.

I eksemplet med den fysisk handicappede person findes der i journalerne et eksempel, som formentlig ikke er enestående: Der fandtes en journalpost til en anden person. Der er formentlig tale om en fejl fra den ansatte, som, at dømme ud fra datoerne, havde for vane er indtaste flere dages journaler på samme dag, og derfor havde glemt at skifte klient i systemet. Hvis der bestilles materiale til denne person eller gives tilladelse til at søge på kun denne person, risikerer man dermed, at andre personers informationer kommer ud til uvedkommende uden tilladelse. Da antallet af poster i KMD Care 2020 (2023), hvor journalen i eksemplet findes, er meget overvældende, er det ikke sikkert, at alle 2000-4000 poster tjekkes igennem for denne type fejl inden det sendes til dem, der har bestilt det. Særligt, hvis man slet ikke er klar over, at der findes denne type fejl i systemet, hvilket dokumentationen ikke antyder. Man ville heller aldrig finde denne journalpost, hvis man søger efter oplysninger på den person, den egentlig omhandler. Ved indsamling af kontekstdokumentation til arkiveringsversioner er der, som før nævnt, en kategori til dokumentation af kvalitetssikring, der er let at misforstå. Her kan man måske i stedet *specifikt* spørge efter typiske fejl ved indtastning eller journalisering i systemet.¹⁴⁰ Man ville ikke kunne være opmærksom på alle typer fejl med sådan en liste, men man ville vide, hvilke fejl myndigheden selv kender til.

Tilgængelighed har også at gøre med begrænsninger. Selvom man er nødt til at søge specifikt for at finde noget i en arkiveringsversion er det også med stor lethed muligt (i hvert fald i det nuværende setup) at finde information, man ikke skal bruge og måske slet ikke burde have. F.eks. ved at åbne tabeller uden en søgning. På denne måde kan man f.eks. finde samtlige felter med tekster fra visitationer på borgere - meget personfølsomme oplysninger - hvis blot man ved, hvilken tabel, det står i. Derfor

¹⁴⁰ En udfordring her kan måske være, at de ansatte i kommunen nok ikke er så glade for at indrømme den slags.

kan man også argumentere for, at den nuværende tilgang, hvor det er arkivet, der skal finde data for en, også har sin plads.

I en dansk kontekst har det også handlet meget om børnesagerne. Der er bare ikke blevet udgivet noget endnu.¹⁴¹ De er de mest komplekse oplysninger at skulle give adgang til, men det er langt fra alle egen-accesssager, der handler om dem. Hos Rigsarkivet er alle såkaldte "særlige personsager" samlet i én, og dem var der 3250 af i 2024. Til sammenligning var der 5000 henvendelser om eksamensbeviser og 6000 om fødselstidspunkt.¹⁴² Disse to typer oplysninger er langt simplere at finde end hele personsager. Her er der også tale om oplysninger, der har konkrete praktiske formål for borgere, og hvor mange af de overvejelser, vi har som forskere, er mere eller mindre ligegyldige. Det er i hvert fald lige indtil, at man ikke kan finde en eller flere karakterer til genrejsning af et eksamensbevis. Hos kommunerne er den seneste statistik over såkaldte "tilgængelighedssager"¹⁴³ fra 2020, hvilket er før lovændringen om egen-access. Denne oversigt findes som bilag til Rigsarkivets rapport om tilgængelighedspraksis 2020, og antallet var generelt *meget* lavt i de enkelte kommuner.¹⁴⁴ Der skelnes heller ikke mellem typer af oplysninger, der ansøges om, eller om det er digitalt eller papir.

Andre vinkler på emnet

Denne undersøgelse har kastet flere åbne spørgsmål af sig. Nogle er værd at undersøge, mens andre snarere giver anledning til diskussion om, hvem der bruger arkiverne og hvordan. Jeg har kun været omkring ganske få af de mange perspektiver på tilgængelighed og tilgængeliggørelse, der findes om arkiveringsversioner. En oplagt undersøgelse i forlængelse af denne kunne være, hvordan arkiveringsversionerne bruges lige nu af arkiver og deres klienter (borgere og forvaltninger). Hvor meget bruger de f.eks. søgepakker og hvor ofte, søger de manuelt efter oplysninger? En dybere undersøgelse af arkiveringsversionernes indhold og struktur i praksis er også relevant

¹⁴¹ Denne oplysning stammer fra Marianne Rostgaard som medlem af Tilgængelighedsudvalget.

¹⁴² Ibid.

¹⁴³ Adgangstilladelser til ikke umiddelbart tilgængeligt materiale.

¹⁴⁴ Rigsarkivet, Rapport om praksis på tilgængelighedsområdet i 2020, jf. arkivlovens § 39, stk. 3, bilag 2.

i denne sammenhæng. For KMD Care specifikt er der f.eks. meget data fra 2006-udgaven, der ikke kommer med i KOMDAs søgepakkesøgning. Her kunne man også undersøge, hvor meget data der *ikke* bliver brugt, og som der slet ikke er fokus på. Som jeg tidligere har nævnt, gemmer man meget ekstra data, som potentielt godt kan bruges, men som ikke bliver det (endnu). En konsekvens af kun at fokusere på de vigtigste data i en arkiveringsversion kan være, at alt det andet, som ligger deri ikke kommer med i arkivbeskrivelsen og ikke bliver forklaret i kontekstdokumentationen. Det kan gøre tilgængeliggørelse og brug af disse data sværere.

Undersøgelsen rejser også spørgsmål om, hvor meget man skal vide i forvejen om både myndighedens virke og IT for at kunne bruge arkiveringsversioner. Jeg har generelt argumenteret for mere og bedre dokumentation, men hvor meget mere og præcis hvad, der er behov for, afhænger af, hvilke forudsætninger man kan forvente af brugerne. Længe inden de bliver interessante for historikere, kan de være interessante for f.eks. sociologer, for at nævne et eksempel, der er relevant for Care specifikt. Er de trænede i kildekritik og i at kunne gætte kontekst ud fra relativt sparsomt materiale som historikere er? Er de trænede i databaser og datavidenskab (ikke at historikere nødvendigvis er det)? Det er værd at undersøge hvor tilgængelige arkiveringsversionerne er i forskellige scenarier og for forskellige brugergrupper. Her tænker jeg særligt på, hvor meget dokumentation, der skal med og hvilken. Det kommer let til at hænge sammen med både de forudsætninger en bruger har, og hvor meget af arbejdet, arkivet vil udføre for en. Hvis man kommer som sociolog og beder om en udtræk af journaler fra x antal klienter i y periode, og arkivet kan skaffe dette for en, så behøver man stort set ikke læse kontekstdokumentationen. Journaler er allerede et kendt begreb i offentlig forvaltning, og man får dem på denne måde tilsendt i et format, der er til at arbejde med. Hvis man derimod vælger en løsning med langt mere selvbetjening, kræver det en langt højere grad af brugervenlighed og forklaringer (både af kontekst og det tekniske). Som arkiveringsversionerne hos KOMDA fungerer nu, er denne løsning ikke blevet brugt ret meget endnu, men hvis man for alvor vil lade arkiveringsversioner blive brugt til forskning, er det formentlig selvbetjening, der skal sikre, at man ikke ender med en årelang sagsbehandlingstid eller skyhøje omkostninger til ansatte, som kommunerne ikke nødvendigvis kan eller vil bære. Her er sagen måske en anden for Rigsarkivet, der har flere ressourcer at arbejde med, men samtidig også har *mange* flere brugere.

For de ældste arkiveringsversioners vedkommende er det også værd at undersøge samspillet mellem papirarkivalier og digitale arkivalier. Som beskrevet i undersøgelsen er der ingen dokumenter i KMD Care 2002 (2003). Det betyder som sagt ikke, at man slet ikke brugte dokumenter i arbejdet med ældre og handicappede. Overgangen fra papir til digital skete ikke fra den ene dag til den anden. Links-tabellen i KMD Care 2002 (2003) er tom, og der står ikke noget om, hvor eventuelle dokumenter ligger. Hvis man vil have adgang til dokumenter, der eventuelt er bevarede om en person, der f.eks. var på plejehjem i 1998, skal man således søge om adgang til mindst to forskellige systemer (de kunne måske ligge i et ESDH-system) og potentielt til papirarkivalier.

Endelig vil det være oplagt at kortlægge, hvordan arkiverne og arkivsamarbejderne tester og kvalitetssikrer arkiveringsversioner udover det obligatoriske (ADA).

Konklusion

Den danske bevaringsstrategi har siden 1970'erne først og fremmest handlet om at bevare IT-systemers indhold. Rigsarkivet stiller ingen krav til, hvordan digitale arkivalier skal tilgængeliggøres. Det er i stedet op til de modtagende arkiver eller arkivsamarbejder (KOMDA i dette tilfælde). Jeg har undersøgt og diskuteret både de valg, KOMDA har taget på den front, men også hvordan dette er en konsekvens af den bevaringsstrategi vi har i Danmark. Man konverterer data til bevaringsformater, og ændrer således på arkivaliers, men man forsøger samtidig at beholde dem i deres originale struktur, eller deres originale orden, hvis man overfører dette begreb til databasesystemer. De digitale arkivaliers proveniens og kontekst dokumenteres i kontekstdokumentationen, og det lykkedes for mine eksempler også langt hen ad vejen. Der er dog også dele af indholdet, som ikke får meget opmærksomhed.

I de specifikke eksempler, jeg har brugt som cases, findes der både personsager, personoplysninger om ansatte og oplysninger om forvaltningens drift (f.eks. Hjælpermidler og ruter), som ikke er personoplysninger. Det er førstnævnte, der er hovedårsag til, at arkiveringsversionerne overhovedet betragtes som bevaringsværdige. Det er derfor disse oplysninger det meste dokumentation og langt de fleste søgninger har fokus på. For nu er denne prioritering også ganske legitim, da arkivalierne er relativt unge. I fremtiden vil arkiveringsversionernes ekstraoplysninger dog åbne for mulige

historiske studier, som vil kunne drage fordel af en tilgang, hvor alle bevarede typer oplysninger som minimum nævnes og defineres tydeligt i kontekstdokumentationen. I mine cases har jeg også kunne dokumentere, at den danske bevaringsstrategi med migrering betyder, at man ændrer nok på arkivaliernes originale orden (struktur) til, at det har betydning for fremsøgning.

For at kunne håndtere adgangstilladelse og brug af data fra offentlige institutioner, skal de være organiserede og bevaret på en hensigtsmæssig måde. Med andre ord skal den tekniske tilgængelighed være i orden. Det gælder særligt, hvis man vil forsøge at gøre materialet lettere tilgængeligt for brugere, uden at det skal igennem en arkivar først. Særligt det sidste punkt er man ikke helt nået til endnu.

Der er endnu mange perspektiver på emnet, der kan undersøges, herunder arkiveringsversionernes faktiske brug, deres kvalitet og test deraf, sammenhængen mellem papir og digitale arkivalier, samt flere brugerperspektiver på arkiveringsversionernes tilgængelighed i praksis.

Bibliografi

Litteratur

Aalborg Stadsarkiv. *Kortlægning, aflevering og genrejsning af IT-arkivalier: Afsluttende rapport fra IT-Arkiv projektet i Hals, Frederikshavn og Aalborg kommuner*. 2003.

Bak, Greg. "Digital provenance." *Archival Science*, 2024, 24: 847-869.

Bak, Greg; Rostgaard, Marianne. *The Nordic Model of Digital Archiving*. Routledge, 2023.

Bloch, Elisabeth; Damsgaard, Peter; Hansen, Else; Nielsen, Lise Qwist. *Dokumentation i en digital tid - ESDH, arkiver og god forvaltningsskik*. Arkivforeningen, 2017.

Breinstrup, Thomas. "KMD køber Rambøll Informatik." *Berlingske*, 1. juli, 2011. URL: <https://www.berlingske.dk/virksomheder/kmd-koeber-ramboell-informatik>

Hoyle, Victoria; Shepherd, Elizabeth; Lomas, Elizabeth; Flinn, Andrew. "Recordkeeping and the life-long memory and identity needs of care-experienced children and young people." *Child & Family Social Work*. 2020;25:935–945.

MacNeil, Heather. "Archival Theory and Practice: Between Two Paradigms." *Archivaria*, 1994-04, Vol.37 (37), p.6-20.

Mordell, Devon. "Critical Questions for the Archives as (Big) Data." *Archivaria*, 2019, 87: 140-161.

Murray, Suellen. "Compassion and Compliance: Releasing Records to Care-Leavers under Privacy and Freedom of Information Legislation." *Social Policy & Society* (2014) 13:4, 493–503.

Pacheco, André; Carlos Guardado Da Silva; Maria Cristina Vieira De Freitas. "A metadata model for authenticity in digital archival descriptions." *Archival Science* (2023) 23: 629–673.

Rigsarkivet. *Årsrapport 2024*. 2025.

Rigsarkivet. *Rapport om praksis på tilgængelighedsområdet i 2020, jf. arkivlovens § 39, stk. 3.* 2021.

Sexton, Anna; Shepherd, Elizabeth; Duke-Williams, Oliver; Eveleigh, Alexandra. "The role and nature of consent in government administrative data." *Big Data & Society*, 2018: 1–17.

Sørensen, Maja. "Indsamling af kontekstdokumentation til kommunale arkiveringsversioner." *Aalborg Universitet*, 2024.

Zhang, Jane. "Original order in digital archives." *Archivaria*, 2012-10, Vol.74, p.167.

Lovgivning

Aktivloven, LBK nr. 1201 af 28/09/2016.

Lov om ændring af arkivloven (Ret til indsigt i egne personoplysninger). LOV nr. 1943 af 15/12/2020.

Bekendtgørelse om bevaring og kassation af arkivalier i primærkommunerne, BEK nr. 1000 af 28/09/2004

Statens Arkiver, Cirkulæreskrivelse om kassation af dokumenter i kommunernes it-systemer, 2010.

Bekendtgørelse om bevaring og kassation af digitalt skabte data og dokumenter fra kommunerne, BEK nr. 183 af 26/01/2018.

Bekendtgørelse om ændring af bekendtgørelse om bevaring og kassation af digitalt skabte data og dokumenter fra kommunerne, BEK nr. 282 af 09/03/2023.

Cirkulære om statslige myndigheders aflevering af elektroniske arkivsystemer til Statens Arkiver, CIR nr. 25 af 08/03/2002 (IT-afleveringscirkulæret).

Bekendtgørelse om arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer, BEK nr. 342 af 11. marts 2004.

Bekendtgørelse om arkiveringsversioner, BEK nr. 1007 af 20/08/2010.

Bekendtgørelse om arkiveringsversioner, BEK nr. 128 af 12/02/2020.

Øvrig dokumentation

Rigsarkivet. *Strategi for digital bevaring 2025*. URL: <https://www.rigsarkivet.dk/wp-content/uploads/2023/04/Strategi-for-digital-bevaring-2025.pdf>

Rigsarkivet. *Vejledning til bekendtgørelse om arkiveringsversioner*. URL: <https://www.rigsarkivet.dk/wp-content/uploads/2022/02/Vejledning-til-bekendtgørelse-nr.-128-af-12.-februar-2020.pdf>

Rigsarkivet. *Vejledende kassationsliste for primærkommunernes IT-systemer*. URL: <https://www.rigsarkivet.dk/wp-content/uploads/2024/02/Kassationsliste-for-primærkommunernes-it-systemer.pdf>

KOMDAs medlemsoversigt:

<https://www.aalborgstadsarkiv.dk/Komda.asp?Komda=KomdaDeltagere>