

Det elektroniske forsknings- og undervisningsarkiv

DEFVA

Januar/februar 1999

Afreportering af projektet ved:

Overbibliotekar Niels S. Clausen
Førstebibliotekar Birgitte Lau
Lektor Niels Erik Wille
Projektkoordinator Katja Guldbæk

Roskilde Universitetsbibliotek

Universitetsvej 1, box 258
DK - 4000 Roskilde
tlf.: + 45 46 74 20 00
fax: + 45 46 74 30 90

RUB

Det elektroniske forsknings- og undervisningsarkiv

Indholdsfortegnelse

1. Resumé (overordnet formål og hovedresultater)	s. 3
2. Projektets baggrund og formål	s. 3
3. Projektets forløb:	
3.0.0. Skematisk oversigt over DEFUA	s. 5
3.1.0. Udvælgelse af relevant litteratur	s. 6
3.2.0. Rettighedsforhandlinger	s. 7
3.2.1. Ophavsretlige forudsætninger	s. 7
3.2.2. Arbejdet med rettighedsklarering	s. 8
3.2.3. Aftalemodeller	s. 10
3.2.4. Erfaringer/konklusioner i forbindelse med rettighedsklarering	s. 11
3.3.0. Digitalisering af dokumenter	s. 13
3.3.1. Indskanning og editering	s. 13
3.3.2. Dokumentstrukturering	s. 14
3.3.3. Dokumentregistrering	s. 14
3.3.4. Dokumentlagring og import/eksportformater	s. 14
3.4.0. Tilgængeliggørelse af digitaliserede dokumenter	s. 15
3.4.1. Docuweb fra Xerox	s. 15
3.4.2. New Publiotech fra Electric Press	s. 16
3.4.3. DOC fra DDE	s. 16
3.4.4. ERes fra Dokutek	s. 17
3.4.5. Digital Library fra IBM	s. 18
3.4.6. Konkluderende bemærkninger	s. 18
4. Projektets modelkarakter	s. 19
5. Erfaringsudveksling	s. 19
6. Videreførelse af projektet	s. 20

Bilag

1. Projektets organisering
2. Anvendelse af bevilling

Afsnit 3.2.

3. Oversigt over materialetyper
4. Oversigt over rettighedshavere
5. Udkast til rettighedsaftale
6. Standardskrivelse til rettighedshavere
7. Aftaleskema
8. Arbejdsrapport
9. Administrativ database

Afsnit 3.3. og 3.4.

- 10-11. Beskrivelse af Xdod og Docuweb
12. New Publiotech fra Electric Press
13. DOC fra DDE
14. ERes fra Dokutek
15. Digital Library fra IBM
16. Kravsspecifikationer (2 stk.)

1. Resumé (overordnet formål og hovedresultater)

Formålet med projektet, Det elektroniske forsknings- og undervisningsarkiv, DEFUA, har været at skaffe adgang til en større mængde central forsknings- og undervisningslitteratur i digital form inden for fagområdet kommunikation ved Roskilde Universitetscenter. Via et elektronisk forsknings- og undervisningsarkiv skal studerende og lærere sikres mulighed for at konsultere materialer efter behov samt udprinte kopier af tekster til brug i undervisningen.

Projektet har i den forbindelse:

- Afprøvet metoder til massedigitalisering af dokumenter.
- Undersøgt/afprøvet metoder til distribuering af digitale dokumenter via elektronisk netværk i en universitetssammenhæng, herunder belyst de tekniske problemer forbundet hermed.
- Belyst juridiske problemer forbundet med etablering og drift af et digitalt arkiv, herunder særlig ophavsretsmæssige spørgsmål i et elektronisk miljø.

Det har ikke i den afsatte projektperiode (godt og vel 1 år) været muligt at nå frem til en egentlig driftsfase, hvorfor målelige resultater i relation til brugerne endnu ikke kan aflæses. Projektledelsen var fra begyndelsen klar over, at den vanskeligste del af projektet - juridisk og tidsmæssigt - var forhandlingerne med rettighedshaverne. Den manglende mulighed for at indgå centrale/kollektive aftaler med rettighedshaverne har forhindret arkivet i på relativt kort tid at opnå den kritiske masse, som ville muliggøre en målelig driftsfase.

Ud over at teste og beskrive forskellige tekniske løsninger er én af projektets konklusioner, at der snarest bør etableres et kollektivt aftalegrundlag med rettighedshaverne. Som et konkret resultat af de mange forhandlinger projektet har gennemført fremlægges et standardkontraktforslag til aftale mellem bibliotek og rettighedshaver(e) (bilag 5).

2. Projektets baggrund og formål

Roskilde Universitetsbibliotek, herefter blot kaldet Rub, betragter formidling af institutionens forskningsresultater som en væsentlig del af sin opgavevaretagelse. Det sker gennem indsamling, registrering og opbevaring af universitetets publicerede forskning. I den forbindelse har biblioteket igangsat forskellige projekter til oparbejdning af erfaring inden for elektronisk publicering. Det er sket i form af netpublicering af indholdsfortegnelser og resumeer af samtlige Ph.d.-afhandlinger ved Roskilde Universitetscenter, samt udvalgte Ph.d.-afhandlinger i fuldtækst. I forlængelse af dette arbejde har man haft et særligt ønske om netpublicering af andet forsknings- og undervisningsrelevant fuldtækstmateriale.

Som bibliotek for et universitet med mange fagområder har Rub oplevet generelt stigende problemer med at kunne tilgodese behovene for litteratur i traditionel trykt form (bøger og tidsskrifter). En kraftig stigning i antallet af studerende på kommunikationsuddannelsen (ca. 1100 studerende) har betydet et stadigt større pres på især dette fags kernelitteratur, hvor mange har brug for adgang til materialerne samtidig. Universitetets notesalg har ligeledes haft stigende vanskeligheder med at tilfredsstille de studerendes behov for kontinuerlig forsyning af særtrykte artikler, kursusnoter og kompendier.

Behovet for at kunne stille central forsknings- og undervisningslitteratur til rådighed *uden begrænsninger* i form af f.eks. fysiske eksemplarer og under *betalingsmæssigt afklarede forhold* har således været stigende gennem mange år. Dette er baggrunden for, at Rub vil søge at bidrage til løsning af problemerne med informationsforsyningen ved at skabe elektronisk tilgængelighed til en større mængde central litteratur inden for faget Kommunikation via et digitalt arkiv.

Formålet med projektet DEFUA har været at undersøge og afprøve dokumenthåndterings-systemer, der kan gøre det muligt for Rub at tilbyde forskere og studerende ved kommunikationsuddannelsen ubegrænset elektronisk adgang til litteratur i fuldtækst. Arkivet skal indeholde alle typer af litteratur, hvoraf noget vil være udgivet af forlag og derfor copyrightbeskyttet, mens andet vil være udgivet via institutter, organisationer m.v. og derfor have en mere blandet ophavsretsmæssig status. Dokumenterne indlægges i arkivet enten via skanning eller kopiering direkte fra elektronisk kilde.

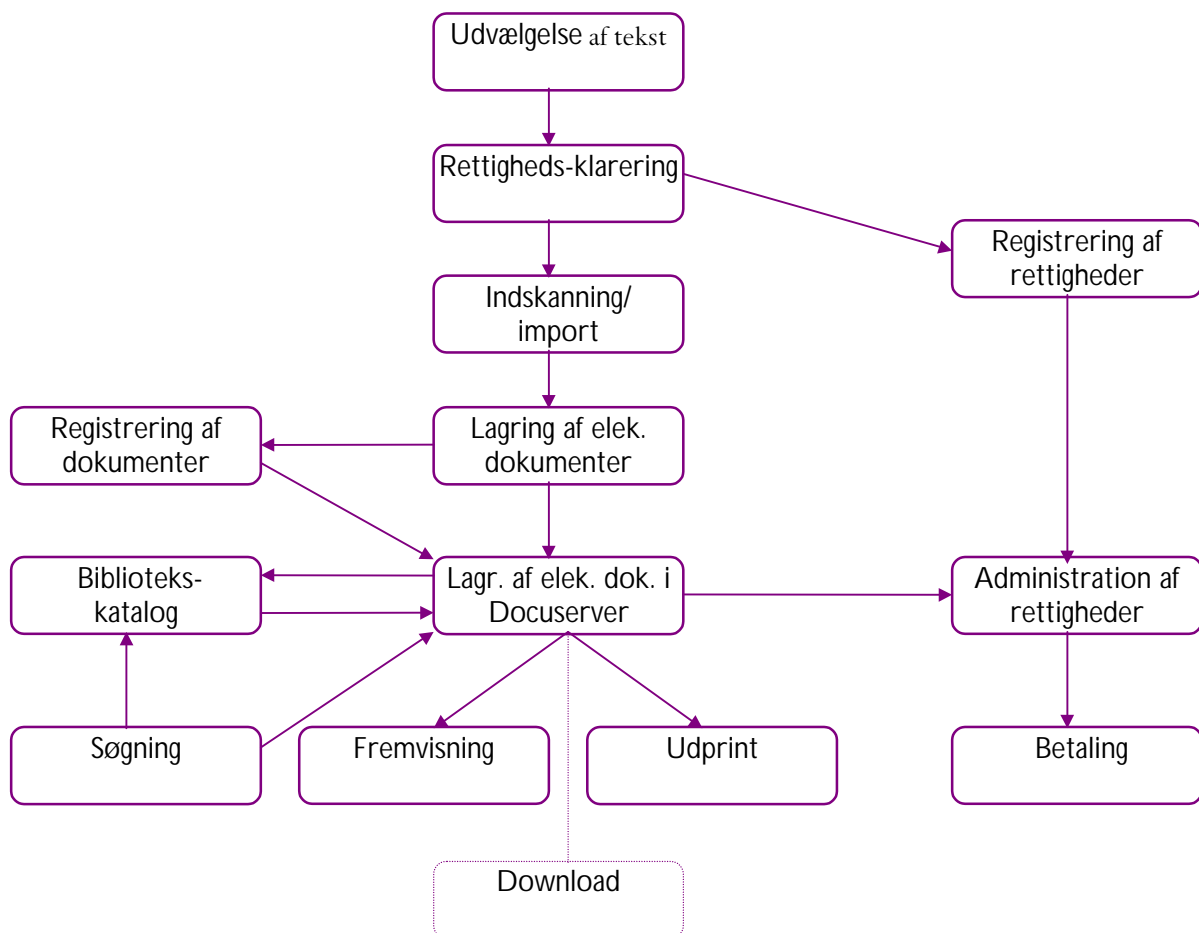
Brugerne skal have adgang til materialerne dels via standard-webbrowsere, dels via biblioteket's almindelige bibliotekssystem, således at man fra en søgning i katalogen (Rubikon) umiddelbart kan klikke sig videre til fuldtækstversionen i det digitale arkiv. Søgning af dokumenter vil enten kunne ske gennem brugen af specificerede søgeord, omfattende de vigtigste bibliografiske data, fritekstsøgning i abstracts, indholdsfortegnelser m.m. Eller ved at browse gennem forskellige dokumentstrukturer, organiseret efter dokumenternes emne eller de kurser, moduler eller workshops, hvori teksterne anvendes.

Systemet skal rumme mulighed for, at der kan udprintes hele eksemplarer eller sider af materiale i arkivet enten på det pågældende instituts lokalprintere, udvalgte printere på biblioteket samt, i tilfælde hvor der kræves større oplag, universitetets centraltrykkeri. Der skal være mulighed for elektronisk bestilling af teksterne og evt. betaling skal kunne ske via betalingskort eller internt konteringsystem.

Da arkivet som nævnt indeholder materiale af blandet ophavsretsmæssigt indhold, er det et vigtigt formål med projektet at belyse de ophavsretsmæssige problemer, der er forbundet med etablering og drift af et sådant arkiv. Projektets formål er på den ene side at undersøge de juridiske aspekter, der sætter rammerne for anvendelsen af materialer i arkivet. På den anden side at undersøge muligheder og betingelser for klarering af brugsrettigheder i større målestok.

3.0. Det elektroniske forsknings- og undervisningsarkiv

En skematisk oversigt:



3.1. Udvalgelse af relevant litteratur

Faget Kommunikation på Roskilde Universitetscenter har en speciel profil sammenlignet med tilsvarende fag på landets øvrige universiteter. Hovedvægten lægges på teori og praksis i forbindelse med formidling til forskellige målgrupper af faglige information. Der anvendes et stort udvalg af faglitteratur, fortrinsvis af nyere dato, med relation til en række forskellige fagområder. For at få et billede af litteraturen, der anvendes i den igangværende undervisning på uddannelsen, er følgende informationsveje blevet benyttet:

- Semesterplaner og kursusbeskrivelser fra foråret 1996 til efteråret 1998.
- Fagets officielle liste med relevante litteraturreferencer (ligger elektronisk på uddannelsens hjemmeside).
- Uformelle samtaler med forskere og undervisere fra uddannelsen.
- Udlånsstatistik fra Rubikon.

Semesterplanerne opdateres umiddelbart for hver ny semesterstart og peger derfor på litteratur, der *benyttes* direkte i undervisningen, mens den officielle litteraturliste ligeledes henviser til litteratur af mere *generel relevans* for faget. Disse giver derfor tilsammen et rimeligt dækkende billede af den litteratur, der måtte være relevant i forhold til et forskningsarkiv inden for fagområdet. De uformelle samtaler med undervisere, samt udlånsstatistik fra Rubikon kan endvidere pege i retning af, hvilke referencer der særlig er pres på. Både semesterplanerne og den officielle litteraturliste udarbejdes i forvejen af uddannelsen og det store arbejde med at indsamle læseplaner og litteraturlister fra undervisere forud for hver semesterstart, behøver biblioteket derfor ikke i første omgang at anvende ressourcer på.

For i projektperioden at opnå størst mulig erfaring med rettighedsklarering af forskellige typer materialer, er i alt 157 titler udvalgt med henblik på en vis differentiering i forhold til:

- Udgiver (forlag, organisationer, forsknings- og uddannelsesinstitutioner m. v.).
- Materialetype ("monografier", tidsskriftartikler, afhandlinger m.m.).
- Alder.

De udvalgte titler er dog alle danske (samt enkelte fra andre nordiske lande). En prioritering begrundet i et ønske om i første omgang at få belyst de komplicerede ophavsretslige spørgsmål på nationalt niveau. Bilag 3 viser bl.a. den konkrete fordeling af udvalgte materialer i forhold til ovenstående grupper.

3.2. Rettighedsforhandlinger

3.2.1. Ophavsretlige forudsætninger

Skanning og indlæsning af et dokument i en elektronisk database er omfattet af ophavsmændenes enerettigheder, som de er fastlagt i den danske lov om ophavsret (OHL), ligesom den efterfølgende tilgængeliggørelse i digital eller trykt form. Derfor forudsætter både indlæsning og brug af ophavsretligt beskyttet materiale, at biblioteket har de nødvendige tilladelser.

Hverken bestemmelserne om kopiering til privat brug (OHL § 12), til undervisningsbrug (OHL § 13) eller til biblioteksbrug (OHL § 16) giver ret til at udnytte materialet på den påtænkte måde, uden at ophavsmændene (eller dem de har overdraget de relevante rettigheder til) spørges først, ligesom man normalt må regne med at skulle betale et vederlag i forbindelse med anvendelsen.

Tilladelsen skal gives af den, der aktuelt er indehaver af rettighederne. Ved udgivne dokumenter kan dette være såvel de oprindelige ophavsmænd som udgiveren, f.eks. et forlag - afhængigt af indholdet af udgivelsesaftalen. Praksis varierer meget, så man kan ikke på forhånd vide, hvordan det forholder sig, ligesom parterne i mange tilfælde kan være uenige om, hvorvidt der overhovedet foreligger en aftale om sekundære udnyttelser som denne, og hvad aftalen i givet fald indebærer.

Det enkelte dokument kan desuden have adskillige ophavsmænd, udover en eller flere forfattere, typisk ophavsmændene til illustrationer og lignende. Det er nødvendigt at sikre, at alle relevante rettigheder er "klarere" ("cleare") med de relevante rettighedshavere, og at der foreligger klare aftaler om f.eks. betaling.

Normalt skal man også sikre rettigheder til at gengive illustrationer (diagrammer, tegninger, fotografier m.v.) i forbindelse med de indlæste værker, men i nogle tilfælde, nemlig hvor illustrationerne fungerer som "billedciter", kan dette formentlig undlades, når der er tale om kritiske eller videnskabelige fremstillinger. Ifølge OHL § 23 kan kunsværker (herunder fotografiske værker) og beskrivende billeder gengives uden særlig tilladelse og uden vederlag i kritiske eller videnskabelige fremstillinger i tilslutning til teksten. Dette har biblioteket formodet også gælder for en digital reproduktion af den pågældende fremstilling.

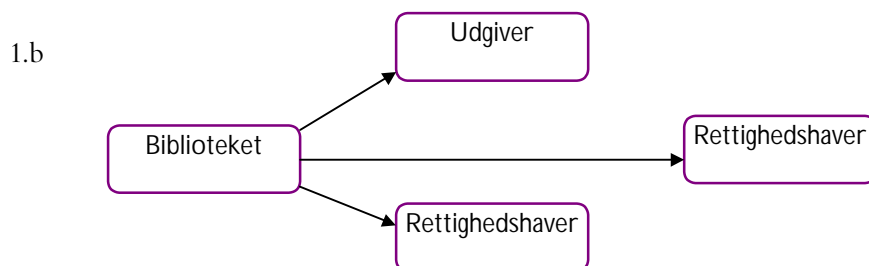
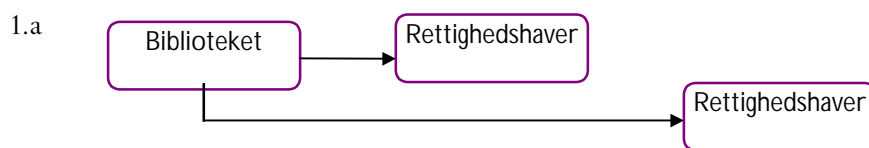
Ophavsretsloven gør det muligt at håndtere rettighedsklareringen gennem såkaldte aftalelicenser, som f.eks. CopyDan-aftalen om fotokopiering til undervisningsbrug. Det indebærer, at man kun behøver at henvende sig ét sted for at opnå de nødvendige tilladelser, ligesom en sådan kollektiv forvaltningsorganisation kan sørge for betalingen til rettighedshaverne. Der har imidlertid ikke hidtil været mulighed for at træffe kollektive aftaler med aftalelicens-virkning på det digitale område.

Men med ændringen af Ophavsretloven i juni 1998 blev der åbnet mulighed for at der, også på det digitale område, kan træffes kollektive aftaler med aftalelicens-virkning for kopier til undervisningsbrug (OHL § 13). Det forudsætter dog, at en forvaltningsorganisation, evt. CopyDan, bemyndiges af sine medlemsorganisationer til at træffe sådanne aftaler, hvilket hidtil ikke har været tilfældet. Det vil derfor indtil videre fortsat være nødvendigt at finde frem til de relevante rettighedshavere for hvert enkelte dokument der ønskes klareret og at forhandle med dem om mulighederne for, og betingelserne i forbindelse med, en udnyttelse af deres værker i et elektroniske dokumentarkiv som det foreliggende.

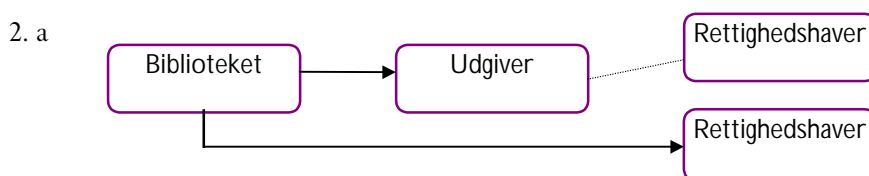
3.2.2. Arbejdet med rettighedsklarering

Med udgangspunkt i den danske lov om ophavsret, behandlet i forgående afsnit, kan de materialer, der ønskes rettighedsklareret og indlagt i arkivet, inddrages i de efterfølgende grupper. De skematiske illustrationers fuldt optrukne pile angiver, at Rub har måtte forhandle direkte med den pågældende part, mens de stiplede pile angiver, hvem der indirekte har indgået i forhandlingerne gennem underskriveren.

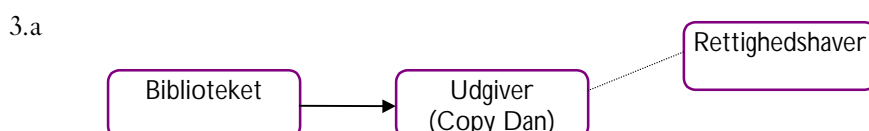
1. Materialer, hvor de relevante rettigheder tilhører de enkelte ophavsmænd, enten fordi disse selv har stået for udgivelsen, eller fordi rettighederne til sekundære udnyttelser som den foreliggende ikke er overdraget. Eneretskonstruktionen medfører at samtlige ophavsmænd skal kontaktes, herunder ophavsmænd til eventuelle illustrationer. (Som nævnt ovenfor kan der være tale om undtagelser fra dette i forbindelse med "billedciter").



2. Materialer, hvor nogle af ophavsmændene har overdraget de relevante rettigheder til en tredjepart, normalt udgiveren, mens andre ikke har gjort det. Her skal både udgiveren og en række enkeltpersoner kontaktes. Det er ofte ophavsmænd til illustrationer, der giver denne slags problemer.



3. Materialer, hvor alle de relevante rettigheder er overdraget til en tredjepart, typisk et forlag, som kan indgå aftaler på ophavsmændenes vegne og varetage administrationen af vederlag. Dette betyder, at der i forbindelse til rettighedsklareringen udelukkende skal søges om tilladelse ét sted. På områder, hvor der indgået aftale på aftalelicensvilkår, er det i stedet en forvaltningsorganisation som CopyDan, der indtræder i denne rolle.



Arbejdet med rettighedsklarering er foregået i to perioder. I første periode fra d. 20.10.1997 og frem til d. 15.1.1998 blev der særlig lagt vægt på titler udgivet af uddannelsesinstitutioner, akademiske foreninger, organisationer m.m. samt forlag knyttet til forskellige akademiske sammenslutninger. I anden periode fra d. 15.1.1998 til d. 1.11.1998 blev der sat særlig fokus på materiale udgivet af danske forlag, både universitetsforlag og andre kommercielle forlag.

Der blev i løbet af første periode taget kontakt til 135 rettighedshavere vedr. 87 forskellige titler. Af disse er der foreløbigt givet tilladelse til, at 57 titler kan anvendes til arkivet. De resterende 30 kan ikke indlægges, enten fordi én eller flere af de pågældende rettighedshavere ikke har ønsket det pågældende materiale anvendt i forbindelse med Det elektroniske forsknings- og undervisningsarkiv, eller fordi en brugsaftale endnu ikke ved rapportens afslutning var forhandlet færdige.

I anden periode blev 27 forskellige danske forlag kontaktet vedr. i alt 70 forskellige titler. På nuværende tidspunkt er der blevet givet tilladelse til, at 21 af disse kan indlægges i arkivet, mens de resterende 49 titler, af grunde beskrevet ovenfor, stadig ikke kan anvendes.

Listen over forlag, forskningsinstitutter og -sammenslutninger m.m., der i forbindelse med projektet har modtaget en forespørgsel fra Rub om anvendelsen af materialer i Det elektroniske forsknings- og undervisningsarkiv findes i bilag 4.

Til at administrere de mange forskellige informationer i forbindelse med rettighedsforhandlingerne, blev der, som ét af de første skridt i projektet, specialkonstrueret en database baseret på Access. Databasen gør det muligt at registrere de forskellige titler, deres bibliografiske data; informationer om relevante personer og udgivere, korrespondance, aftaleinformationer m.m. (se bilag 9)

3.2.3. Aftalemodeller

Der findes endnu ikke nogen (dansk) praksis for, hvordan aftaler vedr. anvendelse af publikationer i et elektronisk dokumentarkiv skal indgås, herunder hvordan den økonomiske kompensation for anvendelsen fastsættes. Aftalebetingelserne har derfor skulle afklares i hver enkel forhandlingssituation, uanset om der har været tale om forhandlinger med enkelte forskere, tidsskriftredaktioner, institutioner eller deciderede forlag. På baggrund af de forhandlinger, Roskilde Universitetsbibliotek på nuværende tidspunkt har indgået med rettighedshavere, kan der skelnes mellem 5 forskellige typer af aftaler beskrevet i det efterfølgende.

Aftaletype 1: Fri anvendelse

Særlig i første periode af rettighedsklaringen, hvor der blev fokuseret på titler udgivet af uddannelsesinstitutioner, organisationer eller lign., indvilligede en stor del af rettighedshaverne i at lade deres materialer anvende i projektperioden uden betaling af vederlag og uden nogen form for udprintningsbegrænsning. Dette skyldtes bl.a., at en del rettighedshavere havde udarbejdet materialet i forbindelse med deres forskningsarbejde og derfor anså deres produktion som betalt via deres løn¹.

Aftaletype 2: Model Copy-Dan

En del rettighedshavere har givet tilladelse til anvendelse af deres materialer på betingelser, der ligger på linie med den kopieringsaftale, som CopyDan forvalter (nævnt tidligere), hvor der betales en afgift pr. udprintet side, der svarer til CopyDan-afgiften for kopierede sider. Denne opgøres i øjeblikket til 19.4 øre pr. side. Fordelen ved denne betalingsmodel er, at betalingen af vederlag til rettighedshaverne udelukkende gøres afhængig af den konkrete udprintning.

Nogle af disse rettighedshavere har desuden ønsket at fastholde den sidemæssige begrænsning, der er en del af CopyDan-aftalens konstruktion. Andre har fundet det tilstrækkeligt, at der finder en registrering og efterfølgende afregning sted pr. udprintet side og har ikke ment det nødvendigt at fastholde den sidemæssige begrænsning.

Aftaletype 3: Betaling pr. udprint

En anden stor gruppe af rettighedshavere har ligeledes ønsket betaling pr. udprintet side, men har angivet en anden takst end CopyDan-afgiften. Disse takster varierer fra 21.25 øre og til 100.0 øre plus moms og oplyses i flere tilfælde at være fastsat ud fra en vurdering af "tab pr. ikke - solgt side" og afhænger derfor af faktorer som f.eks. publikationens udsalgspris, dens produktionspris, aktualitet, alder, tilbageværende oplag m.m. Et enkelt forlag har dog ønsket at sidestille udprintning fra DEFUA med kopiering uden for CopyDan aftalen og har, som følge heraf, peget på en afgift på 100.0 øre plus moms pr. side udprintet.

Aftaletype 4: Betaling for digitalisering samt udprintning

I modsætning til de to forgående betalingsmodeller, hvor vederlag er gjort afhæng af den konkrete udprintning, har en større gruppe af forlag (alle er medlemmer af den nystiftede universitetsforlagsforening, se bilag 4) gjort tilladelsen til digitalisering og efterfølgende udprintning i projektets løbetid, afhængig af at de betales et engangsbeløb på 18.00 kr. excl. moms pr. 1.000 enheder digitaliseret og efterfølgende 18.0 øre pr. side, der udprintes fra de pågældende tekster.

Denne model og det konkrete krav om økonomisk kompensation betyder imidlertid, at det er forbundet med store omkostninger blot at etablere arkivet. For indskanning af bare en enkelt udgivelse (150 sider med

¹ For uddybende behandling af ophavsretlige forhold i forbindelse med forskningspublikationer, forfattet af videnskabelige medarbejdere tilknyttet de højere læreanstalter og finansieret af offentlige bevillinger, læs evt. "Elektronisk informationsformidling og forskningspublicering", Handelshøjskolen i Aarhus, 1998.

ca. 2100 enheder pr. side), skal der alene for indskanningen betales ca. 5670 kr. til rettighedshaverne, endnu inden der er udprintet fra publikationen.

Aftaletype 5: Engangsbetaling

En lille gruppe af rettighedshavere har lagt op til, at der skal betales et éngangsbeløb for siden ubegrænset anvendelse af deres værker. Dette er sket ud fra en vurdering af, at det vil være lettest for dem selv, såvel som for biblioteket, og de konkrete beløb angives af rettighedshaverne til at være fastsat ud fra en vurdering af, i hvor stort omfang det pågældende materiale vil blive anvendt, og på hvilket niveau en éngangsbetaling derfor med rimelighed bør lægges.

3.2.4. Erfaringer/konklusioner i forbindelse med rettighedsklarering

Som nævnt, er det en forudsætning for anvendelsen af et værk i Det elektronisk forskning- og undervisningsarkiv, at biblioteket identificerer de aktuelle rettighedshavere og forhandlet brugsrettigheder med dem hver især. Klareringsprocessen har derfor vist sig at være en utrolig tidskrævende proces.

Langt de fleste rettighedshavere ansat ved forskningsinstitutioner eller -institutter, har vist sig positive i forhold til anvendelsen af deres publikationer i DEFUA. Dette skyldes især, at distributionsformens fordele i forhold til studerendes adgang til undervisningsmaterialer har vakt stor interesse i forskningsmiljøerne generelt. De, der ikke ønskede at medvirke, begrundede det oftest med, at publikationen enten havde været udgivet ved et forlag, eller at man håbede på en senere forlagsudgivelse. Forlagene har derimod generelt været mere tilbageholdende. Dette afspejles bl.a. i bibliotekets vanskeligheder ved at få udgiverne til at indgå bredere og mere generelle aftaler om anvendelse af et større antal af deres udgivelser. Udgiverne har i stedet insisteret på, at der skal ske en løbende rettighedsklarering af de enkelte værker og at de konkrete betingelser skal aftales fra titel til titel.

I forbindelse med klarering af illustrationer i videnskabelige værker udgivet ved forlag, har der vist sig særlige problemer. Selvom ikke alle illustrationer i videnskabelige værker kræver rettighedsklarering, har det ofte vist sig vanskeligt at fastholde retten til gengivelse af "billedciter" uden særlig tilladelse. Flere udgivere har eksplicit gjort Rub's anvendelse af et værk afhængig af, at der ligeledes sker en rettighedsklarering af værkets illustrationer. Det kan i så tilfælde være problematisk, hvis biblioteket alligevel vælger at indlægge værket (med illustrationer) uden yderligere rettighedsklarering, ligesom forlagene, med henvisning til kvalitetsforringelse, vil modsætte sig en udeladelse af de pågældende illustrationer.

Det må imidlertid formodes, at processen med rettighedsklarering tidsmæssigt vil forkortes, efterhånden som rettighedshaverne, herunder forlagene, bliver mere afklaret i forhold til spørgsmålet om elektronisk kopiering. Langt den største del af de adspurgte forlag har endnu ikke en egentlig politik på området og har derfor fundet det nødvendigt, at få diskuteret sagen på højeste niveau før der kunne afgives et endelig svar på bibliotekets forespørgsel.

Udgiverne har sjældent på forhånd de fornødne rettigheder til at kunne tillade denne type sekundære udnyttelse af værkerne. En tilladelse i forhold til DEFUA forudsætter derfor, at udgiverne først selv får indhentet de nødvendige rettigheder fra samtlige ophavsmænd. I en del tilfælde er der blevet gjort opmærksomt på, at man ikke har de nødvendige ressourcer til dette arbejde. Da sagen i øvrigt betragtes som af stor principiel betydning, har flere udgivere følt sig nødsaget til at afvente udfaldet af drøftelserne om digitale rettigheder, som foregår bl.a. i Dansk Forelæggerforening, Dansk Journalistforbund m.v.

Det har vist sig vanskeligt at fastholde en homogenitet i de aftaler, det har været muligt at indgå med rettighedshaverne. Spørgsmålet om hvordan den økonomiske kompensation for den pågældende anvendelse afregnes, har vist sig at dele rettighedshaverne, ligesom ikke alle er enige om, hvilke udnyttelsesformer der kan tillades. Disse aftalemomenter kan variere dels mellem de forskellige værker, dels de enkelte værkers rettighedshavere imellem. Dette kan imidlertid indebære, at der skal udbetales forskelligt vederlag til et værks forskellige rettighedshavere, ligesom den mest restriktive brugstilladelse, indgået i forbindelse med et værk, sætter grænsen for, hvad værket må anvendes til. Afvises anvendelsen fuldstændig af blot én enkelt rettighedshaver, må dokumentet opgives. Det kan som følge heraf overvejes, om et alternativ kunne være, at sætte fastere rammer for deltagelse i projektet, sådan at brugstilladelsen ikke kan gradueres.

I følge ophavsretsloven overdrages udelukkende retten til de udnyttelsesformer, der er specificeret eksplicit i aftalen mellem parterne. Dette gør det særlig vigtigt at det konkretiseres i kontrakten, hvilke udnyttelsesformer der ønskes rettigheder til, så der hverken overdrages for mange rettigheder, med forøgede klareringsomkostninger til følge, eller for få, så teksten ikke kan anvendes efter hensigten. På baggrund af projekterfaringerne, kan der peges på nogle retningslinier til udformning af en egentlig kontrakt vedr. elektronisk kopiering. Disse retningslinier er vedlagt som bilag (bilag 5)

Rub kan som følge af de erfaringer, arbejdet med rettighedsklarering i forbindelse med DEFUA har bragt med sig, særlig pege på behovet for, at der, evt. i DEF-regi, for det første arbejdes videre med formulering af en egentlig standardaftale på området. For det andet, at der på længere sigt arbejdes i retning af en centralisering af rettighedsklarering i forbindelse med digital kopiering inden for undervisningsområdet. En central løsning vil fjerne byrden forbundet med klareringen fra brugerne og dermed fremme mulighederne for, at der kan ske en udvidet distribuering af undervisningsmateriale via elektroniske netværk.

3.3. Digitalisering af dokumenter

Ved valg af system til indskanning af dokumenter, er der blevet lagt vægt på følgende faktorer:

- Skanningskvalitet og -hastighed.
- Dokumentstrukturering.
- Dokumentrensning.
- Billedbehandling.
- Fleksibilitet mht. lagringsformater.
- Mulighed for effektiv printkontrol af hensyn til rettighedshavere.
- Integration med Roskilde Universitetetscenters trykkerisystem.
- Integration med Rub's almindelige bibliotekssystem RUBikon.

Ønsket om en sammenhæng med trykkerifunktionen samt sikkerhed for effektiv printkontrol har været afgørende for vores valg af samarbejde med Xerox om en afprøvning af programmet XDOD 4.1.(Xerox-Document On Demand). Systemet var på dette tidspunkt (1997) i brug ved Københavns og Oslos universitet som produktionssoftware ved kompendiefremstilling. Aftalen med Xerox har indebåret en anskaffelse af indskanningsudstyr og -software, samt en gratis prøveinstallation af softwarepakken DOCUWEB til web-baseret distribution på nettet. (se pkt. 3.4 samt beskrivelserne i bilag 10-11)

Arbejdsgangen for dokumenthåndteringen har indebåret en række operationer:

- Indskanning og editering
- Dokumentstrukturering
- Dokumentregistrering
- Dokumentlagring og import-/eksportformater

3.3.1. Indskanning og editering

Skanningshastigheden er 15 - 20 enkeltsider pr. minut (600 dpi, A4), og der kan indskannes fra begge sider af papiret. Der er fleksible faciliteter til skalering af indskannede sider. Der er automatiske funktioner til oprettelse af skævheder i siderne, samt faciliteter til rensning af evt. urenheder stammende fra originalen. Softwaret indeholder avancerede faciliteter til særlig behandling af f.eks. billeder og illustrationer med henblik på en høj billedkvalitet, men håndterer ikke farver.

Bortset fra, at farvehåndtering må anses for et klart krav til et fremtidigt system, har vi vurderet, at skannings- og editeringsfaciliteter i XDOD er af høj kvalitet og helt tilfredsstillende. Softwaret er nemt at arbejde med, og der kan opnås en høj kvalitet i forhold til arbejdsindsatsen.

3.3.2. Dokumentstrukturering

Det indskannede dokument kan struktureres hierarkisk i kapitler eller afsnit, afhængigt af det trykte dokumentets struktur. Opdelingen er relevant i forbindelse med distribution af dokumentet, hvor brugeren, afhængigt af den indgåede brugsaftale i forbindelse med dokumentet, kan rekvirere print af dele af dokumentet. Faciliteterne i XDOD til dokumentstrukturering vurderes som tilfredsstillende til projektets formål.

3.3.3. Dokumentregistrering

Det har været et højt prioriteret krav, at der skal etableres en forbindelse mellem digitaliseringssystemet og Roskilde Universitetsbiblioteks katalog Rubikon, så dokumenternes bibliografiske data er søgbare i katalogen, og at der kan etableres en link fra katalogen til det digitaliserede dokument. Forbindelsen tænkes etableret mellem dokumentserveren og bibliotekskatalogen.

I XDOD indgår en mulighed for brugerdefinerede registreringsfelter, hvis data er søgbare i XDODs SQL database og i Xerox dokumenthåndteringssystem Docuweb (se afsnit 3.4). Der er i forbindelse med projektet defineret et antal dataelementer til registrering af dokumenterne, som svarer til den ønskede katalogpost, og som vil kunne anvendes som grundlag for dannelse af metadata for dokumentet i forbindelse med web-publicering.

Registreringsdata tilføjet i XDOD overføres imidlertid ikke ved konvertering til pdf-format, og de indtastede informationer vil derfor ikke være søgbare i forbindelse med andre dokumenthåndteringssystemer end Xerox' eget. På denne baggrund må det konkluderes, at XDODs faciliteter på dette punkt er mangelfulde. Det må betragtes som et krav til et editeringssystem, at det kan formatere metadata i overensstemmelse med såvel Dublin Core standarden som et valgfrit marcformat. Faciliteter til en selektiv OCR på f.eks. indholdsfortegnelser og abstract burde også kunne indgå som en del af registreringen.

3.3.4. Dokumentlagring og import/eksportformater

Lagring i XDOD systemet efter endt editering foregår i TIFF(RDO), som er et format der tillader høj hastighed i print fra Xerox' trykkerisystem Docutech. Herudover lagres dokumentet i PDF med henblik på eksport til dokumentserver. XDOD kan kun importere dokumenter i TIFF eller PS formater, og er derfor ikke velegnet som et generelt system til konvertering og dokumentbehandling forud for den elektroniske publicering. Da PDF inden for de seneste år har fået karakter af de facto standard, må det skønnes at et nyere program som Adobes Acrobat Capture 2.01 i langt højere grad vil være relevant for det videre digitaliseringsarbejde. Ud over, at dette program efter vores foreløbige skøn stiller nogle faciliteter til rådighed, som kan afhjælpe XDODs mangelfulde behandling af PDF formatet, er der også tale om et billigere produkt som indbyder til, at også små biblioteker kan deltage i digitaliseringen af ældre materiale med henblik på elektronisk publicering. Specielt mht. valg og håndtering af lagrings- og importformater er det vores vurdering at XDOX 4.1 ikke er tilfredsstillende, og at der nu er produkter på markedet, som er mere velegnede, når målet er netpublicering og ikke kun kompendieproduktion.

3.4. Tilgængeliggørelse af digitaliserede dokumenter

Det har været et selvstændigt mål med projektet at undersøge markedet for softwareløsninger, der kunne håndtere en web-baseret netpublicering af materiale omfattet af copyright, hvor afregning af betaling til rettighedshavere sker efter forbrug.

Krumtapperne i afklaringen af de specifikke krav har været

- De forskellige brugergrupperes behov for adgang til at browse, søge, printe, samt bestille arkivets tekster.
- Rettighedshavernes (herunder forlagenes) krav til sikkerheden omkring adgangsbegrænsning og registrering af forbrug.
- Bibliotekets krav til størst mulig transparens i adgangen til dets service, samt til integrationsmulighederne med institutionens og biblioteksområdets øvrige systemer.

På basis af disse afvejsninger er der blevet udarbejdet en kravspecifikation (se bilag 16), som har været udgangspunkt for afsøgning af markedet. Dette har resulteret i en nærmere undersøgelse af i alt 5 forskellige produkter, som kort karakteriseres i det følgende:

3.4.1. Docuweb fra Xerox

I forbindelse med anskaffelsen af XDOD arbejdsstationen i januar 1998 indgik Roskilde Universitetsbibliotek en aftale med Xerox om afprøvning af programmet Docuweb. Docuweb er et NT 4.0 baseret system, som Xerox har udviklet som en extension af XDOD med henblik på netpublicering.

Programmet gør det muligt:

- At opbygge digitale kollektioner af dokumenter til forskellige formål og med forskellige typer af tilknyttede rettigheder og adgang.
- At søge dokumenter på basis af registreringsdelen af XDOD.
- At vise dokumenter som thumbnails og/eller i jpg-billeder.
- At give mulighed for downloadning af PDF versioner af dokumenter med fri adgang.
- At bestille print af det samlede dokument eller dele deraf.

Systemet var efter det oplyste i brug hos National Academy Press i USA, et forlag som står for elektronisk publicering og print-on-demand af alle statsligt finansierede amerikanske forskningsrapporter.

Kort evaluering:

Systemet er ikke umiddelbart brugbart til DEFUAs formål. Blandt manglerne er bl.a. mangelfulde faciliteter til administration af brugsaftaler, forbrugsregistrering og afregning i forhold til aftalerne med rettighedshaverne. Faciliteter til printstyring er utilstrækkelige, f.eks. til håndtering af valg mellem printmuligheder (lokalprint, print på trykkeri). Systemet kan ikke håndtere brugerregistrering på to niveauer, så den enkelte bruger kan tilknyttes en brugergruppe eller en konto. Konventionen for navngivning af dokumenterne er dårligt egnet til linkning fra bibliotekskatalogen og systemet kan ikke levere filer til eksport af katalogoplysninger til bibliotekets katalog.

Samlet må man vurdere, at Docuweb i den version RUb har afprøvet (version 2.0), ikke egner sig til en distributionsform, hvor dokumenter med forskellige adgangsbegrænsninger skal stilles til rådighed for

slutbrugere, der efterfølgende skal debiteres efter forbrug. Programmet er i sin nuværende form udviklet til at sikre netadgang til XDOD-systemets dokumenter med forbindelse til trykkerisystemet Docutech, men kan ikke umiddelbart integreres med andre systemer.

3.4.2. New Publiotech fra Electric Press

I forbindelse med undersøgelsen af Docuweb blev det konstateret, at der var udviklet et alternativt produkt ved et samarbejde mellem bl.a. Carnegie Mellon University og National Academy Press. Produktet var efter afsluttet udvikling på markedet under navnet Publiotech (producent: Electric Press)

Dette produkt var udviklet til integration med XDOD, men indeholdt en række moduler som ikke var udviklet i Docuweb, f.eks. faciliteter til betalingsformidling, fleksible muligheder for forskellige prissætninger af dokumenter, bestilling af dele af dokumenter, etc. Vi vurderede, på basis af kontakter med Electric Press, samt en gennemgang af detaljerede produktbeskrivelser og undersøgelser af sites hvor produktet var i drift, at Publiotech ville være et interessant produkt til distribution af digitale dokumenter omfattet af betalingsaftale.

De efterfølgende forhandlinger om muligheder for anskaffelse eller subsidiært leje af systemet endte dog uden resultater. Det blev klart, at produktet var beregnet for et kommercielt forlagsmarked, dvs. egentlige kommercielle forlag, og producenten var ikke aktuelt interesseret i det institutionelle marked. Efterfølgende annoncerede Electric press en interesse for "academic licensing", og i denne forbindelse forsøgte vi at bringe en repræsentant til landet for at demonstrere systemet for interesserede biblioteker. Kontakten blev afsluttet, da firmaet kort efter aflyste arrangementet.

Kort evaluering

Systemet er internationalt udbredt som softwarepakke for elektroniske forlag med print-on-demand produkter, og virker meget gennearbejdet til sit formål. Det vurderes at systemet er for dyrt til have en bred interesse til biblioteks- eller universitetsformål, med mindre en form for national licensing, eller samarbejde mellem flere universiteter, kan komme på tale.

3.4.3. DOC fra DDE

DDE, som er leverandør af RUB's integrerede bibliotekssystem, har i januar 1999 startet betatest på et dokumentserverssystem, som bl.a. Rub skal teste. Den version der er under afestning, indeholder kun generelle serverfaciliteter, såsom søgning, scanning visning samt print fra browser. Herudover indeholder den en række af de ønskede faciliteter mht. til systemintegration i form af produktion af automatiske katalogposter til bibliotekssystemet på basis af fuldtekstdokumenternes metadata. Versionen indeholder ikke faciliteter til en differentieret forbrugsregistrering og -afregning, ligesom den heller ikke indeholder det ønskede adgangskontrolsystem.

Biblioteket forhandler i øjeblikket om en kontrakt på udvikling af sådanne faciliteter, med udgangspunkt i et ønske om dataintegration mellem dokumentserverens brugerstruktur og det integrerede bibliotekssystems lånerregister. Tidshorizonten for udvikling af de ønskede udvidelser er fastlagt til ca. medio 1999. Det skal understreges, at systemudviklingen fra DDEs side ikke specifikt sigter på

integrationsmuligheder med DDE-bibliotekssystemet. Dokumentserverssystemet er et selvstændigt system, som vil kunne opnå samme type af integration (lånerregisterkoordinering og eksport af katalogposter) med andre bibliotekssystemer.

Kort evaluering

Fra Rub's synsvinkel er perspektivet omkring DDE's DOC ikke mindst interessant, fordi det umiddelbart vil tillade en højere grad af systemintegration med bibliotekets integrerede bibliotekssystem, og fordi en kontrakt muliggør en påvirkning af indhold og udviklingsprocess. Ulempen ved valg af løsningen er, at det ikke foreligger som et brugbart og færdigt system til løsning af DEFUA opgaven. Det forventes, at et kontraktgrundlag vil foreligge i slutningen af januar 1999.

3.4.4. ERes fra DOCUTEK

DEFUA projektet er beslægtet med en række systemer, der i USA og UK betegnes som "electronic reserve systems", dvs. systemer der håndterer administrationen af elektroniske "lån" af efterspurgt kursusmateriale.

Der findes en række sådanne systemer, hvoraf Rub har udvalgt et til nærmere undersøgelse i form af prøveadgang til systemet i efteråret 1998. Det drejer sig om ERes 3.0, som er udviklet af firmaet Docutec Information Systems (Docutec), og som er udbredt på et stort antal store amerikanske universiteter. En del af baggrunden for produktets succes er sandsynligvis dets brugervenlighed. ERes er et enkelt web-baseret administrationssystem for organisering og distribution af dokumentamlinger eller kursuspakker fra fælles NT eller Unix dokumentserver.

Det tillader håndtering af en række gængse formater på enkel vis, opererer med IP- eller passwordstyret adgangskontrol på dokumentamlingsniveau, og åbner mulighed for lukket chat-funktion inden for brugergruppen. Der er dog ikke tale om et egentligt dokumentserverssystem med f.eks. søge-, scan- og visningsmuligheder. Systemet indeholder heller ikke i den nuværende version muligheder for afgiftsberegning til rettighedshavere, ligesom der heller ikke er mulighed for at identificere en individuel bruger af systemet. Producenten har tidligere oplyst (og senest i januar 1999 bekræftet), at systemet vil blive frigivet i en release 4.0 senest juni 1999, hvor dokumentbestillings- og betalingsfaciliteter vil være integreret.

Kort evaluering

I sin nuværende version ser vi programmet som en mulig "front-end" til en kommende dokumentserver. Hvad der især synes interessant er, at ERes giver mulighed for at undervisere på kommunikationsuddannelsen selv, på en simpel måde, kan udarbejde egne kursuslister direkte on-line med mulighed for evt. tilføjelse af eget materiale som fuldtekst. På den måde kan ERes muligvis være med til at reducere den ellers til tider ganske langsommelige proces, fra en underviser sætter en relevant tekst på sin kursuslitteraturliste, og til teksten forefindes som elektronisk dokument tilgængeligt for de studerende. Programmet er anskaffet til en begrænset brugergruppe på RUC, hvor det vil blive afprøvet i praksis af nogle få undervisere på RUCs kommunikationsuddannelse.

3.4.5. Digital Library fra IBM

IBM har igennem længere tid markedsført produktet Digital Library, som er et fuldt udviklet fuldtekst dokumenthåndteringssystem, som er i drift på et stort antal universiteter i USA. Systemet er ikke i brug i Danmark, men f.eks. i Essen i Tyskland. Systemet kan efter det oplyste leveres med de forbrugsregistrerings- afregnings- og adgangskontrolfunktioner som er specificeret i kravene til DEFUA. Der har været afholdt en demonstration af systemet på Rub, hvor systemkravene blev gennemgået. På basis heraf er der efterfølgende indhentet et tilbud på leverance.

Systemet er opbygget omkring et kernesoftware, som er selve dokumentserverens søgesystem, samt en række værktøjer til opbygning af det individuelle system hos den enkelte kunde. Dette afspejles i prisstrukturen, hvor softwareleverancen udgør en mindre del af den samlede udgift, set i relation til udgifter til tilretninger af systemet.

Kort evaluering

IBMs strategi sigter på den individuelle tilpasning af et generelt værktøj. Der er i vid udstrækning tale om en individuel applikationsudvikling, hvor det ønskede system, i lighed med flere af de øvrige ovenfor, er en udviklingsopgave, firmaet tilbyder at påtage sig, snarere end et færdigudviklet produkt. Det er dog godt gjort, at faciliteter svarende til projektets krav om adgangskontrol og afregning overfor rettighedshavere er implementeret på andre IBM installationer af produktet.

Bibliotekets kontakt med IBM om det foreliggende tilbud er ikke afsluttet, men det er projektstyrergruppens opfattelse, at alene prisniveauet vil forhindre en større udbredelse i det danske biblioteksvæsen.

3.4.6. Konkluderende bemærkninger

Eftersøgningen af software, der tilfredsstillende opstillede kravspecifikation, har været lang og har endnu ikke ført til en færdig installation af en dokumentserver. Projektstyrergruppen har valgt indtil videre se bort fra de to løsninger, som tilsyneladende kunne håndtere systemkravene, men hvis udgiftsniveau vurderedes som for højt, nemlig Publiotech og IBM's Digital Library. I stedet er det besluttet at deltage i betatesten af DDE DOC og satse på en udvikling af fase to af dette projekt, før en endelig beslutning om anskaffelse af dokumentserverssystem træffes i midten af 1999. I mellemtiden vil der blive foranstaltet en afprøvning af ERes systemets version 3 på brugerniveau, og version 4.0, så snart den foreligger.

4. Projektets modelkarakter

Projektet har taget udgangspunkt i et generelt problem for de fleste forskningsbiblioteker: at kunne stille central studielitteratur til rådighed uden begrænsninger i form af antal trykte eksemplarer - et behov som har været stigende gennem mange år. Projektet har sigtet på at belyse en række generelle problemstillinger i forbindelse med dokumentudvælgelse, dokumenthåndtering, forhandling omkring copyright m.v. og dets hovedkonklusioner vil være umiddelbart anvendelige for andre fagområder end det aktuelle.

Teknologisk set har projektet arbejdet med problemløsninger, der i videst mulig grad har taget udgangspunkt i allerede kendt og anvendt teknologi samt på softwaresiden anvendelse af standardprogrammer.

En væsentlig del af projektet har været helliget forskellige problemstillinger i forbindelse med klarering af rettigheder. I den forbindelse har man belyst ophavsretsmæssige problemstillinger af generel karakter, samt udviklet arbejdsmetoder og retningslinier, der vil være anvendelige i lignende projekter. Projektet peger på en række initiativer i forhold til en effektivisering af klareringsprocessen, hvis gennemførelse vil indebære en forenkling af proceduren også i forbindelse med andre DEF-projekter.

5. Erfaringsudveksling

I løbet af projektperioden har der været skabt flere informelle fora for kommentering og erfaringsudveksling. Der har været kontakt til Ålborg Universitetsbibliotek, Handelshøjskolens Bibliotek i Aarhus, Kompendieudsalget ved KUA, samt Kompendie- og noteudsalget ved Oslo Universitet.

Repræsentanter fra projektgruppen har både løbende og ved div. netværksmøder videregivet erfaringer vedrørende tekniske og juridiske spørgsmål i forbindelse med internetdistribueret materiale. Biblioteket har desuden orienteret omverdenen via sin hjemmeside, ligesom der har været skrevet i det lokale organ, RUC-Nyt. Der har været orienteret på kommunikationsuddannelsen dels via den interne elektroniske nyhedskanal, dels på interne møder for ansatte ved instituttet.

Endelig har projektet været præsenteret for en bredere kreds ved et fagligt seminar, hvor samtlige pilotprojekter i forbindelse med DEF var deltagere (26.11.1997). Ved et lignende seminar i april 1999 vil der blive givet den seneste statusrapport.

6. Videreførelse af projektet

Projektet forsætter foreløbig frem til oktober 1999, hvor man håber at kunne iværksætte en egentlig driftsfase. Frem til dette tidspunkt vil der blive arbejdet med betatest af DDE's dokumentserverssystem samt afprøvning af ERes softwarepakken. Man håber i den forbindelse at få mulighed for afprøvning af et

adgangskontrol- og afregningssystem, som rettighedshaverne vil finde tilfredsstillende. Sideløbende hermed vil projektledelsen, i samarbejde med BS/DEF, forsøge at nå frem til et aftalekompleks med rettighedshaverne, evt. på midlertidig forsøgsbasis.