

Dokumentutveksling på internett

– **Adobe PDF, et åpent OG proprietært format**

Internett og Makt: Semesteroppgave for kandidat 154796

Institutt for medievitenskap

Universitetet i Bergen

Våren 2003

Innholdsfortegnelse

1 Introduksjon og avgrensning	3
2 Åpne formater	4
2.1 Eksempler på åpne og lukkede formater	4
2.2 «Open Source Software» og «Free Software»	5
2.3 Bruk og misbruk av begrepet «åpen kildekode»	5
2.3 DMCA	7
3 Proprietære formater	8
4 Dokumentutveksling med PDF-formatet	9
4.1 Portabilitet	10
4.2 Båndbredde og kompresjon	10
4.3 Sikkerhet	11
4.4 PDF – åpent og proprietært	12
4.5 Driver Adobe veltedighet?	12
4.6 En «vinn-vinn»-situasjon?	13
4.7 PDF som offer for begrepsbruk	13
6 Oppsummering	14
7 Referanser	16

1 Introduksjon og avgrensing

Dokumentutveksling på internett er noe man i informasjonssamfunnet kan regne som trivielt, portabilitet angår de fleste som skal lese eller produsere et elektronisk dokument. Adobe Portable Document Format (PDF) har i løpet av de siste ti årene blitt et dokumentformat som ofte brukes til dette formål. Dette formatet er «åpent» og «proprietært», noe som kan gi litt ambivalente følelser, da disse begrepene ofte brukes ukorrekt som antonymer. Antonymet til «åpent» er «lukket», og antonymet til «proprietært» er «ikke-proprietært».

Like misbrukt som det feilaktige antonymparet brukes begrepet «åpen kildekode (eng. Open Source)». Den strategiske kampen rundt uttrykket «Open Source» har tatt av like etter at begrepet ble tatt i bruk (først i 1998), kanskje først og fremst fordi det ble knyttet til positive assosiasjoner og suksessbedrifter som Linux (fritt operativsystem) og World Wide Web. Det er dokumentformatene, ikke programvaren, som er tema for oppgaven – men åpne formater har mye å gjøre med åpen programvare.

Åpne formater har for alvor kommet for å bli, med internett som overlegent distribusjonsnett og utviklingsforum. Mange av disse formatene brukes nettopp på internett. Når det skrives om åpne formater i et maktperspektiv, blir de ofte glorifisert overfor proprietære formater. Dette med sin fulle rett, hvis det da menes lukkede proprietære formater. Skjebnen til lukkede proprietære formater er totalt overgitt de som sitter med eierskapet, og de som benytter dem utøver mikromakt (Slaatta (red.) 2002:44) overfor mottakerene av dokumentene.

Opgaven omhandler dokumentutveksling på internett, og tar i denne sammenheng opp de «problematisk» begrepene nevnt ovenfor, med den nødvendige historiske bakgrunn og illustrerende eksempler. Etter en utgreiing rundt begrepene «åpne formater» og «proprietære formater» kommer selve diskusjonen rundt PDF som dokumentutvekslingsformat på internett tas opp. Gjennomgående eksempler på dokumentutvekslingsformater er XML, PDF og Microsoft Word-formatet, «.doc».

2 Åpne formater

Et åpent dokumentformat innebærer at formatet er åpent og kjent og kan brukes av alle. Papiret og alfabetet er gode eksempler fra den analoge hverdag.

Som åpent format regnes enhver form for digital koding av informasjon som tilfredstiller følgende krav (Hannemyr på web 24.05.03):

1. Den tekniske dokumentasjonen som beskriver formatet er offentlig tilgjengelig.
2. Kildekode til minst en brukbar implementasjon for lesing og skriving av formatet er offentlig tilgjengelig.
3. Det er ingen restriksjoner med hensyn til tredjeparts anledning til å gjøre bruk av, eller å skape programvare som lagrer, leser, overfører, mottar eller på andre måter gjør bruk av data som er kodet i henhold til formatet.

Antonymet til «åpent format» er «lukket format». Som lukket format regnes ethvert digitalt format som ikke tilfredstiller de overnevnte tre kravene. Gramstad har en god allegori med papiret som lukket proprietært format (Gramstad på web 24.05.03): Hvis papir hadde vært et lukket proprietært format kunne papirformateieren gjort det umulig eller forbudt å skrive på papiret med andre enn de han eller hun godkjenner, evt. selger eller leier ut. Formateieren kunne ha gjort det forbudt å bruke A4-formatet uten at brukerne har betalt dyre lisenser for hver side. Papirformateieren kunne faktisk ha hindret hvem som helst i å skrive eller lese, og kunne også ha slettet eller gjort utilgjengelig tidligere skrifter ved å gjøre endringer i papirformatet.

2.1 Eksempler på åpne og lukkede formater

For å illustrere kan man dra fram Extensible Markup Language (XML), som er et eksempel på et åpent format. Formatet er utviklet og forvaltes av organisasjonen World Wide Web Consortium (W3C) som alle interesserte parter har adgang til og innsyn i prosessene til. XML er et enkelt, fleksibelt tekstformat med utspring fra SGML (International Organization for Standardization (ISO), standard 8879). Originalt ble det laget for å imøtekomme utfordringene til elektronisk publisering, men XML spiller også en stadig større rolle som utvekslingsformat over World Wide Web (W3C på web 24.05.03). Spesifikasjonene til

formatet er publisert og det er ingen restriksjoner med hensyn til tredjeparts anledning til å bruke, eller skape programvare som lagrer, leser, overfører, mottar eller på andre måter gjør bruk av data som er kodet i henhold til dette format.

Et eksempel på lukket format er Microsoft Word-formatet (doc). Microsoft Word er en tekstbehandler som skaper doc-filer. Word-formatet har kommet i flere versjoner, som tilbyr stadig mer komplisert tekstbehandling og formatering. Koden som lagres i doc-filer er lukket. Dette betyr at tredjepart må gjette hvordan dokumentet skal dekodes, noe som nå blir forbudt å prøve pga DMCA, se avsnitt 2.3. Kildekoden er vanskelig tilgjengelig/ikke tillatt å modifisere, og det er dermed problematisk å konvertere doc-filer til andre plattformer; Word-formatet er dermed begrenset til brukere av Microsoft og samarbeidspartneres produkter (Gramstad på web 24.05.03).

2.2 «Open Source Software» og «Free Software»

Den første lisensen knyttet til det som senere har fått betegnelsen «åpen kildekode programvare», ble utviklet av Free Software Foundation (FSF) i forbindelse med deres GNU 11 (Gnu's Not Unix) programvare i 1984. FSF, med Richard M. Stallmann i spissen, hører hjemme i hackermiljøet. GNU Public License (GPL) var hans fremste virkemiddel i GNU-prosjektet. Denne spesielle lisensavtalen går ut på at programvare som er underlagt GPL kan brukes, endres, kopieres og formidles videre til tredjepart, men at den som gjør dette ikke kan knytte bestemmelser, betingelser eller dataformater til produktet som begrenser mottakerens rettigheter i forhold til den opprinnelige lisensen (Slaatta (red.) 2002:48).

Begrepet «Open Source» ble første gang brukt av Christine Peterson, i forbindelse med at kildekode til nettleseren Netscape Navigator, Mozilla, ble publisert i 1998 (Williams 2002). Open Source Initiative (OSI) ble stiftet med Eric S. Raymond som leder samme år (OSI på web 24.05.03). OSI publiserte en ny definisjon på begrepet, Open Source Definition (OSD), som per definisjon er samme fenomen som det som av FSF var blitt kalt for «fri programvare». I følge Stallmann er det nye begrepet en måte å få programvaren til å virke mindre skremmende for kapitalister – enn et begrep som inneholder ordet «free», som kan henlede til å tro at det handler om gratis programvare, ikke fri (i betydning frihet) programvare (Stallmann 1998:57). OSI kaller GPL en av mange mulige lisenser for åpen programvare som faller inn under OSD. Forskjellen på hvordan FSF og OSI ser på fri

programvare kan uttrykkes slik (Stallmann 1998:55): For OSI er lukket programvare en ikke-optimal løsning, for FSF er lukket programvare et sosialt problem – og fri programvare er løsningen.

2.3 Bruk og misbruk av begrepet «åpen kildekode»

Både FSF og OSI mener at begrepet «Open Source» nå har blitt så utbredt i bruk at betydningen har mistet presisjon. OSI foreslår å fortsatt bruke begrepet, men å referere til OSD (OSI på web 24.05.03), mens FSF står fast ved at «fri programvare» er det beste begrepet å bruke (Stallmann 1998:59). Så utenfor sammenheng har begrepet havnet at Microsoft (v/Jason Matusow) hevder: «at delt eller åpen kildekode verken er nytt eller fremmed for Microsoft eller andre i bransjen, men at dette har vært vanlig i over 30 år» (Computerworld 16.05.03:28). Når Microsoft beskriver «Open Source Software Model» er ikke OSI eller OSD nevnt med et ord, men når de omtaler «Free Software» handler det om GPL (Microsoft på web 24.05.03).

Begrepet «åpen kildekode» har blitt et ullent begrep som mange prøver å utnytte, det synes som en strategisk kamp er i gang, hvor formater presenteres som «åpen kildekode» uten å oppfylle OSD – og nye begreper, som «delt kildekode» og «intern åpen kildekode» dukker opp. Direkte feilaktig bruk av begrepet «åpen kildekode» finner man mange eksempler på når bedrifter tror at de ikke trenger å oppfylle alle betingelsene i OSD. Mange bedrifter prøver å gi begrepet en ny mening (Stallmann 1998:59): På Open Source Developers Day i August 1998 sa mange av de kommersielle utviklerne at de bare ville gjøre deler av sine programmer som fri programvare eller åpen kildekode. Fokuset deres er på å produsere lukkede tillegg til programmene eller lukket programvare som opererer på åpne operativsystemer som Linux. Disse bedriftene prøver å utnytte begrepet åpen kildekode for lukkede produkter ved å ha en eller annen relasjon til åpen programvare. Hvis en bedrift lager selvstendig programvare som er åpen kan de mer rettmessig kalle seg en bedrift som bruker åpen kildekode, selv om de også lager lukkede programmer.

I 2001 lanserte Microsoft Shared Source Initiative (SSI), som de mener har de beste aspektene av både kommersiell programvareutvikling og åpen kildekode programvareutvikling, men uten aspekter fra fri programvare (Microsoft på web 24.05.03). Ut fra definisjonene er fri programvare og åpen programvare det samme; når Microsoft presenterer sitt nye begrep ,«delt

kildekode», sørger de for å samtidig lage sin egen definisjon av «åpen kildekode». Det kan høres ut som Microsoft lar brukerne se på kildekode i denne modellen, men dette er forbeholdt større aktører, som offentlige myndigheter. Norske fagfolk har nå fått tilgang til kildekode fra Microsoft for å kunne gå gjennom sikkerheten i produktene.

Næringsdepartementet, med statsråd Ansgar Gabrielsen i spissen mener at det ikke er villedighet fra Microsoft, men skyldes at norske myndigheter krever å kunne evaluere sikkerheten i Microsofts produkter. (Computerworld 16.05.03:28). Delt kildekode gir ingen rett til å benytte kildekode, eller å rette opp evt. feil – men det skal rapporteres til Microsoft, som velger hvordan de vil behandle det.

En annen taktikk som brukes rundt begrepet «åpen kildekode» er at produsenter tar i bruk åpne standarder og utvider disse til å bli noe bare deres egne programmer støtter. Dette fenomenet har etterhvert blitt kjent som «embrace and extend». Etter at Microsoft har vært med å utvikle XML-formatet i W3C bruker de nå sine egne tillegg, slik at formatet bare kan leses skikkelig i Microsofts programvare. Det samme ser man enkelt hvis man lager et HTML-dokument fra Word: Dokumentet blir formatert med mange HTML-elementer som ikke er definert i standarden fra W3C, men som fungerer ypperlig med Microsofts nettleser Internet Explorer.

2.3 DMCA

Hackernes forsøk på å sabotere lukkede proprietære formater betraktes med motvilje av maktens representanter, særlig statsmakten (Slaatta (red.) 2002:49). I USA har dette i første rekke gitt utslag i Digital Millennium Copyright Act (DMCA). Haag-konvensjonen vil i realiteten gjøre DMCA gjeldende innen norsk jurisdiksjon. DMCA gjør det bl.a. til en forbrytelse å drøfte offentlig virkemåte til teknologi konstruert for å beskytte intellektuell eiendom (Slaatta (red.) 2002:42), og gjør det straffbart med reverse engineering (rekonstruksjon), selv om formålet kun er å få tilgang til egne data (Gramstad på web 24.05.03). Mangler man programvare for å eksempelvis lese et Word-dokument av eldre versjon har man altså ikke lov til å lage et leserprogram, selv om man bare er ute etter å tolke sitt eget dokument.

3 Proprietære formater

Bokmålsordbokas definisjon av proprietær er bare av substantivet «proprietør» (en gammel norsk betegnelse på godseier) (Bokmålsordboka på web 24.05.03).

Når ordet proprietær brukes om formater på norsk har det grovt sett samme betydning som den engelske «proprietary» (Gisle Hannemyr 08.04.03) – som bl.a. har denne definisjonen (Encarta på web 24.05.03):

proprietary (pr pree tère) adjective

1. relating to owners or ownership: relating to, involving or associated with an owner, ownership or something owned

Et format som ikke er fritt, men som «eies» av et selskap i den forstand at selskapet eier spesifikasjonen av formatet og kontrollerer den videre utviklingen av det, kalles proprietært. At et format er proprietært vil som oftest bety at det er et lukket format. Da vil formateieren hemmeligholde spesifikasjonene, og kan ta betalt for bruken av formatet. Et eksempel her kan være Microsoft Word-formatet (se 2.1), som er eid av Microsoft. Dette er et format med utbredt bruk – et format som er i en tiltagende monopolsituasjon, bl.a. p.g.a. stivhengighet; jo flere mennesker som bruker en bestemt teknologisk løsning, desto vanskeligere blir det for et alternativ å vinne innpass, selv om det strengt tatt er bedre (Slaatta (red.) 2002:178). Doc er ikke ett format, det er mange. Kompatibiliteten med tidligere og fremtidige versjoner er dårlig – man kan ikke være sikker på at et dokument er leselig i framtida, i hvertfall ikke hvis Microsoft går konkurs eller legger ned formatet (Gramstad på web 24.05.03).

Det motsatte av proprietære formater er ikke-proprietære formater. Disse forvaltes av organisasjoner som alle interesserte parter har adgang til og innsyn i prosessene i. Dette er åpne formater, med publiserte spesifikasjoner uten restriksjoner med hensyn til tredjeparts anledning til å bruke, eller skape programvare som lagrer, leser, overfører, mottar eller på andre måter gjør bruk av data som er kodet i henhold til disse formatene. XML-formatet forvaltes av organisasjonen W3C og er et eksempel på ikke-proprietære formater.

4 Dokumentutveksling med PDF-formatet

I 1974 gav hackeren Theodore Holm Nelson ut boka *Computer Lib/Dream Machines* med sin framtidsvisjon om hvordan datamaskiner kan brukes i et samfunn der all informasjon er åpent tilgjengelig og kan deles mellom likemenn. Nelson foreslår at informasjonen på de ulike maskinene skal være lenket sammen (Slaatta (red.) 2002:46). Engelskmannen Tim Berners-Lee skapte med World Wide Web et fungerende hypertekstsystem. I tråd med hackertradisjonen er det ikke restriksjoner på bruken av World Wide Web, og teknologien bak ikke patentert eller beskyttet på annet vis. Dataformat som HTML og XML, som muliggjør datautveksling globalt gjennom internett er fullstendig åpne og fritt tilgjengelige. Brukerne kan selv velge nettleser, plug-ins og startside.

Når man kommuniserer er det essensielt at mottakeren har mulighet til å tolke informasjonen. Dette gjelder ikke minst dokumentutveksling på internett, herunder filer lagt ut på WWW, eller dokumenter sendt som epostvedlegg. Dokumentformatet avgjør om dokumentet er lesbart, evt. blir rett formatert, for mottakeren.

Den utbredte bruken av lukkede proprietære utvekslingsformater for dokumenter (MS Word og MS Excel) er et eksempel på mikromakt som gir mottakerene beskjed om at de bør skifte ut sin nåværende plattform med datasystemer fra Microsoft (Slaatta (red.) 2002:62). Disse formatene kan ikke (pga DMCA og lukket kildekode) leses av programvare fra andre produsenter.

PDF er de facto-standard for sikker og pålitelig dokumentutveksling av elektroniske dokumenter og skjemaer over hele verden (Adobe på web 24.05.03). I motsetning til en de jure-standard, bestemt av en standardautoritet som IEEE, ISO, ANSI o.l. er en de facto-standard bare et format med utbredt bruk. PDF-formatet har nå eksistert i ti år, og gratisleseren Acrobat Reader er distribuert i mer enn 500 millioner kopier (Adobe på web 24.05.03). PDF er et universelt filformat som bevarer skriftene, bildene, grafikken og layouten i et hvilket som helst kildedokument, uansett hvilket program eller hvilken plattform det ble opprettet i. Filene er kompakte og komplette, og de kan deles, vises og skrives ut av alle som har en tilgjengelig leser for filformatet, f.eks. Acrobat Reader som er gratis tilgjengelig fra Adobe Inc.

4.1 Portabilitet

Portabilitet er et nøkkelbegrep innen elektronisk dokumentutveksling. Er et dokument portabelt vil det være lesbart for mottakeren. Tidligere trodde man at plattformuavhengighet hang nøye sammen med portabilitet, men når mer enn 80% av datamaskinene benytter en Microsoft-plattform er dette ikke noe som plager verken utviklerene eller brukerne (Hannemyr på web 24.05.01). Hvis man likevel skal våge å tro at trendene skal endre seg, må plattformuavhengighet fremdeles settes i høysetet.

PDF er et format som lar seg lese på alle plattformer, med lesere gratis tilgjengelig. Det er imidlertid relativt få som har anledning til å redigere formatet, i motsetning til millioner av brukere som har anledning til å redigere Word-filer. Microsoft har klart å skape en motsetning mellom portabilitet og plattformuavhengighet, og tvinger dermed brukerne til å velge portabilitet framfor plattformuavhengighet. På den annen side truer Internett med å skape en markeds plass hvor åpne standarder etterspørres og hvor portabilitet og plattformuavhengighet går hånd i hånd (Hannemyr på web 24.05.01), formater som PDF kan dermed true Microsofts dominans blant dokumentutvekslingsformater. Microsofts mottaktikk, som har vist seg bl.a. ved deres bruk av XML og HTML, «embrace and extend», er ikke tatt i bruk mot PDF – de vil tilby konkurrerende formater de selv kontrollerer, som det nye XML-formatet xdoc.

4.2 Båndbredde og kompresjon

Dokumenter som sendes over internett må være tilpasset mottakerens overføringskapasitet. Det er vanskelig å forutse om brukeren sjekker epost på mobiltelefon eller på en bredbåndsterminal. Dokumentene må også være tilpasset brukerens behov. Dokumenter ment for utskrift krever minst dobbelt så høy oppløsning som dokumenter ment for skjermvisning. Dokumenter ment for trykk og presse vil kreve enda høyere kvalitet. Tekst satt med tilgjengelige skrifttyper (fonter) vil ikke kreve større filstørrelse for bedre kvalitet, men når dokumentene inneholder bilder og grafikk blir dokumentene markert tyngre å laste jo bedre kvaliteten er. For å få ned filstørrelsen på dokumentene for raskere utveksling over internett kan en rekke kompresjonsteknikker tas i bruk.

PDF-formatet bygger på to eksisterende ISO-standardar: American Standard Code for Information Interchange (ASCII) og Standard Page Definition Language (SPDL), som er en ISO-kodifisering av Adobes egen PostScript-standard. Adobe bruker ISO-standardar og de

tyngste industristandardene CCITT (teleorganisasjonenes internasjonale standardiseringsorganisasjon) og Lempel Ziv Welch komprimeringsalgoritmer (LZW), til å håndtere grafikk og komprimering (Odin på web 24.05.03). LZW-komprimering er underlagt Unisys-patentet som legger restriksjoner for bruk. Det eksisterer imidlertid en rekke såkalte «LZW-frie» implementasjoner av PDF (som typisk bruker ZIP/flate-komprimeringsteknikker som erstatning for LZW). PDF lagres ASCII-kodet (tekstkode i motsetning til binærkode) – som er en fordel med tanke på gjenbruk, men krever oppsett for font-håndtering. De delene av en font som behøves for å gjengi et dokument kan inkluderes i PDF-dokumentet.

På tross av komprimering inneholder PDF-filer ofte informasjonsmengder som gjør at mottakeren bør bli varslet om filstørrelsen før nedlasting, evt. vedleggsending på epost.

4.3 Sikkerhet

Det kan være vanskelig å beskytte sine opphavsbeskyttede verk ved publisering på internett. Digitale dokumenter kan kopieres tapsfritt, og mange formater kan enkelt redigeres. Publiserer man et Word-dokument kan mottakerene enkelt kopiere tekst og bilder. Et PDF-dokument kan i laveste modus være sperret for utskrift og tekstkopiering, kun tilgjengelig som skjermbilde. Formatet har flere tilgangsnivåer innebygget, hvor det kan være passord for utskrift og/eller redigering av dokumentet.

Makrovirus kan enkelt implementeres i Word-dokumenter. PDF har blitt sett på som et virusfritt format til Peachy-viruset dukket opp, en trojansk hest Hannemyr hadde forutsett mange år tidligere (Hannemyr på web 24.05.03). Da det er tillatt å legge vedlegg inni PDF-filer vil det være mulig å lage virus som kjøres ved at vedlegget åpnes i redigeringsprogram. PDF-filer som kun leses, f.eks. i Acrobat Reader, kan ikke kjøre vedlegg, og derfor ikke inneholde slike virus.

Sikkerhet innen dokumentutveksling kan også enkelt ivaretas med kryptering. Dette kan gjøres med alle slags elektroniske dokumenter med ønsket grad av autentisering og kryptering.

4.4 PDF – åpent og proprietært

PDF er en åpen filformatspesifikasjon som er tilgjengelig for alle som ønsker å utvikle verktøy for å opprette, vise eller manipulere PDF-dokumenter. Det er altså slett ikke nødvendig å bruke Adobes produkter for å bruke PDF – det finnes mer enn 1800 leverandører som tilbyr PDF-baserte løsninger (Adobe på web 24.05.03).

Adobe Inc. har utviklet PDF-formatet, og det er aldri plassert i det fri. Det er altså Adobe Inc. som eier formatet i den forstand at selskapet eier spesifikasjonen av formatet og kontrollerer den videre utviklingen av det. Spesifikasjonen er imidlertid publisert, og Adobe tillater også tredjeparter å benytte PDF-formatet i sine produkter, samt å lage programvare som på ulike måter gjør bruk av formatet uten å betale for dette. Adobe forvalter altså eierskapet til dette åpne formatet på en ikke-monopolistisk måte. MS Word-formatet er eksempel på et format som er lukket og proprietært. XML er et format som er åpent og ikke-proprietært. PDF er et åpent proprietært format! Microsoft har også et åpent og proprietært format, men forvalter det dårlig. Dette er et format for enkel tekstformatering, ved navn Rich Text Format (RTF). De publiserte spesifikasjonene stemmer ikke overens med spesifikasjonene Microsoft selv bruker, og dette gjør det vanskelig å bruke formatet på forskjellige applikasjoner.

4.5 Driver Adobe veldedighet?

Man kan undre seg over at PDF-leseren er gratis, og desto mer over at formatet er åpent – hva har et kommersielt selskap å hente på dette? Ved å gi ut en gratis leser har Adobe senket terskelen for at folk kan ta i bruk PDF-filer – de skaper på denne måten et marked for alle dem som produserer PDF-filer. De som produserer PDF-filer blir markedet for Adobe, som leverer programvare for produksjon av PDF-løsninger.

Resonnementet over er kjapp logikk, men fungerer fint uten at formatet er åpent (bare med gratis leser). At formatet er åpent innebærer at Adobe lar hvem som helst lage programvare både for og med PDF-formatet, leverandører kan dermed dra nytte av et godt format og bidrar ytterligere til å gjøre PDF-formatet til en de facto-standard. Adobe er markedsleder på løsninger for PDF-formatet, og satser kanskje på å være nettopp det – som utvikler av formatet kan Adobe sitte på den beste kompetansen. De kan være føre-var, først ute med nye løsninger, og full kontroll med utviklingen fordi formatet er proprietært.

4.6 En «vinn-vinn»-situasjon?

Brukerne av PDF-formatet kan hente gratis lesere fra Adobes nettsider. De siste versjonene av leseren vil alltid støtte alt spesifikasjonene inneholder, da Adobe lager spesifikasjonene. Leseren er bakoverkompatibel slik at eldre PDF-filer vises slik de skal. Om Adobe skulle gå konkurs eller gi opp formatet kan fremdeles hvem som helst lage en leser ut fra de publiserte spesifikasjonene, og man er dermed sikret at PDF-formatet er lesbart i framtida. Blant mange leverandører av PDF-løsninger finnes det også gratis programvare for å lage PDF-filer. Dette kommer brukere til gode som ikke krever det ypperste av funksjonalitet, da mange av de gratis versjonene ikke inneholder det mest avanserte, som elektroniske skjemarutiner, brukertilgang o.l.

Leverandører av PDF-løsninger nyter godt av de åpne spesifikasjonene av Adobe. De står fritt til å bearbeide spesifikasjonene og lage sine egne produkter. At spesifikasjonene av et åpent format styres av Adobe alene kan muligens være et ankepunkt for noen, men dette sikrer også at formatet utvikles enhetlig, og det går kanskje raskere å utvikle spesifikasjonene, når man slipper å gå gjennom alle slags instanser og godkjenninger for hver endring.

4.7 PDF som offer for begrepsbruk

Uten presisering havner PDF-formatet lett i bås med lukkede proprietære formater og deres tilbøyeligheter. En uformell empiri-test fra søkemotoren Google gir 10.200 treff på PDF + «proprietary format», men bare 4.130 treff på PDF + «open format». Dette kan illustrere at PDF ofte nevnes blant proprietære formater, uten presisering om at det også er åpent.

Fordi PDF-formatet er knyttet så tett opp til Adobe, kan det være lett å tro at PDF bare kan produseres med Adobes produkter, som selvfølgelig anbefales i den gratis leseren fra Adobe, Acrobat Reader. Horvei, som er opptatt av problemstillinger rundt formatbruk (Horvei på web, 24.05.03), antar at PDF er lisensbilagt og dermed kan gi dokumenter som med tiden kan bli uleselige, og har derfor ingen planer om å ta formatet i bruk (Tom Ivar Horvei, 18.05.03). Hvis det er en alminnelig oppfatning at PDF er et lukket format kan dette virke negativt for omdømme og kan gi en høyere terskel for å ta formatet i bruk.

6 Oppsummering

I arbeidet med innhenting av stoff til denne oppgaven var det frustrerende og misvisende at terminologiene «åpent» og «proprietært» ble brukt som antonymer. Denne begrepsbruken er så utbredt at selv Hannemyr, som på et senere tidspunkt lagde en definisjon som klarer opp i dette, bruker begrepene ukorrekt i pensumlitteraturen til denne oppgaven (Slaatta (red.) 2002:47).

Vanligvis er begrepsbruken kanskje ikke et problem, da de fleste proprietære formater er lukkede, og de fleste åpne formater er ikke-proprietære. Begrepsbruken kompliserer når man skal gjøre rede for Adobe PDF, som er et åpent proprietært format. Et skille går mellom proprietære og ikke-proprietære formater, hvor de ikke-proprietære formatene ofte ivaretas av organisasjoner (som f.eks. W3C eller ISO), mens proprietære formater eies av kommersielle aktører. Et annet går mellom «åpent»/«lukket»: Antonymet til «åpent» er «lukket», og det er også her det viktigste skillet mellom formatene går (Gisle Hannemyr 09.04.03). Begrepene «åpent format» og «åpen kildekode» har blitt utydelige bl.a. gjennom mediebruk. Det pågår en strategisk kamp rundt uttrykket «Open Source»: Det har blitt populært å bruke begrepet «Open Source» uten at det har rot i definisjonene; derfor bør man gå produktene nærmere i sømmene om de benytter dette begrepet. Hvis et format eller en programvare oppfyller GPL bør man kalle det «fritt/fri» (Stallmann 1998:57), hvis det oppfyller OSD bør det referere til OSD (OSI på web 24.05.03).

PDF-programvare fra Adobe er ikke åpent, men det er PDF-formatet. At PDF er et åpent format er ikke et hinder for at Adobe kan være en kommersiell aktør med lukket programvare, tvert i mot har det gjort PDF-formatet til en de facto-standard, med et stort marked for tilhørende programvare.

PDF-formatet er godt egnet til dokumentutveksling på internett. Formatets egenskaper forsvarer at det er godt egnet, ved at det tar vare på layout og alltid fremtrer slik det er laget – men desto viktigere er det at leseren er gratis tilgjengelig, og formatet åpent. Dette har gitt en portabilitet som gjør at PDF-filer kan leses og redigeres på alle plattformer. Formatet er også godt egnet til lagring, da man med større sikkerhet får tolket formatet i fremtiden enn for lukkede proprietære formater.

Det pågår en politisk kamp for åpne formater. Å eie et format er noe helt annet enn det å eie et verk (det er det siste som har med opphavsrett og lover om åndsverk å gjøre). Folk kan tjene penger på å produsere papir, eller verk på papir, uten at noen eier papirformatet. I fremste rekke finner man fremdeles hackerne, men offentlig forvaltning har også fått øynene opp for problemet. Hannemyr har lagt fram et lovforslag om at all offentlig informasjon skal utveksles og langtidslagres i åpne formater (Hannemyr på web 24.05.03). Staten sa i 2002 opp sin innkjøpsavtale med Microsoft, slik at offentlige etater står fritt til å bruke alternativ programvare (NRK på web 24.05.03). Professor i rettsinformatikk, Olav Tørvund, mener at offentlighetsloven bør fremme bruken av åpne formater. Forvaltningsloven, som tar opp kommunikasjon mellom offentlige og private aktører, samt arkivloven som har med langtidslagring å gjøre, må også fremme åpne formater slik at det offentlige ikke risikerer retten til å lese formatet (Digi på web 24.05.03). UiB har i sin IKT-plan 2003-2007 gått inn for at format som HTML og PDF er ønskelige for dokumenter som i utgangspunktet er ment for lesing og ikke redigering for mottakeren (UiB på web 24.05.03). Politiske beslutninger må tas for at ikke bruk av teknologi skal kunne effektivere re-delegering av makt fra politiske institusjoner til private aktører (Le Monde Diplomatique 11.04.03:28). Bruk av åpne formater, som PDF, lar brukerne beholde makten over egne dokumenter.

Referanser

Adobe: *Hva er Adobe PDF?*

<<http://www.adobe.no/products/acrobat/adobepdf.html>>, besøkt 24.05.03

Bokmålsordboka: *Artikkel fra bokmålsordboka*

<<http://www.dokpro.uio.no/perl/ordboksoek/ordbok.cgi?OPP=propriet%E6r&bokmaal=S%F8k+i+Bokm%E5lsordboka&ordbok=bokmaal&alfabet=n&renset=j>>, besøkt 24.05.03

Digi: *Offentlighetsloven kan brukes for åpne formater*

<http://www.digi.no/dtno.nsf/pub/dd20021031145808_ero_63693751>, besøkt 24.05.03

Encarta: *MSN Learning & Research - Dictionary*

<<http://encarta.msn.com/encnet/features/dictionary/DictionaryResults.aspx?refid=1861736468>>, besøkt 24.05.03

Eriksen, Thomas Hylland (2002): *Ordets makt: The power of Word*, s. 174-191 i Slaatta, Tore (red.) *Digital makt*. Oslo, Gyldendal Akademisk.

Gramstad, Thomas: *Brave New Word*

<<http://www.efn.no/brave-new-world.html>>, besøkt 24.05.03

Gramstad, Thomas: *Hvem skal eie formatene?*

<<http://www.efn.no/hvem-eier-formatet.html>>, besøkt 24.05.03

Hannemyr, Gisle: *Adobe PDF and Trojan Horses*

<<http://folk.uio.no/gisle/trap.html>>, besøkt 24.05.03

Hannemyr, Gisle: *Digital tilgjengelighet*

<<http://folk.uio.no/gisle/essay/lovofb.html>>, besøkt 24.05.03

Hannemyr, Gisle (2002): *Foucault i kyberrommet*, s. 41-63 i Slaatta, Tore (red.) *Digital makt*. Oslo, Gyldendal Akademisk.

Hannemyr, Gisle: *Internett slik du vil ha det*

<<http://folk.uio.no/gisle/essay/ms01.html>>, besøkt 24.05.03

Hannemyr, Gisle: Microsofts skjulte imperium, s. 26-28 i *Le Monde Diplomatique*, Oslo, fredag 11. april, s. 26

Horvei, Tom Ivar: *Går din informasjon tapt?*

<<http://www.horvei.com/artikler/infotap.html>>, besøkt 24.05.03

Johnsen, Frank: Vinduet på gløtt, s. 29 i *Computerworld*, Oslo, vol 20, nr. 27, fredag 16. mai, s. 28

Microsoft: *Basic Principles of Software Source Code Licensing*

<<http://www.microsoft.com/resources/sharedsource/Articles/LicensingBasics.msp>>, besøkt 24.05.03

NRK: *Gir Microsoft på båten*

<<http://www.nrk.no/nyheter/okonomi/1987724.html>>, besøkt 24.05.03

ODIN: *Elektronisk post i statsforvaltningen*

<<http://odin.dep.no/odinarkiv/norsk/dep/ad/1995/publ/034005-990044/index-hov007-b-f-a.html#fn4>>, besøkt 24.05.03

OSI: *The Open Source Definition*

<<http://www.opensource.org/docs/definition.php>>, besøkt 24.05.03

Stallmann, Richard M. (1998): *Free Software, Free Society: Selected Essays of Richard M. Stallman*. Fri kopi fra <<http://notabug.com/2002/rms-essays.pdf>>, besøkt 24.05.03

UiB: *Alternativ programvare og standard filformat*

<<http://www.uib.no/it/doc/altpro-stafil.html>>, besøkt 24.05.03

W3C: *Extensible Markup Language (XML)*

<<http://www.w3.org/XML/>>, besøkt 24.05.03

Williams, Sam (2002): *Free as in Freedom*, Fri kopi fra
<<http://www.oreilly.com/openbook/freedom/ch11.html>>, besøkt 24.05.03

Intervju v/epostutveksling

Gisle Hannemyr, intervjusvar sendt på epost: 08.04.03 kl 21:25, 09.04.03 kl 14:13

Tom Ivar Horvei, intervjusvar sendt på epost 18.05.03 14.01