UNIVERSITETET I OSLO
Institutt for informatikk

Kan metadata sikre validiteten til digitale dokumentbevis?

Masteroppgave

Astri Ek Larsen

1. mai 2007

Sammendrag

Stadig mer bevismateriale i rettstvister er digitalt, og stadig større mengder materiale sikres i etterforskningen av norske straffesaker. Påtalemyndighetens plikt til å belyse straffesaker best mulig medfører store mengder fremlagt bevismateriale, og dommerne må ta stilling til om bevisene skal antas eller avskjæres. Lover og praksis i norsk strafferett henger etter den teknologiske utviklingen, og det fins i dag ingen retningslinjer for hvordan digitalt materiale skal håndteres. Den teknologiske utviklingen skjer så raskt at advokater og dommere ikke evner å tilegne seg den tekniske kompetansen de trenger for å henge med i utviklingen.

Denne oppgaven tar for seg digitalt bevismateriale og håndtering av dette. Oppgaven gir en introduksjon til norsk straffeprosess, og med et teknologisk perspektiv forklares de tre fasene i digital etterforskning. Verdikjeden til et digitalt dokumentbevis forklares også. I tillegg ser oppgaven på relevant forskning i Norge, Europa og USA.

For at et digitalt dokument skal antas som bevis i strafferetten må det ha validitet. I oppgavens problemstilling stilles spørsmålet om metadata kan sikre denne validiteten. Metadata er informasjon om et dokuments egenskaper, men for at denne informasjonen skal kunne benyttes for å si noe om validiteten må den være korrekt, ekte og ikke ha blitt manipulert med. Oppgavens forskning går ut på å undersøke hvilke metadata som genereres når et digitalt dokument opprettes. Det undersøkes også om metadataene kan manipuleres, eventuelt hvordan og hvilke aktører som kan gjøre det.

Forskningen viser at det er en rekke metadata som genereres. Det er dessuten mulig å manipulere og opprette nye metadata i etterkant av opprettelsen av dokumentet. Flere ulike aktører kan gjøre dette – både programvare- og personrolle-aktører. Argumentene som taler for at metadata kan sikre validitet, som at tidsstempler er sikker informasjon og at sakkyndige kan oppdage manipulasjon er ikke godt nok fundamenterte. Oppgaven konkluderer med at metadata ikke kan sikre validiteten til digitale dokumentbevis fordi denne typen informasjon kan manipuleres.

Innhold

Tabeller

Figurer

Forord

Denne oppgaven er skrevet som en del av min mastergrad ved Institutt for Informatikk, Universitetet i Oslo.

Jeg vil takke min veileder ved Universitetet i Oslo, Odd Aurmo, for god støtte og hjelp gjennom arbeidet med denne oppgaven. Takk til Kjell Thorvaldsen, gjesteforsker ved Norsk Regnesentral, som har kommet med mange idèer og gode innspill. Tusen takk for veiledning og korrekturlesing av Maria Astrup Hjort. Jeg vil også rette en takk til daVinci Consulting AS for lån av utstyr.

En stor takk til medstudenter, venner og familie som har hjulpet meg gjennom denne oppgaven. Deres oppmuntring, veiledning, tips og støtte har gjort det mulig for meg å gjennomføre masterstudiet. En spesiell takk til Frode Gundersen som har bidratt med uvurderlig hjelp i form av konstruktiv kritikk, korrekturlesing og ikke minst oppmuntring og støtte.

Oslo, mai 2007
Astri Ek Larsen

Kapittel 1

Introduksjon

I avsnitt 1.1 presenteres motivasjonen bak denne masteroppgaven. I avsnitt 1.2 blir problemstillingen lagt frem, og avsnitt 1.3 beskriver metodene jeg har benyttet for å svare på spørsmålene i problemstillingen. I avsnitt 1.4 vises hvordan denne masteroppgaven bidrar til forskning og samfunn, og avgrensninger jeg har gjort blir beskrevet i avsnitt 1.5. Det siste avsnittet i dette kapittelet, avsnitt 1.6, vise denne masteroppgavens struktur.

1.1 Motivasjon

I følge organisasjonen Cybex er mer enn 90 prosent av alle dokumenter i et firma digitale, og digital kommunikasjon har hatt en kraftig vekst de siste årene [40]. I norsk strafferett har påtalemyndigheten plikt til å belyse en sak best mulig. Det gir et insentiv til å ønske å fremlegge mange bevis. Dette medfører at stadig mer digitalt materiale blir knyttet til rettstvister, og etterforskere sikrer, analyserer og presenterer oftere digitale bevis nå enn for bare noen år siden.

Lover, regler og praksis i norsk strafferett har imidlertid ikke rukket å tilpasse seg den nye teknologiske virkeligheten. Det medfører at norsk lovgivning ikke har et regelverk som gir retningslinjer for hvordan spesielt digitalt materiale skal håndteres. Advokater og dommere, som skal tolke og håndheve de lovene som faktisk eksisterer, har heller ikke særlig teknisk kompetanse. Det er bekymringsfullt fordi det er opp til dommerne å avgjøre om et potensielt bevis skal antas eller avskjæres. I tillegg fins ingen formelle krav til hvem som kan avgi vitneprov som sakkyndig.

Bevismateriale som påtalemyndigheten legger frem i straffesaker sikres og analyseres av etterforskere. Spesiell programvare benyttes til dette, og slike verktøy kan for eksempel speilkopiere og analysere beslaglagte hard-disker og andre lagringsmedier. Verktøyene som brukes av norske etterforskere har ofte opprinnelse i USA der en helt annen jurisdiksjon gjelder. Dermed er ikke verktøyene tilpasset norske forhold

og norsk lovgivning. Det er lite hensiktsmessig fordi det da kan oppstå uoverensstemmelse mellom lovgivning og programvareprosedyrer. Et økt tverrfaglig samarbeid mellom juridiske og teknologiske forskningsmiljøer vil kunne bedre situasjonen. Ett resultat kan være programvareverktøy utviklet for norsk lovgivning og norske prosessregler.

Forskning utført av John D. Gregory viser at at integriteten til fysiske dokumenter blir mer eller mindre tilfeldig behandlet når de vurderes for bevisbruk [10]. Stephen Mason påpeker i sitt arbeid gjennom ARMA International Educational Foundation i tillegg at det samme kan sies å gjelde for dokumenter på digitale formater [16, 36]. Selv om disse to forskerne har sett på forhold i henholdsvis USA og Storbritannia, kan vi anta at situasjonen ikke er vesentlig annerledes ved norske domstoler.

Advokat Arve Føyen er kjent med denne problematikken [37]. I en kommentar i magasinet Computerworld Norge fra 2006 påpeker han at det er stor usikkerhet rundt håndtering av digitale dokumenter i norsk rettspraksis. Det fins ikke spesifikke retningslinjer for hvordan digitale bevis skal ilegges beviskraft [9]. Føyen sier seg enig i påstanden om at norske advokater har begrenset kunnskap om forholdet mellom et opprinnelig digitalt dokument og en fremlagt papirutskrift.

Digitalt materiale inneholder metadata. Dette er strukturert informasjon som forteller noe om materialet. Metadataene til et digitalt dokument kan for eksempel fortelle hvem som opprettet det, når det ble opprettet og når det sist ble endret. Ved å formalisere slik informasjon kan dokumentets verdikjede dannes. Verdikjeden til et dokumentbevis beskriver historien eller livsløpet til beviset. Den sier noe om hendelser som har inntruffet (opprettelse, utskrift, lagring), hvem som har fått hendelsene til å skje (aktører) og når de skjer (tidsstempler). Hvis verdikjeden til et bevis er kjent, vet man også mer om egenskaper som brukbarhet, integritet og troverdighet.

I dagens strafferett fins det verken rutiner eller retningslinjer for kartlegging av verdikjeden til et digitalt bevis på denne måten. For at en dommer skal kunne anta et bevis, er det likevel viktig at bevisets validitet er dokumentert. Validiteten bygger på autentisiteten til beviset, og denne bygges opp av de tre egenskapene brukbarhet, integritet og troverdighet. Når påtalemyndigheten skal sikre, analysere og presentere bevis bør de derfor legge vekt på disse egenskapene for å få en mest mulig korrekt bevisfremlegging. Det burde også finnes rutiner og retningslinjer for å kartlegge verdikjeden til digitale bevis i norsk strafferett slik at domstolen bedre kan avgjøre validiteten til slike bevis før de antas eller avskjæres.

For å få oversikt over et digitalt dokuments verdikjede kan man altså se nærmere på metadataene til dokumentet. Denne informasjonen er med på å gi oversikt over verdikjeden og er vesentlig når man skal avgjøre dokumentets autentisitet og validitet. For at metadataene kan gi troverdig informasjon for å avgjøre dette, må man imidlertid være sikker

på at informasjonen er korrekt, ekte og ikke manipulert. Kan metadata manipuleres? Hvem kan eventuelt manipulere metadata, og hvordan kan det gjøres? Kan metadata sikre validiteten til digitale dokumenter? Dette er spørsmål som denne masteroppgaven tar sikte på å besvare.

1.2 Problemstilling

Hovedfokus i denne masteroppgaven er digitale dokumentbevis. Spesielt ønsker jeg å se på metadata tilknyttet digitale dokumenter. Denne formen for data gir informasjon utover innholdsdataene i dokumentet. Jeg vil undersøke hvilke metadata som fins og hvordan disse potensielt kan endres. Formatet som har fokus er dokumenter opprettet i Microsoft Word 2003.

Jeg vil gi en generell introduksjon til norsk strafferett og se spesielt på digital etterforskning og digitale dokumenters verdikjede. Forskning innen teknologi relevant for digitale data vil også få fokus. Blant annet vil jeg gi en kort gjennomgang av kjente analyse- og manipuleringsverktøy.

Digitale dokumenter og metadata er fokus i følgende problemstilling:

P: Kan metadata sikre validiteten til digitale dokumentbevis?

- P_1: Hvilke metadata lagres?
- P_2: Hvilke aktører genererer metadata?
- P_3: Hvordan kan metadata manipuleres?
- P_4: Hvilke aktører kan manipulere metadata?

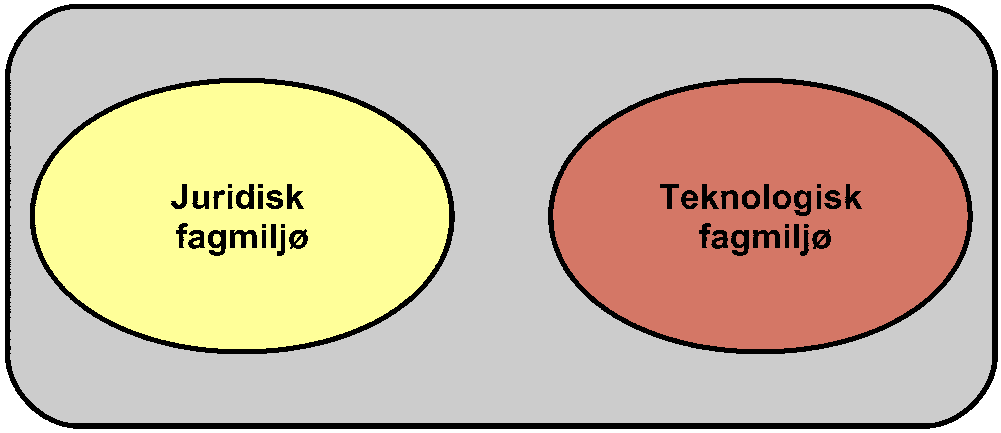
1.3 Forskningsmetode

Spørsmålene i problemstillingen vil bli besvart ved hjelp av programvaredokumentasjon og ved å gjennomføre metadata-analyse og metadata-manipulasjon. Undersøkelse av metadata vil være basert på MS Word 2003s behandling av disse.

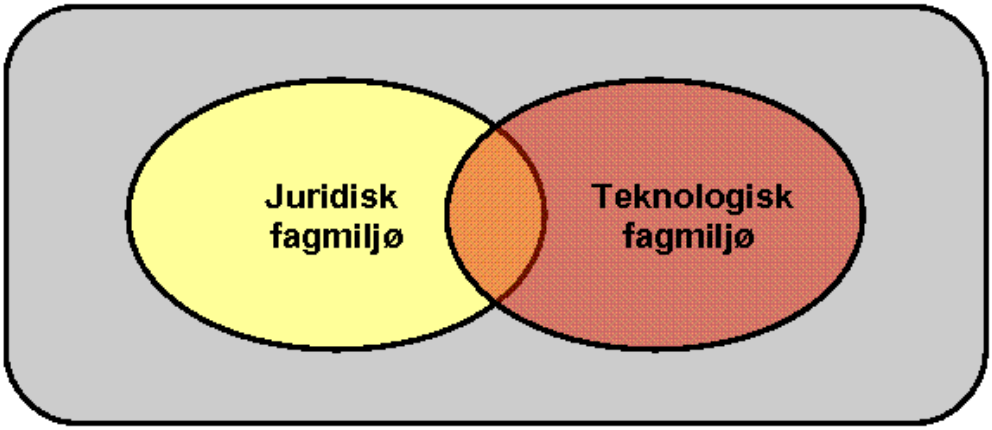
Metadata-analyse og -manipulasjon vil bli gjennomført ved hjelp av verktøyene beskrevet i kapittel 4.2. Disse verktøyene blir brukt i både offentlige og private organisasjoner i Norge så vel som i internasjonal sammenheng. Analyseverktøyene vil bli brukt for å undersøke hvilke metadata som genereres og lagres. Deretter vil metadataene bli utsatt for forsøk på manipulasjon for å undersøke hvilke aktører som kan endre, manipulere og slette metadata. Resultatene fra testene vil bli brukt i diskusjonen om hvorvidt metadata kan sikre validiteten til digitale dokumentbevis.

1.4 Bidrag

Både digital etterforskning generelt og bruk av digitale bevis spesielt, krever kompetanse både innen teknologi og jus. Forskere og andre profesjonelle aktører fra de teknologiske og juridiske fagmiljøene bør jobbe sammen for å fremme forskning rundt denne oppgavens tema. Samarbeidet mellom de to fagmiljøene er i dag svært begrenset. Heller ikke kommunikasjonen er så god som man kunne ønske. Dette illustreres i figur 1.1.



Figur 1.1: To adskilte fagmiljøer



Figur 1.2: To overlappende fagmiljøer

Universitetet i Oslo har Avdeling for Forvaltningsinformatikk som arbeider med informasjons- og kommunikasjonsteknologi på tvers av

fakultetsgrensene ved universitetet. Selv om denne avdelingen er et skritt i retning av bedre samarbeid mellom de to fagmiljøene, er det fortsatt en lang vei å gå.

Det er viktig å utvikle teori og prosedyrer som er i skjæringspunktet mellom fagdisiplinene. Da øker sannsynligheten for riktige prosessavgjørelser med hensyn til digitale bevis. Det må skapes et miljø der man har bred innsikt i både juridiske og teknologiske spørsmål knyttet til digitale bevis og digital etterforskning. God teknisk kontroll kan skapes ved å forene det tekniske forskningsmiljøet med både det juridiske og det politifaglige miljøet. Inger Marie Sunde påpeker i boken "Lov og rett i cyberspace" at det kan utvikles ulike digitale etterforskningsverktøy som er spesielt tilpasset norsk straffeprosess og rettstradisjon. Det krever imidlertid bevisst politikk og at utviklingen foregår innen et tverrfaglig miljø [27]. Området der sirklene overlapper hverandre i figur 1.2 illustrerer et ønsket tverrfaglig samarbeid mellom juridiske og tekniske forskningsmiljøer.

Denne masteroppgaven er en del av prosessen det er å bringe fagmiljøene tettere sammen. Dette gjøres ved å fokusere både på juridiske og teknologiske aspekter ved digitale dokumentbevis. For å gi teknologer en innføring i rettslig praksis gis det i kapittel 2.2 en lettfattelig gjennomgang av prosesser, bevisføring, digital etterforskning og bruk av sakkyndige. Samtidig har denne oppgaven også et teknologisk perspektiv, og det vil bli benyttet tekniske metoder for å besvare problemstillingene presentert i avsnitt 1.2.

1.5 Avgrensninger

Det er digitale dokumenter brukt som bevis i norske straffesaker som er i fokus i oppgaven. Det betyr at det ikke tas hensyn til andre digitale bevistyper. Det er spesielt straffeprosessen som beskrives, og selv om sivilprosessen nevnes er ikke denne vesentlig i denne oppgaven.

Med dokumentbevis menes tekstinformasjon gitt i form av typisk brev og kontrakter skrevet i et tekstbehandlingsprogram, slik begrepet blir definert i avsnitt 2.1. I denne oppgaven fokuseres det på dokumenter opprettet i tekstredigeringsapplikasjonen Microsoft Word 2003 [51]. Det betyr at andre dokumentformater som for eksempel OpenOffice Writer-dokumenter eller dokumenter som er lagret på pdf-format ikke vil bli undersøkt og analysert.

Microsoft Word er programvare som stadig er i utvikling. Nylig ble MS Word 2007 sluppet på markedet. Denne applikasjonen synes å virke på en ganske annen måte enn MS Word 2003 og bruker metadata i mye større grad enn sin forgjenger. MS Word 2007 vil imidlertid ikke bli undersøkt i denne masteroppgaven. MS Word 2003 er, og vil i lang tid være en markedsledende programvare for dokumentproduksjon, og jeg

anser derfor en undersøkelse av denne programvareversjonen for å være fortsatt aktuell til tross for nyere versjoner.

E-post er også et digitalt format som stadig oftere blir lagt frem som bevis. Denne oppgaven vil ikke omhandle denne typen data. Heller ikke informasjon produsert av operativsystemer og servere vil bli behandlet i denne oppgaven. I tillegg faller databaseinnhold og system- og datalogger utenfor rammene av denne oppgaven.

1.6 Struktur

I dette kapittelet har jeg gitt en introduksjon til og motivasjon for denne masteroppgaven. Problemstilling og metoder som skal benyttes har blitt presentert. Bidraget denne oppgaven gir til forskning og hvilke avgrensninger som har vært nødvendige har blitt gjennomgått og her gjennomgås oppgavens struktur.

Kapittel 2 gir bakgrunn for og kontekst til denne masteroppgavens tema. Begreper som blir brukt defineres og presiseres i avsnitt 2.1. I avsnitt 2.2 gir jeg en kort oversikt over norsk straffeprosess, og i avsnitt 2.3 forklares de ulike fasene i digital etterforskning. Verdikjeden til et bevis beskrives i avsnitt 2.4.

I kapittel 3 gir jeg en oversikt over relatert forskning. Her beskrives organisasjoner og forskningsprosjekter som jobber med ulike problemstillinger innen digital etterforskning og digitale bevis.

I kapittel 4 beskriver jeg først hvilke forskningsmetoder jeg bruker for å besvare spørsmålene som problemstillingen reiser. Analyse- og manipuleringsverktøy som brukes av etterforskningsinstitusjoner og som jeg vil benytte meg av i min analyse og undersøkelse, beskrives i avsnitt 4.2.

Resultater fra undersøkelsene som inngår i denne masteroppgaven legges frem i kapittel 5.

I kapittel 6 diskuterer jeg resultatene jeg presenterte i kapittel 5. Forslag til mulige tiltak for å løse utfordringene ved bruk av digitalt bevismateriale legges frem i avsnitt 6.2, og validiteten til oppgavens resultater diskuteres i avsnitt 6.3.

I kapittel 7 presenteres konklusjonen og forslag til videre arbeid.

Kapittel 2

Bakgrunn

Dette kapittelet gir bakgrunnsinformasjon og kontekst til temaet i denne masteroppgaven. Begreper som blir brukt defineres og presiseres i avsnitt 2.1. I avsnitt 2.2 gis en oversikt over norsk rettsbehandling, og i avsnitt 2.3 forklares de ulike fasene i digital etterforskning og eksempler på hvilke organisasjoner som jobber innenfor dette. Verdikjeden til et bevis beskrives i avsnitt 2.4.

2.1 Begreper

I dette avsnittet defineres og presiseres en rekke aktuelle begreper brukt i denne oppgaven.

Dokument Justisdepartementet skriver i rapporten fra kartleggingsprosjektet at dokumentbegrepet tradisjonelt er knyttet til papir [13]. Det er ikke alltid opplagt hva som er et dokument. I forarbeidene til offentlighetsloven blir ikke digitale dokumenter nevnt, men i utgangspunktet var det utskriftene som ble ansett som dokumenter. I statens generelle kravspesifikasjon om elektronisk saksbehandling er dokument "... en avgrenset og sammenhengende informasjonsmengde, fremstilt for et bestemt formål. Informasjonen kan bestå av en kombinasjon av tekst, data, grafikk, bilder og multimedia. Et dokument kan også bestå av flere dokumenter (sammensatte dokumenter)."

Straffelovens § 179 definerer dokument slik:

> "Ved Dokument forstaaes i denne Lov enhver Gjenstand, som i Skrift eller paa anden Maade indeholder et Tilkjendegivende, der enten er af Betydning som Bevis for en Ret, en Forpligtelse eller en Befrielse fra en saadan eller fremtræder som bestemt til at tjene som Bevis."

I ØKOKRIMs skriftserie nummer 9, "Datakriminalitet", fra 1995 konkluderes det med at metadata er tilkjennegivende slik tradisjonelle dokumenter også er.

I denne oppgaven vil begrepet dokument brukes slik som Kunnskapsforlagets fremmedordbok definerer det: "Eksemplar av et medium som lagrer informasjon for senere overføring, lytting eller lesing" [8]. Begrepet brukes både om informasjonen på papir og på digitalt format. Metadata knyttet til informasjonen er en del av dokumentet.

Digital På Wikipedia defineres et digitalt system som et system som bruker diskrete verdier, spesielt binære tall, for blant annet input, prosessering og lagring. Dette til forskjell fra det analoge kontinuerlige spekteret av verdier [32].

Eksempler på digital data er alt som har blitt opprettet eller lagret på en datamaskin, slik som digitale dokumenter, e-post, musikk på mp3-format og digital radiokringkasting.

Analogt dokument Begrepet brukes i denne oppgaven om dokumenter som tradisjonelt fins på papir og som ikke har blitt opprettet digitalt.

Original og kopi I Justisdepartementets kartleggingsprosjekt blir det forklart hva som menes med original – noe ekte og opprinnelig [13]. Innholdet endres ikke over tid, og det fins ofte bare èn original. Da kan originalen kan dermed bare være på ett sted av gangen. Når man snakker om papirdokumenter kan man på den annen side også bestemme hvor mange originaler som opprettes. Selv om èn original ofte brukes, forekommer det også at flere originaler utstedes slik som for eksempel der det undertegnes kontrakter i det antall underskrifter som skal påføres.

I kartleggingsprosjektet poengteres det at begreper som original og kopi tradisjonelt er knyttet til papirbaserte dokumenter. Kopi er en gjengivelse av et originalt dokument og "dokumentet utsettes for et generasjonstap(visuelt) ved kopiering". Dette blir basert på at en kopi ikke opprettholder alle detaljer fra originalen slik som farge og blekk ved underskrift.

Rapporten fortsetter med at det vil være noe mer problematisk å definere original og kopi når man har med digitale dokumenter å gjøre, da man ikke vil ha et generasjonstap slik papirdokumenter har.

United Nations Commission on International Trade Law (UNCITRAL) er en FN-kommisjon og består av representanter fra 60 medlemsland som velges for 6 år av gangen. Norge er for tiden ikke et av disse. Kommisjonen jobber med å harmonisere og forene

den internasjonale handelsloven. I UNCITRALs modellov artikkel 8 brukes begrepet original også om digitale dokumenter, men i enkelte deler av loven, som i regnskapsloven, brukes imidlertid begrepet original kun for papirbaserte dokumenter [19].

Kartleggingsprosjektet foreslår at man kan benevne dokumenter som original "så fremt det viser hvem som sendte det, at det ikke er blitt endret, kan leses av mottakeren mv. - uavhengig av hvilken type av medium som er blitt benyttet" [13].

Manipulering Begrepet benyttes i denne oppgaven for både modifikasjon og fabrikkering. Med modifikasjon menes endring av data, og med fabrikkering menes tilføyelse av data, slik begrepene defineres i "Lov og rett i cyberspace" av I. M. Sunde [27].

Autentisering Med autentisering menes i denne masteroppgaven verifisering av avsenderens eller mottakerens identitet. Begrepet brukes i samme betydning i Justis- og politidirektoratets høringsuttalelse "Tilrettelegging for elektronisk kommunikasjon med domstolene – Forslag til endringer i rettspleielovene" fra september 2001 [24].

Dataintegritet Med begrepet dataintegritet menes her at man kan verifisere at innholdet i et dokument ikke er endret fra det ble skrevet til det blir fremlagt som bevis i en rettstvist. Begrepet brukes i samme mening som definert av Vanstone et.al. i 1996 [17]:

> "(...) the property whereby digital data has not been altered in an unauthorized manner since the time it was created, transmitted, or stored by an authorized source."

Sporbarhet Sporbarhet er et prinsipp i offentlig forvaltning som sikrer at behandlingen av en sak kan rekonstrueres i ettertid. Begrepet brukes også i regnskapslovgivningen og betyr at det skal være mulig å etterspore hva som har skjedd (eng: audit trail).

Formater og lagring over tid Justisdepartementets rapport fra kartleggingsprosjektet tar opp problemet knyttet til kravet om at dokumenter må arkiveres og lagres over lang tid og mange år [13]. Fordi papirdokumenter ikke krever verktøy for å leses vil man, forutsatt at trykket holder, kunne lese et papirdokument flere tiår etter at det ble produsert. I den digitale verdenen utvikles stadig ny programvare og gammel programvare forsvinner ut. Dermed mister man også støtte for å lese dokumenter skrevet for lang tid siden. For å hindre dette, må man konvertere dokumenter over til andre formater. Som departementet påpeker, medfører konverteringen at det digitale dokumentet kan miste sin unikhet ved at dokumentlikheten forsvinner.

Nummer	Element	Nummer	Element
1	Title	12	Date
2	Subject	13	Format
3	Description	14	Identifier
4	Type	15	Language
5	Source	16	Audience
6	Relation	17	Provenance
7	Coverage	18	RightsHolder
8	Creator	19	InstructionalMethod
9	Publisher	20	AccrualMethod
10	Contributor	21	AccrualPeriodicity
11	Rights	22	AccrualPolicy

Tabell 2.1: Kjerne-elementer i DCMIs definisjon av metadata

Ved konvertering av et dokument med digital signatur oppstår et nytt problem. I slike tilfeller vil det ikke lenger være mulig å knytte den opprinnelige signaturen til det konverterte dokumentet. I henhold til avnsitt 8.4 i UNCITRALs modellov er det innholdsintegritet og tilgjengelighet som er viktig. Dokumentet er fortsatt en original selv om lagringskonverteringen påvirker dokumentets opprinnelige utseende [13].

Metadata I denne oppgaven brukes begrepet metadata slik som Dublin Core Metadata Initiative (DCMI) har definert det. DMCI er en organisasjon som utvikler metadata-standarder. Organisasjonen ønsker å fremme utvikling av metadata slik at digitalt materiale kan beskrives bedre for å oppnå mer korrekte informasjonssøk [43].

DCMI definerer metadata som strukturert data om data. Det vil si at metadata er informasjon som beskriver et objekt eller en resurs – fysisk eller digitalt. Begrepet metadata er i seg selv relativt nytt, men man finner igjen konseptet om beskrivende informasjon i arkiv- og bibliotekskontekst. Der har informasjon alltid blitt organisert og kategorisert for bedre lagring og enklere gjenhenting [7].

DCMIs definisjon er utviklet for å være enkel og konsis og for å beskrive nettbaserte dokumenter. Den har 22 kjerne-elementer som alle er valgfrie og kan finnes i flere eksemplarer. Elementene kan presenteres i valgfri rekkefølge, og de listes opp i tabell 2.1 [6].

NISO, en gruppe tilknyttet American National Standards Institute (ANSI), forklarer at det fins tre typer metadata. *Deskriptive metadata* beskriver kilden for å forenkle gjenfinning og identifisering. Her finner man elementer som tittel, forfatter og nøkkelord. *Strukturelle metadata* forteller hvordan sammensatte objekter er bygget opp. *Administrative metadata* gir informasjon om for eksempel når og hvordan et dokument er opprettet, dokumentets filtype samt rettigheter knyttet til dokumentet [22, 35, 52].

Metadata-analyse Med metadata-analyse menes strukturert gjennomgang av eksisterende metadata som tilhører et digitalt dokument. Dette kan gjøres på flere ulike måter, for eksempel ved hjelp av MS Word 2003s "Egenskaper"-visning, en xml-editor eller analyseprogrammer.

Scrubber En scrubber er en type programvare laget for å manipulere et digitalt dokuments metadata. Gjennom et enkelt grensesnitt tilbys brukeren å endre lagrede metadata. En scrubber har i tillegg funksjonalitet for å slette metadata.

RDF Resource Description Framework (RDF) er et deklarativt språk som definerer en XML-standard for å representere metadata. Dette gjøres ved hjelp av utsagn om egenskaper og relasjoner mellom elementer på nettet. Ved bruk av digital signatur kan RDF for eksempel brukes til å uttrykke informasjon om hva som signeres, hvilken betydning signaturen har og hvor lenge den er gyldig [53].

Tidsstempel Et digitalt tidsstempel er dato og tid lagret av et digitalt medium [34]. Det fins ulike formater for slik lagring. Tidsstempler genereres for eksempel når filer blir opprettet, endres eller åpnes. Det er også vanlig at datasystemer har loggings-funksjonalitet over systemprosesser og at disse tidsstemples.

Digital etterforskning Med begrepet menes praktisk bruk av metoder, rammeverk og verktøy for bevaring, innsamling, validering, identifisering, analyse, tolkning, dokumentasjon og presentasjon av digital informasjon.

Fane Begrepet brukes i denne oppgaven som det engelske uttrykket "tab" kjent fra nettlesere og annen programvare. Faner deler vinduer inn i områder.

2.2 Beskrivelse av rettsbehandlingen

I dette avsnittet beskriver jeg hvordan rettsbehandlingen er i Norge. Jeg går igjennom prosedyrer ved bruk av sakkyndige, ulike bevistyper og bevishåndtering beskrives så før jeg til slutt gir et overblikk over straffelovens dokumentfalskbestemmelser. Det er straffeprosessen som beskrives, og behandling av sivile saker etter sivilprosessen blir det ikke redegjort for.

2.2.1 Straffeprosessen

En norsk domstol kan i hovedsak gjøre to typer avgjørelser; dommer og kjennelser. Dommer avslutter en sak og avgjør det materielle i saken, for eksempel om den tiltalte er skyldig. Kjennelser avgjør derimot prosessuelle spørsmål, for eksempel hvilke bevis som skal føres. Anke av kjennelser gjøres i form av kjæremål (saken påkjæres).

I norsk strafferett er hovedregelen at påtalemyndigheten sitter med bevis-byrden i en straffesak. Det heter seg at så lenge det fins en rimelig tvil i skyldspørsmålet, skal tvilen komme den tiltalte til gode. Man antar at tiltalte er uskyldig helt til man eventuelt finner et fellende bevis. Denne uskyldsantagelsen innebærer at dommerne må være oppimot 100 prosent sikre på at tiltalte har begått en forbrytelse før de kan dømme vedkommende.

Johs. Andenæs skriver i første bind av ”Norsk Straffeprosess” at en dom kan bare begrunnes med de bevis som har kommet frem under den muntlige hovedforhandlingen [3]. Dette betyr at man ikke kan dømme noen på grunnlag av å ha studert sakens dokumenter, og man kan ikke supplere muntlige utsagn i hovedforhandlingene med innhold i dokumenter som ikke er nevnt under den muntlige forhandlingen. Dette kalles bevisumiddelbarhet og kommer til uttrykk i straffeprosesslovens § 305.

I dagens norske rettsvesen gjelder prinsippet om fri bevisbedømmelse. Dette prinsippet bygger på at dommeren best kan finne frem til sannheten i en sak dersom hun får bedømme bevisene uten å være bundet av lovregler [3]. Dermed er det opp til den enkelte dommer å gjøre seg opp en mening om bevisene, og ulike dommere vil kunne bedømme bevis på ulike måter.

2.2.2 Sakkyndige

En dommer vil alltid bruke sin egen livserfaring og kjennskap til menneskets natur som grunnlag når hun dømmer i en sak. Av og til trenger hun imidlertid kunnskap om emner brakt opp i retten som ligger utenfor hennes vanlige kunnskapsområde. Da vil retten kunne oppnevne

sakkyndige som typisk har grundig kjennskap til emnet, for eksempel rettsmedisinsk eller teknisk sakkyndige. Selv om partene også kan foreslå sakkyndige og protestere mot oppnevnte sakkyndige, er det i siste instans retten selv som avgjør dette. Dersom de først oppnevnte sakkyndige er uenige eller i tvil i saken de skal uttale seg om, kan retten oppnevne nye sakkyndige.

Det som kommer frem i en sakkyndigs vitneprov er ikke bindende for retten. Siden slike blir oppnevnt nettopp fordi retten selv ikke sitter inne med kunnskap om et tema, er likevel vitneprovet ofte veiledende for retten. Den som avgir vitneprov som sakkyndig forventes å være kyndig og upartisk. I den forbindelse fins det i strpl. § 142 en ugildhetsregel som sier at en person ikke bør oppnevnes som sakkyndig dersom hun ville vært ugild som dommer i saken. Dersom flere sakkyndige oppnevnes, bør de ikke være avhengige av hverandre for eksempel ved at den ene er underordnet den andre i en yrkessituasjon. Siden mange fagmiljøer er små og dette kan være vanskelig å unngå, er heller ikke ugildhetsregelen i § 142 absolutt.

Sakkyndige leverer etter lovens § 299 skriftlig forklaring til retten. Denne kan så leses opp under hovedforhandlingene dersom det blir funnet hensiktsmessig. I tillegg kan partene i saken begjære fremmøte av den sakkyndige etter § 143 for å avgi muntlig forklaring – enten i stedet for å levere skriftlig forklaring eller for å utrede den skriftlige forklaringen nærmere.

Selv om det forventes at sakkyndige sitter med bred og god kjennskap om emnet hun skal gi vitneprov om, er det i dagens norske rettsvesen ingen formelle krav til hvem som kan vitne som sakkyndig. En teknisk sakkyndig som skal uttale seg om en applikasjon, vil naturlig være en som arbeider daglig med den aktuelle programvaren. Proprietær programvare med lukket kildekode vil måtte bli bevitnet av en person som har tilgang til kildekoden – ofte en som er ansatt i firmaet som utvikler programvaren. Denne ansatte og sakkyndige vil da måtte gi en objektiv forklaring om et emne og en aktør hun har et underdanig avhengighets- og maktforhold til. Det kan stilles spørsmål ved hvor objektiv en slik ekspertuttalelse vil kunne være.

2.2.3 Bevis

Et bevis er et middel som benyttes for å godtgjøre en rettslig relevant omstendighet i en sak [5]. Det er tre kategorier bevis:

- Muntlige forklaringer
- Reelle bevis
- Dokumentbevis

Muntlige forklaringer, vitneprov, er av stor betydning som bevis. Siktedes forklaring er viktig, og også andre personers vitneprov anses som vesentlige i en straffesak. Muntlige forklaringer suppleres ofte med reelle bevis eller dokumentbevis. Reelle bevis kalles ofte tinglige bevis, og slike kan for eksempel være gjenstander funnet på åstedet for en kriminell handling.

Det er kategorien dokumentbevis som er særlig interessant for denne oppgaven. Når man skal benytte seg av digitale dokumenter som bevis, er det, som for fysiske bevis, meget viktig å etablere troverdighet rundt beviset. Det er vesentlig at man med sikkerhet kan fastslå et digitalt dokuments egenskaper for å garantere dets integritet og ekthet. Man må også kunne si med sikkerhet om dokumentet i ettertid er endret og i så fall hvor og hvem disse endringene har blitt utført av. Dette er med på å bygge integriteten til beviset.

Bevissikring

I strafferetten kan retten i følge strpl. § 210 pålegge utlevering av en ting som antas å ha betydning som bevis dersom den som har det potensielle beviset plikter å vitne i en sak. Siden siktede ikke plikter å vitne i en sak, trenger heller ikke hun eller hennes pårørende å utlevere potensielle bevis. Retten kan imidlertid gjøre bruk av § 203 og beslaglegge tingen de ønsker å føre som bevis. Dette kan ikke den siktede nekte retten å gjennomføre.

Andenæs påpeker i sin bok at straffeprosesslovens og tvistemålslovens regler om dokumentbevis skiller seg tydelig fra hverandre. Mens tvistemålsloven har nøye nedtegnede regler om fremleggelse, ekthetsspørsmål og beviskraft, fins slike regler ikke i straffeprosessloven. Skriftlige bevis leses opp i strafferetten under hovedforhandlingene, strpl. § 302. Regler for granskning av dokumentbevis dersom det er reist tvil omkring dokumentet fins i strpl. kapittel 12. Siden påtalemyndigheten har plikt til å få en sak best mulig opplyst, fins et insentiv til å ønske å legge frem mange bevis.

Ved bevissikring i straffesaker sier straffeprosesslovens § 197 siste ledd og § 205 første ledd at det må spesifiseres hva som skal sikres. I "Lov og rett i cyberspace" av I. M. Sunde eksemplifiseres dette med at dersom det er en e-postkonto som skal undersøkes, er det denne som skal sikres og ikke datautstyret siktede eventuelt har benyttet [27]. I. M. Sunde påpeker også at det er *ting* som kan beslaglegges, og både fysisk utstyr og data regnes som ting. Ifølge Rt. 1992 er det rettslig avgjort at opplysninger lagret som digitale data og som kan skrives ut, regnes som ting [26, s.905]. Digitale dokumenter inneholder metadata. Det fins funksjonalitet, blant annet i MS Word 2003, for å skrive ut disse metadataene. Derfor kan også denne informasjonen beslaglegges ved sikring av digitale dokumenter.

Avskjæring av bevis

Det er flere tilfeller der et bevis kan være ervervet på ulovlig eller kritikkverdig måte. Av og til kan det være fristende for politiet å innhente bevis uten å ta hensyn til reguleringene i loven, og også forsvarer eller siktede kan fristes til å gå veien utenom loven for å få tak i ønskelige bevis. Også forhørsretten kan gjøre den feil å oppta bevis som ikke er innhentet slik loven foreskriver.

I tilfeller der en av partene mener at bevis er ervervet ulovlig eller på en kritikkverdig måte, kan det reises spørsmål og protesteres ved bevisene under hovedforhandlingene. Retten må da ta stilling til protesten, og beviset blir enten tillatt brukt eller avskåret. Det fins en ankemulighet for å påkjære avgjørelsen til en høyere rettsinstans. Siste instans er Høyesteretts kjæremålsutvalg.

Når retten vurderer om et bevis skal godkjennes eller avskjæres, vurderes det hvilken betydning det har at beviset er ulovlig innhentet. I enkelte tilfeller kan slik bevisinndrivelse svekke bevisets verdi, og også personopplysningsloven kan komme inn i bildet ved at en avskjæring vil ha en beskyttelsesverdi for siktede eller andre parter i en sak.

Dersom et bevis i en straffesak er innhentet i forbindelse med en ulovlig ransakelse fra politiets side, er det ikke noen automatikk i at beviset skal avskjæres. Dersom det under en slik aksjon tas beslag i et reelt bevis i saken, vil bevisverdien være den samme uavhengig av inndrivelsesmetode.

Det fins flere eksempler på saker der retten etter vurdering har besluttet å avskjære fremlagte bevis, og den såkalte Gatekjøkkenkjennelsen fra 1991 er en av dem. Saken gjaldt et gatekjøkken der eieren mistenkte en av sine ansatte for å underslå penger fra kassa. Eieren satte opp videokamera i butikken for å sikre bevis – uten at de ansatte ble varslet. Høyesterett tok hensyn til at bevis som er skaffet til veie på utilbørlig måte kan nektes ført når dette vil medføre en krenkelse av personvernet. Høyesterett uttalte at det ville virke støtende om bevis som er anskaffet på denne måten blir tillatt brukt, og det vil kunne oppfattes som en ny krenkelse av den som blir overvåket" [25, s. 616]. Beviset ble avskåret og nektet ført, og denne dommen har fått stor presedens i senere saker. Kort tid etter at kjennelsen ble avsagt kom et forbud mot skjult fjernsynsovervåkning i straffeloven § 390 b, og denne bestemmelsen er i dag nedfelt i personopplysningsloven kapittel 7 om fjernsynsovervåkning [31].

Analoge dokumentbevis

Av de tinglige bevisene i en sak er det i denne oppgaven dokumentbevisene som er særlig interessante. Ved dokumentbevis er det meningsinnholdet som er relevant. Meningsinnholdet er også det analoge dokumentbevis og digitale dokumentbevis har mest til felles. Analoge dokumenter besitter nemlig en rekke egenskaper som ikke nødvendigvis er å finne i det digitale formatet.

Analoge dokumenter er svært ofte påtegnet dato for opprettelse. Man vil også ofte finne informasjon om forfatteren godt synlig i topp- eller bunntekst, og det fins typisk en bekreftende signatur nederst i dokumentet. Slik informasjon har flere funksjoner – en signatur verifiserer avsender i tillegg til å varsle at dokumentet avsluttes. En signatur medfører også at avsender vanskelig kan nekte å ha sendt ut dokumentet. At dokumentet er trykket på papir øker også dataintegriteten ved at det i stor grad vanskeliggjør modifisering og endring av innhold i ettertid.

Digitale dokumentbevis

Digitale bevis er elektronisk materiale som er samlet inn for å underbygge en rettslig relevant omstendighet [5]. Ikke alle egenskaper man finner hos et analogt dokument fins nødvendigvis hos et digitalt dokument. Metadata knyttet til dokumentet kan gi likevel mye av den samme informasjonen man finner hos et analogt dokument. Denne informasjonen kan imidlertid ligge godt skjult for det sotre flertallet av brukerne og kan ofte være vanskelig å vise i en papirutskrift.

Analoge vs digitale dokumentbevis

Norsk straffeprosess har hatt lang erfaring med og strenge rutiner for innhenting og bevaring av fysiske bevis. Digitale bevis på sin side er et fenomen rettsvesenet først i den senere tid har måttet ta stilling til. Prosessreglene er ikke tilstrekkelig tilpasset for denne bevistypen. Nettopp derfor er det viktig å undersøke konkret hvilke egenskaper digitale dokumenter har og hvordan disse best kan svare på spørsmål knyttet til bevisets ekthet. På hvilke områder er fysiske og digitale bevis like og når må man behandle de to bevistypene forskjellig?

En av de viktigste egenskapene ved et analogt dokument er at man ikke trenger hjelpemidler for å tilegne seg innholdet. Dette gjør papirdokumenter til et viktig medium for informasjonsformidling. Siden papir er et relativt uforanderlig medium, vil heller ikke informasjonsinnholdet endres over tid. Papir har følgelig en viktig bestandighetsfunksjon. Originalfunksjonen til papir er også viktig siden det alltid vil finnes ett opprinnelig originaldokument. I rapporten fra Justisdepartementets kartleggingsprosjekt for å fjerne rettslige hindringer for digital kommunikasjon i

det offentlige Norge, skrives det at i en digital representasjon av et dokument er det mer problematisk skille mellom original og kopi. En kopi av et digitalt dokument vil ikke bli utsatt for et generasjonstap; kopien er like god som originalen. Hvis originalen skal defineres som det mediet informasjonen ble skrevet i første gang, er det ikke mulig å tale om originale digitale dokumenter, tatt i betraktning at mottaker alltid vil få en kopi av dokumentet [13]. Også når det gjelder dokumentfalsk (jamfør kapittel 2.2.4) er original og kopi en vesentlig problemstilling. I en juridisk hovedfagsoppgave om dokumentfalsk skrevet av L. C. Sunde forklares det at det ikke er nødvendig at et originaldokument benyttes for at et dokument regnes som forfalsket. En kopi av et falsk dokument er i like stor grad et falskt dokument som originalen, og en kopi av et digitalt dokument er normalt identisk med originalen [28].

2.2.4 Dokumentfalsk

Datakrimkonvensjonen definerer i artikkel 7 digital dokumentfalsk som ”input, alteration, deletion, or suppression of computer data, resulting in inauthentic data” [1]. Konvensjonen benytter altså ordet ”inauthentic” som betegnelse på at et dokument er falskt. Dette samsvarer med informasjonssikkerhet og beskyttelse av datas autentisitet. I tillegg til å beskytte datas autentisitet verner dokumentfalskreglene i straffelovens kapittel 18 datas integritet.

L. C. Sunde skriver at dokumentfalskreglene ikke først og fremst er til for å beskytte selve bevisførselen i rettssaker. Bevisførselen beskyttes av straffelovens § 132 om bevisforspillelse og kapittel 15 om falsk forklaring. Hensikten med dokumentfalskreglene er å beskytte den alminnelige tilliten til dokumenter. Som bevis kan forfalskede digitale dokumenter stille likt med reelle bevistyper. Dersom dokumentet er forfalsket vil det kunne bli gjenstand for granskning i retten som ethvert reelt bevismiddel.

De viktigste reglene for dokumentfalsk fins i straffeloven §§ 179-186. Bruken av forfalskede dokumenter blir omtalt i § 182 og § 183. Forberedelse på å anskaffe et falskt dokument omtales i § 186, og § 185 gjelder selve forfalskningen eller anskaffelsen. Selve begrepet *dokument* er definert i § 179. Det er ikke avklart i loven om databasert informasjon som for eksempel digitale dokumenter skal regnes som dokumenter og faller inn under reglene om dokumentfalsk. Dette ble imidlertid behandlet i det nye lovforslaget, presentert i NOU 2007: 2 den 12. februar av Datakrimutvalget (jamfør kapittel 3.1.1). Hovedsaklig går forslaget som omhandler dokumentfalskreglene ut på å erstatte begrepet *dokument* med begrepet *bevisbærer*, jamfør utvalgets forslag til ny § 31-1 i ny straffelov.

2.3 Digital etterforskning

Digital etterforskning kan defineres som praktisk bruk av tekniske metoder og verktøy for bevaring, sikring, validering, identifisering, analyse, tolkning, dokumentasjon og presentasjon av digital informasjon.

2.3.1 Organisasjoner og arbeidsgrupper

Etterforskning av digitale medier er et relativt nytt aspekt i rettsaparatet. Willassen og Mjølsnes skriver i artikkelen "Digital Forensics Research" at USA var først ute da FBI etablerte Computer Analysis and Response Team (CART) i 1984 [39, 45, 34]. Ikke før i 1993 ble det startet en arbeidsgruppe innenfor området i Interpol [47]. The International Organisation on Computer Evidence (IOCE) ble så etablert i 1995 [48].

I Norge jobber både offentlige og private aktører innen digital etterforskning. ØKOKRIM fikk i 1995 opprettet en gruppe innenfor datakriminalitet – en gruppe som i 2002 ble Politiets Datakrimsenter (PDS) [49]. Gruppen tar seg av det nasjonale ansvaret for digital etterforskning i politmyndigheten. Den kommersielle firmaet IBAS AS har siden sent på 1990-tallet gjennomført digital etterforskning for både private firmaer og statlige institusjoner [34].

Først de siste årene har det teknologiske fagmiljøet verden over vist særlig interesse for digital etterforskning. Derfor har mange av prosedyrene og arbeidsmetodene en ad-hoc-tilnærming. Nå utvikles det imidlertid nye og bedre teorier og metoder i forskningsmiljøet. I 1998 ble the Scientific Work Group on Digital Evidence (SWIGDE) etablert i USA [54]. Arbeidsgruppen har medlemmer fra mindre politiavdelinger, føderale amerikanske byråer og avdelinger, banker og universiteter.
Gruppen jobber for å sikre kvalitet og konsistens innenfor det digitale og multimedia-relaterte etterforskningsmiljøet. Det internasjonale forskningsmiljøet har også etablert the International Journal on Digital Evidence (IJDE). Dette er et forum for utveksling av idèer, teorier og forskning [46].

2.3.2 Etterforskningens 3 faser

Både amerikanske Brian Carrier og norske Willassen og Mjølsnes deler digital etterforskning inn i 3 faser [4].

1. Sikring av bevis
2. Analyse av bevis
3. Evaluering av bevis

Frem til nå har mesteparten av innsatsen på forskningsområdet blitt lagt til første etterforskningfase. Resultatet er at det nå fins mange

speilkopieringsverktøy og andre lignende metoder for å sikre digitale bevis. Analyse- og evalueringsfasene har midlertid ikke fått særlig fokus, og derfor er dette nå viktige områder for forskning [34].

Sikring av bevis

Første fase, bevissikring, innebærer å lage en nøyaktig kopi av det sikrede digitale mediet. Dette kalles speilkopiering fordi man lager en nøyaktig kopi av for eksempel en hard-disk. Alt digitalt materiale på disken kopieres, inkludert filer, bilder, lydfiler og metadata. Krav som stilles til verktøy som brukes i denne fasen beskrives næremere i kapittel 4.2.3.

I white paperet "Data Integrity Within Computer Forensics" utgitt av SWGDE understrekes viktigheten av å bevare integriteten til sikrede digitale bevis [29]. Hashing er en velbrukt metode for dette, noe også artikkelen til Chat Hosmer som bli nærmere beskrevet i kapittel 3.2.2 diskuterer. I figur 2.1, hentet fra SWGDEs white paper, illustreres fem steg for å sikre dataintegriteten til digitale bevis.

SWGDE forklarer at digital informasjon må sikres både logisk og fysisk for å forhindre uautorisert tilgang til originaldata og etterforskningsmaterialet. Når materialet er sikret, kan integriteten sikres ved å generere hashing-verdier av originalmaterialet. Det genereres også hashing-verdier av det kopierte materialet. Verifisering av at kopien er sunn og komplett gjøres så ved å sammenligne de to hashing-verdiene. Det er viktig at generering og sammenligning av hashing-verdier gjøres før analysefasen starter.

Analyse av bevis

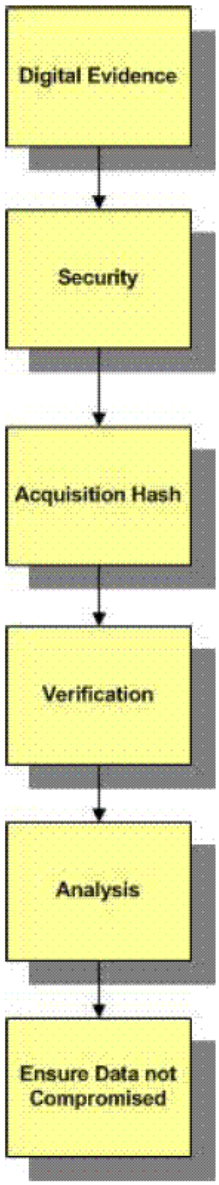
Andre fase, analyse av bevis, beskrives av Willassen og Mjølsnes som en vanskelig fase der man må gjennomgå ett og ett bevis. Vanskeligheten ligger i at det ofte er svært store datamengder som må analyseres. Det er også sjeldent spesifisert på forhånd hvilke data som er av rettslig interesse og som kan ha bevisverdi [34].

I tillegg til å analysere innholdet i filer og filstrukturer i speilkopien, kan slettet materiale gjennopprettes og analyseres. Brian Carrier deler analysematerialet inn i 3 kategorier basert på hvilke bevistyper som analyseres [4]:

1. Skyldsbevis: Bevis som støtter en gitt teori.

2. Uskyldsbevis: Bevis som motsier en gitt teori.

3. Bevis for manipulering: Dette er bevis som ikke kan knyttes direkte til en teori men som beviser at data er manipulert.



Figur 2.1: Fem steg for å bevare integriteten til digitale bevis

Evaluering av bevis

Tredje og siste fase i etterforskningen kalles evalueringsfasen av Willassen og Mjølsnes [34]. I denne fasen skal det avgjøres hvilke følger bevisene har for etterforskningen. Som forfatterne spør i sin artikkel: Hva forteller bevisene oss om bruken av datamaskinen og hvilke handlinger brukeren har gjort?

Det fins ingen definerte rutiner eller retningslinjer i straffeprosessen for hvordan evaluering av digitale bevis skal foregå. Å dele bevismaterialet inn i Carriers tre kategorier over kan være en del av løsningen. Å undersøke et bevis og kartlegge verdikjeden er også en metode for å evaluere beviset og dets potensielle beviskraft.

2.4 Verdikjeden til et bevis

Verdikjeden til et bevis beskriver historien eller livsløpet til et bevis. Den sier noe om hva som hender et bevis, hvem som får hendelsene til å skje og når de skjer eller har skjedd. Når verdikjeden til et bevis er kjent vet man også mer om egenskaper som brukbarhet, integritet og troverdighet. Dette er egenskaper som man potensielt kan få mer kunnskap om ved å undersøke bevisets metadata.

I dagens strafferett fins det ikke rutiner eller retningslinjer for kartlegging av verdikjeden til et digitalt bevis på denne måten. For at en dommer skal kunne anta et bevis, er det likevel viktig at bevisets validitet er dokumentert. Validiteten bygger på autentisiteten til beviset, og denne bygges opp av de tre egenskapene brukbarhet, integritet og troverdighet. Når påtalemyndigheten skal sikre, analysere og presentere bevis bør de derfor legge vekt på disse egenskapene for å få en mest mulig korrekt bevisfremlegging. Det burde også finnes rutiner og retningslinjer for å kartlegge verdikjeden til digitale bevis i norsk strafferett slik at domstolen bedre kan avgjøre validiteten til slike bevis før de antas eller avskjæres.

Tema for denne oppgaven er digitale dokumenter, og det er følgelig verdikjeden til digitale bevis som beskrives her. For å si noe om hva som har hendt med et digitalt dokument, må potensielle hendelser kartlegges. Det må undersøkes hvilke aktører som får hendelsene til å inntreffe og hvordan hendelsene registreres med hensyn til tidspunkt. En del kunnskap rundt disse problemstillingene er kjent, men det forskes fortsatt mye på dette både nasjonalt og internasjonalt. I kapittel 3 omtales blant annet to norske forskningsprosjekter, ”TID – Time stamps in Digital Forensics” (jamfør kapittel 3.2) og ”DESDIFOR – Defining Standards in Digital Forensics” (jamfør kapittel 3.4), der henholdsvis tidsstempling og proaktiv bevissikring er fokus.

Det er et vidt spekter av hendelser som et digitalt dokument kan bli utsatt for. Enkelte hendelser påvirker dokumentet drastisk, mens andre

hendelser har mindre betydning for selve dokumentet og dets innhold. At dokumentet blir opprettet, endret, lagret i ulike versjoner eller slettet er hendelser som påvirker dokumentet mye. Hendelser som at dokumentet skrives ut, sendes som e-post eller anti-virussjekkes er hendelser som i mindre grad påvirker dokumentet. Likevel kan alle disse hendelsene være viktige når man skal avgjøre brukbarhet, integritet og autentisitet.

Type	**Aktør**
Personrolle	Forfatter
	Avsender
	Mottaker
	Redigerer
	Privat firma som rekonstruerer data
	Etterforskningorgan
	Påtalemyndighet
Programvare	Tekstredigeringsapplikasjon
	Filsystem
	Operativsystem
	E-post-klient
	Scrubbers
	XML-editor
	Serversystem
	Saksbehandlingssystem
	Speilingsverktøy
	Analyseverktøy
	Back-up system
	Anti-virusprogram

Tabell 2.2: Potensielle aktører i verdikjeden

For å videre avgjøre hvor relevant og troverdig et digitalt bevis er, er det viktig å kartlegge hvilke aktører som har hatt befatning med hendelsene knyttet til det digitale dokumentet. Det er to typer aktører – personroller og programvare. Med personroller menes faktiske mennesker som har vært i kontakt med dokumentet. Eksempler på

personroller skisseres i tabell 2.2. I tabellen listes det også opp eksempler på programvare som kan fungere som aktører i verdikjeden.

Spesielt viktig for domstolen når bevis skal vurderes, er tidsaspektet. Det er viktig å med sikkerhet kunne fastslå når en konkret hendelse har funnet sted. Når ble dokumentet opprettet? Når ble det endret? Når er ulike versjoner fra? Flere av hendelsene knyttet til et digitalt dokument lagres med et tidsstempel som forteller når dette skjedde. Det er typisk at et digitalt dokument har lagret et tidsstempel for når det for eksempel ble opprettet, når det sist ble endret og når det sist ble skrevet ut.

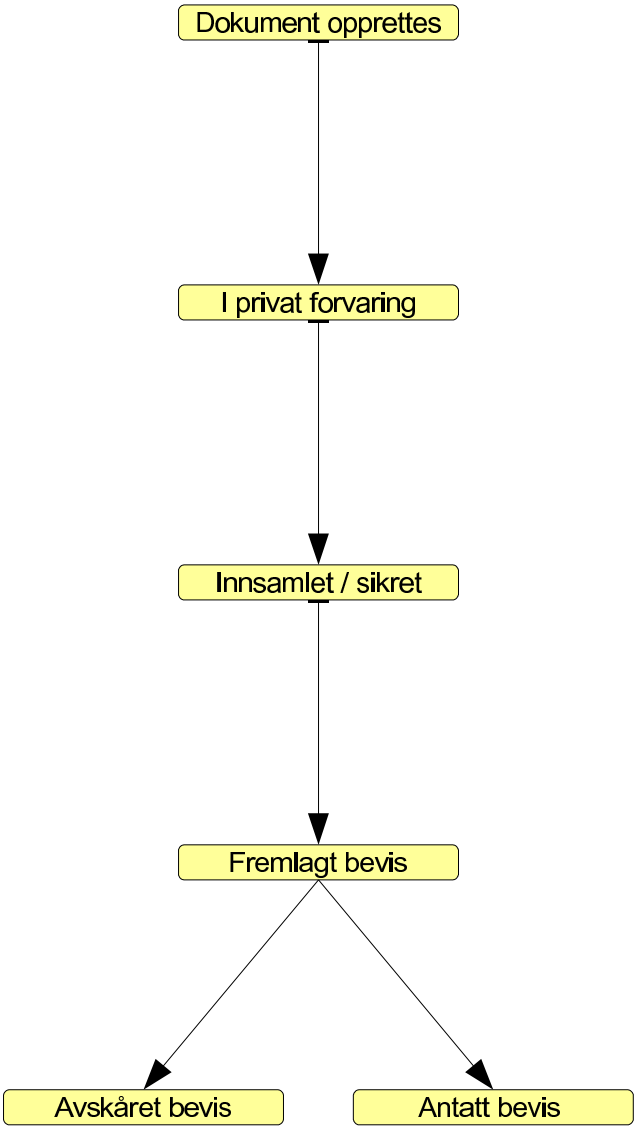
2.4.1 Overordnede faser

Verdikjeden til et digitalt dokument kan på et overordnet nivå deles inn i fire faser – i privat forvaring, innsamlet og sikret, fremlagt som bevis og antatt eller avskåret som bevis. I alle faser er det potensielt flere aktører som har befatning med dokumentet. Mange hendelser kan skje, og det er i varierende grad at tidsspunktet for dette blir lagret med et tidsstempel. I figur 2.2 illustreres verdikjeden til et digitalt bevis. De enkelte fasene vil i det følgende forklares nærmere.

I privat forvaring

I første fase av verdikjeden befinner et digitalt bevis seg som et hvilket som helst digitalt dokument i privat forvaring. Det betyr at dokumentet ikke har fått noen som helst betydning for en konflikt eller rettstvist og ennå ikke har fått status som bevis. Det vil i de fleste tilfeller ikke bli tatt større notis av dokumentet, og det vil ikke bli gjennomført større arbeid for å lagre og sikre det.

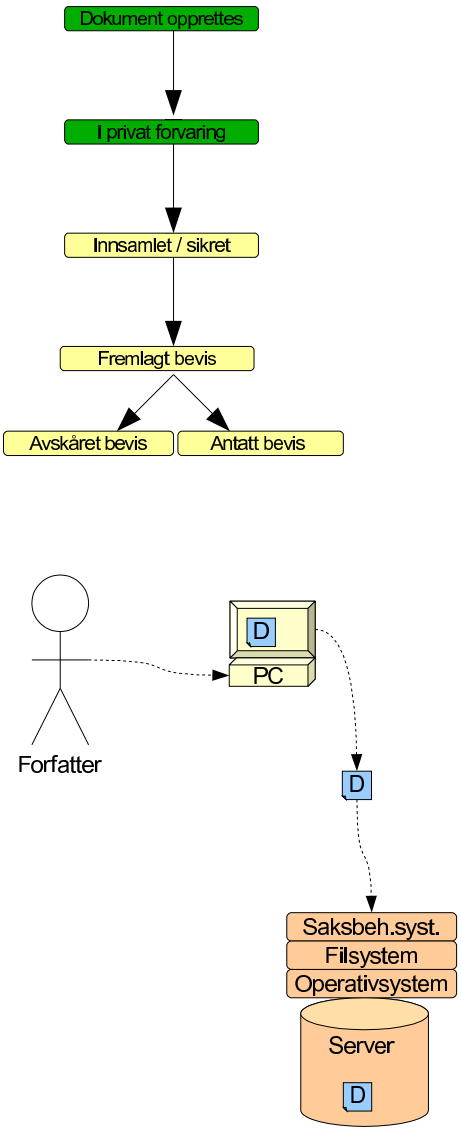
Den første fasen, og samtidig starten på hele dokumentets verdikjede, blir trigget av at det digitale dokumentet opprettes. Dersom man snakker om alle former for digitale formater, som for eksempel databaseinnhold eller datalogger, vil dette kunne skje ved at en applikasjon virker som aktør. Datasystemer, som saksbehandlingssystemer, vil også kunne opprette tekstdokumenter, men siden fokus for denne oppgaven konsentrerer seg om digitale dokumenter fra programvaren MS Word 2003, vil opprettelse av et dokument i all hovedsak skje ved at en fysisk person, en forfatter, oppretter dokumentet. Dokumentet opprettes i et tekstredigeringsprogram og blir håndtert både av et filsystem og et operativsystem. Dersom vi antar at forfatteren oppretter dokumentet på sin arbeidsplass vil også et saksbehandlingssystem kunne være en aktør ved at dokumentet legges inn i systemet. Dessuten er det svært vanlig at dokumentet lagres på en server i tillegg til lokalt på forfatterens datamaskin. Tidspunktet for opprettelsen av dokumentet blir som regel lagret. Dette skjer ved at et tidsstempel lagres i dokumentets metadata. Andre metadata kan være informasjon om forfatter og programvare. Allerede ved den aller første hen-



Figur 2.2: Verdikjeden til et digitalt dokumentbevis

delsen som skjer et dokument har vi altså identifisert 6 aktører som kan påvirke dokumentet: Forfatter, tekstredigeringsprogram, filsystem, operativsystem, saksbehandlingssystem og serversystem (jamfør figur 2.3).

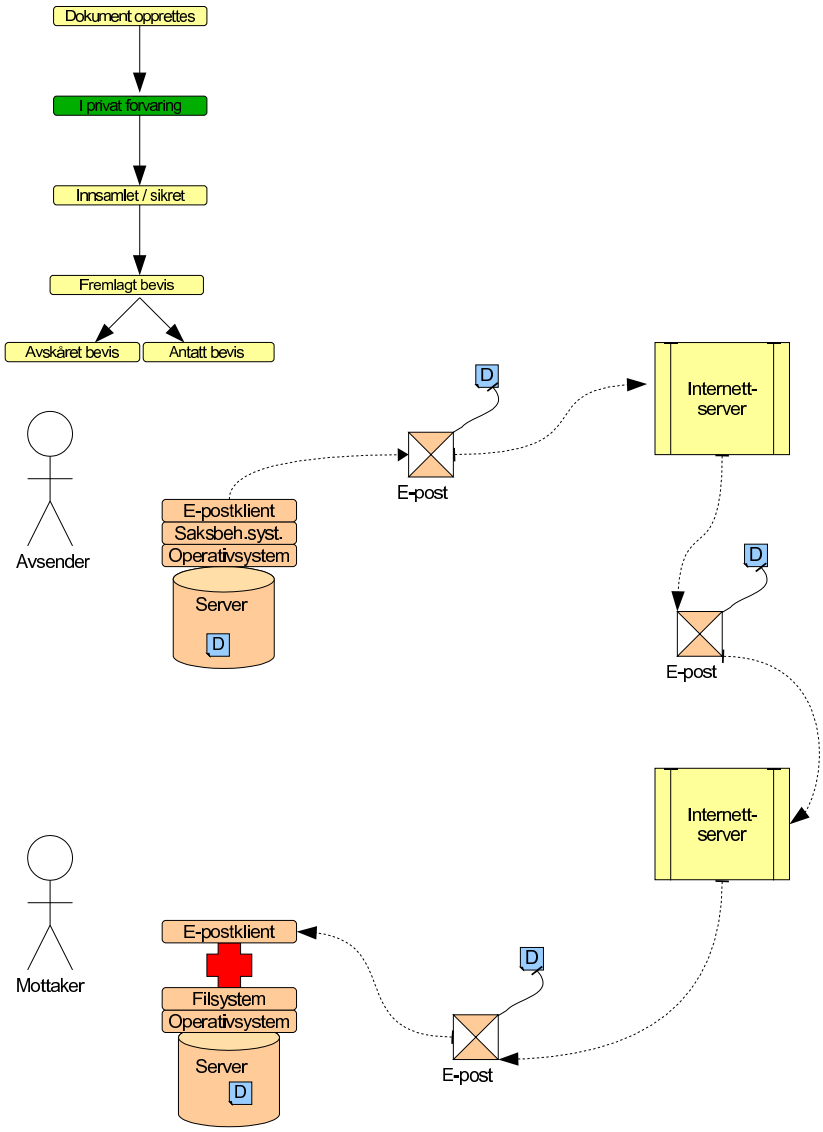
Svært mange hendelser i tillegg til opprettelsen kan skje et digitalt dokument i første fase av verdikjeden. Forfatteren kan endre innholdet i dokumentet flere ganger, og dokumentet kan skrives ut på ulike skrivere. Forfatteren ønsker kanskje også å dele dokumentet med andre, og sender det som vedlegg i en e-post til en mottaker. Både forfatteren som avsender, mottakeren og begges e-postklient vil da virke som aktører som får hendelser til å skje dokumentet. Når det ankommer mottakeren kan et anti-virusprogram trigge hendelser i dokumentet. Slike program går ofte gjennom dokumenter bit for bit ved å plukke



Figur 2.3: Verdikjeden – dokument opprettes

det fra hverandre før det settes sammen igjen. Hvilke aktører som kan spille inn når et digitalt dokument sendes som vedlegg til en e-post illustreres i figur 2.4. Dokumentet er blått og merket med en *D* og befinner seg i utgangspunketet på avsenderens server. Så sendes en e-post med dokumentet som vedlegg via flere internett-servere til en mottaker. Der møter e-posten mottakerens e-postklient, anti-virusprogram, filsystem og operativsystem før det lagres på mottakerens server.

Når dokumentet så lagres på mottakers datamaskin, kan det samtidig lagres på en ny server. Mottakeren kan redigere dokumentet og opprette nye versjoner. Dokumentet kan så sendes tilbake til forfatteren eller videre til andre nye mottakere. Personrolle-aktørene kan benytte seg av programvare, altså programvare-aktører, for å endre dokumentet. Scrubbers kan



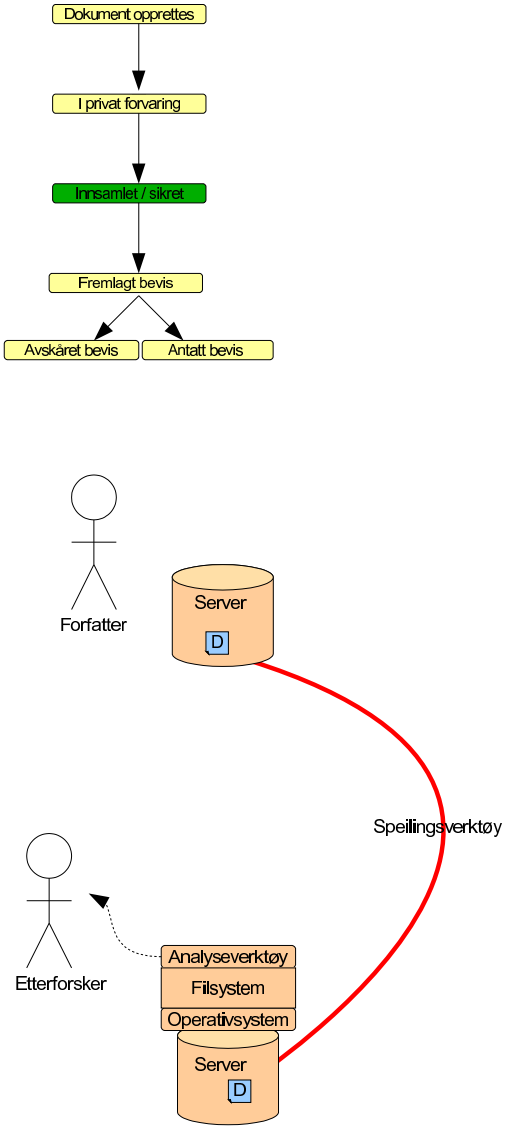
Figur 2.4: Verdikjeden – dokument sendes i e-post

brukes for å endre den lagrede metadata-informasjonen som ligger lagret i dokumentet. Denne programvaren endrer ikke nødvendigvis den synlige informasjonen i dokumentet, men den bakenforliggende informasjonen vil kunne endres.

Innsamlet og sikret

I verdikjedens andre fase blir digitale dokumenter innsamlet og sikret som potensielle bevis. En metode for dette er at private eller offentlige etterforskere speilkopierer hard-disker som mistenkes for å inneholde bevismateriale. Etter speilkopiering blir materialet gjennomgått og analysert for å finne det man er på jakt etter.

I denne fasen er det personrolle-aktører som private gjennopprettings-



Figur 2.5: Verdikjeden – bevis sikres

firma eller etterforskningsorganer som trigger hendelsene som skjer et digitalt dokument. Disse benytter seg av programvare-aktører som server-systemer, speilingsverktøy og analyseverktøy (jamfør figur 2.5).

Striden mellom Datatilsynet og Redningsselskapet fra 2005 gjaldt nettopp sikring av slikt bevismateriale. Redningsselskapet engasjerte IBAS, et privat firma som tilbyr tjenester innen rekonstruksjon av data, for å sikre digitalt bevismateriale i forbindelse med en personaltvist. IBAS speilkopierte servere og sikkerhetskopier av Redningsselskapets informasjonssystem slik som figur 2.5 illustrerer. Slettet materiale fra sikkerhetskopiene ble rekonstruert. I dette eksempelet var IBAS personrolle-aktør, og verktøyet som ble brukt for å speilkopiere var programvare-aktør i verdikjeden.

Det er viktig at sikringsverktøy ikke endrer dataene for at for at det digitale materialet skal kunne legges frem som bevis. Det debatteres om programvare for dette bør være skrevet i åpen eller lukket kildekode. I kapittel 4.2.3 kan man lese om Brian Carriers forskning. Han har undersøkt dette og mener at programvare med åpen kildekode har et fortrinn. Hvordan selve etterforskningen foregår i denne fasen beskrives i kapittelet "Etterforskningens 3 faser", kapittel 2.3.2.

Fremlagt som bevis

I denne fasen legger en av partene i saken frem det digitale dokumentet som bevis. Retten tar stilling til om dokumentet skal brukes som bevis eller om det må avskjæres ved å ta stilling til dokumentets brukbarhet, integritet og autentisitet. Dette gjøres basert på hva man vet om beviset – hvilken informasjon om dokumentet fins utover dets meningsinnhold?

Dersom man på dette stadiet kjenner det digitale bevisets verdikjede, kan man danne et bilde av hvilke hendelser beviset har vært utsatt for, hvilke aktører som har trigget hendelsene og når hendelsene har funnet sted. Noe av denne informasjonen ligger lagret i dokumentets metadata. Som vi har sett lagres for eksempel tidspunktet for når et dokument ble opprettet som et tidsstempel. I dag fins det dessverre få retningslinjer og rutiner for å kartlegge verdikjeden til digitale dokumenter. Dersom slike rutiner hadde vært etablert kunne domstolen i større grad benyttet seg av verdikjeden for å avgjøre validitet og relevans.

Når metadataene skal brukes for å avgjøre bevisets validitet, er det viktig at man kjenner til hvordan informasjonen genereres, lagres og endres. Hvordan metadata håndteres i ett tekstbehandlingsprogram, MS Word 2003, blir undersøkt senere i denne masteroppgaven.

I straffeprosessen har påtalemyndigheten som oppgave å belyse saker. Bevisene er i påtalemaktens bevaring, og de vil da lagres på en server, og i den hendelsen vil et serversystem være aktør. Dersom det er en sivil sak, er det de to partene som hver for seg legger frem bevis. Bevisene er da i deres bevaring, og befinner seg på respektives datalagringsmedium, for eksempel hard-disk. Da er lagringsmediet og filsystemet aktører.

Avskåret eller antatt som bevis

Dersom domstolen ikke finner det riktig å anta et bevis, blir det avskåret. Dette kan for eksempel skje dersom beviset ikke er relevant eller ikke har nok troverdighet. Videre hendelser som inntreffer beviset vil da ikke være relevant for verdikjeden.

Dersom domstolen finner at det digitale dokumentet har sikker validitet og det er relevant for saken, kan det antas som bevis i rettstvisten. Beviset vil fortsatt befinne seg på påtalemyndighetens hard-disk eller server i straffesaker og i en av partenes hard-disk eller server i sivilsaker.

2.5 Oppsummering av kapittelet

I dette kapittelet har en rekke viktige begreper blitt definert og spesifisert. Videre har jeg gitt en overordnet beskrivelse av norsk straffeprosess.

Vi har sett at det er påtalemyndigheten som sitter med bevisbyrden i straffesaker, og at den har plikt til å få en sak best mulig opplyst. Dette medfører et insentiv til å legge frem mange bevis, men det er opp til dommeren å godta eller avskjære bevis. Det er viktig å skape troverdighet omkring bevis som legges frem, og man må være sikker på at beviset er ekte. Kapittel 18 i straffeloven omhandler dokumentfalsk, men det er imidlertid ikke avklart om digitale dokumenter faller innenfor dokumentfalskreglenes virkeområde.

Sakkyndige brukes dersom emner som blir brakt for retten ligger utenfor dommerens vanlige kunnskapsområde. Retten oppnevner den sakkyndige, men dennes vitneprov er ikke bindende, kun veiledende, for retten. Det er ingen formelle krav til hvem som kan vitne som sakkyndig, men det forventes at denne er kyndig og upartisk.

Det er tre faser i digital etterforskning – sikring, analyse og evaluering. Ved sikring av digitale bevis lages en speilkopi av mediet der potensielle bevis er lagret. Hashing er en vanlig metode for å sikre at kopieringen ikke forandrer dataene. I analysefasen må bevis sorteres, og det er både kostnads- og tidkrevende å analysere de ofte store datamengdene som blir sikret. I evalueringen avgjøres hva bevisene kan fortelle om bruken av datamaskinen de var lagret på og hvilke handlinger som har blitt foretatt.

Verdikjeden til digitale bevis beskriver historien eller livsløpet til et bevis. Den sier noe om hva som hender et bevis, hvilke aktører som får hendelsene til å skje og når de skjer. Aktørene deles inn i to grupper: Personroller og programvare. Det burde finnes rutiner og retningslinjer for å kartlegge verdikjeden til digitale bevis i norsk strafferett slik at domstolen bedre kan avgjøre validiteten til slike bevis.

Kapittel 3

Relatert forskning

I dette kapittelet gir jeg en oversikt over forskningsprosjekter og forskning relatert til digitale dokumenter i Norge, Europa og USA. Prosjektene beskrives med både fokusområde og forskere i hvert enkelt avsnitt.

3.1 Datakrimutvalget

Datakrimutvalget ble opprettet ved kongelig resolusjon i 2002. Utvalget, ledet av sorenskriver Knut Rønning, har medlemmer fra ulike institusjoner som arbeider med datakriminalitet – Christina Christensen fra Samferdselsdepartementet, Hanne Gulbrandsen fra Datatilsynet, advokat Birthe Taraldset, statsadvokat Jenny Sellæg, forsker Inger Marie Sunde og doktorgradsstipendiat ved NTNU Svein Y. Willassen.

3.1.1 Lovtiltak mot datakriminalitet

Utvalget skal etter mandat utrede lovtiltak mot datakriminalitet. Første delutretning ble publisert i Norges Offentlige Utredninger (NOU) 2003: 27 "Lovtiltak mot datakriminalitet – Delutredning I". Her tok utvalget for seg Europarådets konvensjon som omhandler bekjempelse av kriminalitet som knytter seg til informasjons- og kommunikasjonsteknologi og hvilke endringer i norsk rett som var nødvendige for å ratifisere Europarådets datakrimkonvensjon fra 2001 [1, 20]. Utfallet av første delutreding var at konvensjonen trådte i kraft med virkning for Norge 1. oktober 2006.

Andre delutredning fra datakrimutvalget ble presentert 12. februar i NOU 2007: 2 [21]. I rapporten kommer utvalget med forslag til bestemmelser som gjelder datakriminalitet som kan tas inn i en ny straffelov.

Utvalget har sett på om det bør fastsettes særskilte straffebestemmelser for datakriminalitet eller om dette i stedet skal inkluderes i de øvrige straffebestemmelsene. Konklusjonen er at det bør legges til et eget kapittel som gjelder datakriminalitet i den nye straffeloven: Vern av data,

databasert informasjon og datasystemer. Utvalget mener at dette vil gjøre det enklere å sette seg inn i strafferegulering. Det andre alternativet, å inkludere datakriminalitet i de eksisterende straffebestemmelsene, mener de derimot vil kunne gjøre reguleringen unødvendig komplisert og utilgjengelig. Utvalget har hovedsakelig lagt vekt på "reglenes innbyrdes sammenheng, felles begrepsbruk, pedagogiske formål, behovet for oversiktlighet og det antatte oppdateringsbehovet" når de har vurdert dette.

I rapporten kommer det frem at det er de profesjonelle aktørene, som for eksempel nett-tilbydere, som har ansvar for å tilrettelegge for sikker bruk av datatjenester og digital kommunikasjon. Private brukere er avhengig av dem for å sikre seg mot virusangrep og lignende.

Datakrimutvalget presiserer videre i sin rapport at det er viktig at norsk lovgivning er på linje med sammenlignbare lands lovgivning. Dersom lovgivningen her er veldig avvikende, kan norske datatjenester være svært attraktive for kriminelle som ønsker å utnytte en eventuell svakere lovgivning.

Dokumentfalsk

I hvilken utstrekning falske digitale data skal vurderes etter samme regler som falske papirdokumenter, er ikke en problemstilling som datakrimutvalget skal utrede. Utvalget mener imidlertid at dette må klargjøres og skisserer enkelte regler som kan være aktuelle. Det er flere forslag om i hvilket kapittel regulering av forfalskning av digitale data skal inn i. I stedet for å omtale digital informasjon i kapittelet om dokumentfalsk slik det er i dagens lovgivning, kan det opprettes en egen paragraf om dette i datakapittelet. Alternativt kan dokumentfalsk reguleres som i dag, men at anvendelsesområdet for bestemmelsene presiseres.

Utvalget mener at et nytt begrep som skal erstatte *dokument* i reguleringen av dokumentfalsk må introduseres. *Bevisbærer* er et godt begrep i så måte. Dette begrepet vil dekke både formater som enkelt lar seg kategoriseres, som e-post og tekstfiler og andre filtyper, som innhold i databaser og datalogger. *Bevismidler* eller *bevisrepresentasjoner* er andre begreper som vil inkludere bevistypene over og vil derfor også være gode løsninger.

I utkastet med forslag til ny straffelov er det stor uenighet omkring flere spørsmål, både blant medlemmene i utvalget og i det juridiske og det teknologiske fagmiljøet. Utkastet til § 11 som omhandler befatning med skadelige dataprogrammer og -utstyr, og forslaget om harmonisering av visse bestemmelser i straffeloven og åndsverksloven er to av emnene det strides om. Flertallet av medlemmene i utvalget går inn for utkastet til § 11. I harmoniseringsspørsmålet er alle medlemmene enige i at regelverket bør forenkles. Det er imidlertid ikke enighet i utvalget om harmoniseringen

bør utredes videre før den eventuelt gjennomføres.

Det tredje spørsmålet datakrimutvalget har et delt syn på er utkastet til § 76b, som omhandler filtrering av nettsteder. Et mindretall av medlemmene er for at dette lovforslaget skal godttas.

Filtrering og stenging av nettsteder

Stenging av nettsteder er et tiltak for å gjøre det vanskelig for norske brukere å få tilgang til ulovlig informasjon og tjenester. For å hindre dette kan nettsteder eller servere som befinner seg på norsk jord og som er eller har vært brukt i straffbare handlinger stenges. Eksempler på slike straffbare handlinger er nettsteder der det foregår ulovlig nedlasting av film eller musikk, nettsteder med mulighet for pengespill og gambling eller nettsteder som har befatning med barnepornografi.

§ 69 i den nåværende straffeloven regulerer stenging av nettsteder. Dersom informasjon og tjenester kommmer fra utenlandske tilbydere og servere, har imidlertid ikke norske myndigheter mulighet til å gjennomføre slik stenging. Som en løsning på dette foreslår to av utvalgets medlemmer, Christensen og Rønning, at det skal åpnes for at nett-tilbydere skal pålegges å filtrere tilgangen til enkelte nettsteder. Slik filtrering kan foregå på nasjonalt nivå eller på tilbydernivå. Kina er kjent for slik filtrering. Der kontrollerer myndighetene all internett-trafikk. Utvalget påpeker at Norge ikke har infrastruktur til å kunne gjennomføre noen slik form for kontroll. Filtrering må skje på tilbydernivå.

Norske nett-tilbydere har allerede utviklet funksjonalitet som filtrerer bort nettsider anbefalt av KRIPOS. Filtreringen gjøres uten lovbestemmelser og stenger ute nettsteder som omhandler barnepornografi. Siden ordningen er frivillig og ikke alle norske nett-tilbydere deltar, er ordningen bare delvis effektiv. Det er også svakheter ved at treff-sikkerheten ikke er maksimal. Filtrering vil også blokkere materiale som ikke er i strid med norsk lov og gi såkalte falske positiver. Dette fore-kommer fordi man filtrerer den konkrete IP-adressen som sender ut ulovlig materiale istedet for å filtrere de uønskede dataene.

Mindretallet i utvalget foreslår en lovhjemmel som pålegger filtrering nettopp fordi dagens filtreringsordninger baseres på frivillighet og på samarbeid med nett-tilbyderne. Forslaget til ny paragraf i straffeloven § 76b har følgende ordlyd:

> "Tjenesteyter kan pålegges å blokkere tilgangen til bestemte steder på internett for sine brukere dersom innholdet ville kunne medføre straffansvar utover bøter i Norge. § 69 tredje ledd og § 76a gjelder tilsvarende. De øvrige regler om inndragning gjelder tilsvarende så langt de passer."

Målet er at denne lovhjemmelen skal gi samme effekt for utenlandske nettsteder som stenging av norske nettsteder har. Filtreringstiltak skal skje ved dom som følge av den prosessuelle fremgangsmåten som gjelder ved stenging av nettsteder med tilhørighet til Norge. Mindretallet er enig med resten av utvalget i at dette ikke gir 100 prosent effekt, men at tiltaket vil gi en betydelig forbedring av dagens situasjon. Likevel mener flertallet i utvalget at det ikke er grunn til å foreslå en egen lovhjemmel for filtrering av nettsteder.

Flertallet mener at hensynet til ytringsfriheten taler mot å sensurere utenlandske nettsider for norske brukere. Ytringsfriheten vil begrense hvilke tiltak som kan treffes. Falske positiver vil bli stoppet og flertallet i utvalget mener at hensynet til ytringsfriheten bør veie tungt. Det er vanskelig å forutsi omfanget av filtrering når denne virker på hele servere eller datamaskiner, og ikke bare på selve dataene som motstrider norsk strafferett. Likevel mener Christensen og Rønning at hensynet til ytringsfriheten ikke taler mot sensur av utenlandske nettsider med straffbart innhold. De to er ikke enige i at den frivillige friltreringsordningen for nettleverandører som allerede fins gjør tilgangen til nettsteder som for eksempel inneholder seksualiserte skildringer av barn vanskelig nok.

Ett av datakrimutvalgets medlemmer, Svein Y. Willassen, teknolog og ph.d.-student ved NTNU (jamfør kapittel 3.2), er sterkt imot mindretallets forslag om å innføre en egen lovparagraf som kan pålegge nettleverandører å stenge utenlandske nettsteder. Han mener lovforslaget innebærer statlig sensur av internett. Han mener også at utvalgets fokus er på feil sted når det gjelder å hindre ulovlig materiale på nettet. Det bør heller handle om å finne frem til hvor materialet spres fra enn å hindre det å komme over norske landegrenser. Dersom man kan avsløre kilden til spredningen vil det i de fleste tilfeller ikke være store problemer knyttet til å få lokale myndigheter til å gripe inn, hevder han.

Det er også problemer knyttet til filtrering på tilbydernivå dersom dette skulle bli aktuelt. Utover filtrering av falske positiver vil også effektiviteten av filtreringen være lav. Dersom man må gjennom prosessuelle fremgangsmåter hver gang et nettsted skal sensureres, vil ikke sensurorganet raskt nok kunne forfølge de kriminelle etterhvert som de eventuelt skifter IP-adresser og lokasjoner å distribuere ulovlig materiale fra. Willassen påpeker i tillegg at norske brukere relativt enkelt kan omgå filtreringen ved å benytte seg av utenlandske servere. Disse kan fungere som proxy-servere, det vil si at det ulovlige materialet sendes via denne serveren til Norge. Proxy-serveren vil ha en annen IP-adresse enn den opprinnelige avsenderens adresse og vil ikke bli fanget opp av den norske nett-tilbyderens filtre fordi den kun fanger opp den opprinnelige IP-adressen.

3.2 TID – Time stamps in Digital Forensics

Ved NTNU pågår det et forskningsprosjekt ledet av Svein Yngvar Willassen – Time stamps in digital forensics (TID). Med støtte fra Forskningsrådet er prosjektet en viktig aktør innenfor forskning rundt digital etterforskning og er det første i sitt slag i Norden [56, 33].

3.2.1 Forskningsprosjektets mål

Prosjektet dreier seg om tidsstempling i datasystemer. Som Willassen og hans kollega fra NTNU, Stig Frode Mjølsnes, påpeker i artikkelen "Digital Forensics Research" er det lite systematisk dokumentasjon av eksisterende formater på tidsstempling. Dokumentasjonen er også mangelfull innenfor den faktiske bruken av tidsstempling, tidssoner og ulike kalendere i datasystemer [34].

Forskningsprosjektet TID ble påbegynt i 2005. Over tre år vil prosjektet gjennomføre analyse og dokumentasjon av eksisterende tidsreferansesystemer, tidssoner og formater samt beskrive hvordan systemoperasjoner påvirker tidsstempler. Det blir også utviklet metoder, programvare og verktøy for å oppdage feilaktige tidsstempler. Et viktig hovedmål med prosjektet er i tillegg å teste og validere metodikk og programvare. Testene skal gjøres både med konstruerte data og med faktisk innhentet digitalt bevismateriale.

ØKOKRIM er en nøkkelsamarbeidspartner i prosjektet. Det private firmaet IBAS AS deltar også, i likhet med Fast Search and Transfer. Disse to internasjonalt aktive aktørene bidrar til at prosjektet tar hensyn til internasjonale krav som blant annet skyldes ulike jurisdiksjoner. Ved Purdue University i Indiana, USA, pågår det også et stort forskningsprogram innen informasjonssikkerhet, CERIAS. Samarbeid med dette forskningsprogrammet sikrer at TIDs resultater er nyttige også i internasjonal sammenheng. Resultater vil bli publisert på internasjonale konferanser i regi av organisasjoner som Digital Forensic Research Workshop, International Organisation on Com-puter Evidence, European Network of Forensic Science Institutes, Scientific Working Group on Digital Evidence (SWIGDE) og Interpol European Working Party on Information Technology Crime [42, 44, 48, 54].

3.2.2 Tidsstempling

Datasystemer og applikasjoner tidsstempler systemoperasjoner eller dokumenter som for eksempel opprettes, endres eller slettes. Tidsstemplene kan brukes til å forsterke eller svekke troverdigheten til et digitalt bevis. For å knytte en rekke prosesser i et digitalt medium til hendelser som skjer i den virkelige verden, må man danne en tidslinje, en

verdikjede, av de digitale hendelsene. Til dette kan etterforskere benytte seg av tidsstemplene.

Kjente problemer

I artikkelen "Digital forensics research" forklarer Willassen og Mjølsnes at problemene til etterforskere av digitalt materiale er mange. Det er lite systematisk dokumentasjon av ulike formater for tidsstempler. Det er i tillegg mangel på dokumentasjon av den faktiske bruken av tidsstempler og hvilke tidssoner det opereres med. Verden over brukes det ulike referansesystemer, som for eksempel ulike kalendere. Det er heller ikke noen standard for antall sekunder som hoppes over ved for eksempel skuddår. Hvordan operativsystemer og filsystemer håndterer tidsstempling ved ulike operasjoner er heller ikke et avklart spørsmål, men det er kjent at disse kan manipulere og endre tidsstempler. Tidsstempler er knyttet opp mot en tidskilde, ofte systemklokker eller nettverkstjenester, som ikke kan regnes for pålitelige kilder.

Kjente løsninger

Chat Hosmer skriver i artikkelen "Proving the integrity of digital evidence with time" at mange metoder som brukes for å bevise integriteten til digitale bevis har blitt hentet fra informasjonssikkerhet og fra informatikk og blitt brukt i domenet til digitale bevis. Tre av disse metodene er sjekksum, en-veis hash-algoritme og digital signatur [11].

Sjekksum er en metode for å sjekke feil i digitale data. Typisk legges en 16- eller 32-bits polynomial til hvert byte av digitale data som man vil beskytte. Resultatet er små heltallsverdier som er 16 eller 32 bits lange og som representerer konkateneringen av dataene. Heltallsverdiene må lagres og sikres. På ethvert fremtidig tidspunkt kan det samme polynomialet legges til dataene for så å sammenligne resultatet med det opprinnelige resultatet. Hvis resultatene er like, fins det en viss grad av integritet.

En-veis hash-algoritme er en metode for å beskytte digital data mot uautoriserte endringer. Metoden lager en bestemt lengde av heltallsverdier (mellom 80 og 240 bits) som representerer de digitale dataene. Metoden har to unike egenskaper: For det første, er det vanskelig å konstruere nye data som gir samme hash. For det andre, er det vanskelig å finne andre data som passer nøyaktig til samme hash-verdi.

Den tredje metoden Hosmer tar for seg for å bedre integriteten til digitale data er digital signatur. Det er sikker måte å knytte identiteten til den som signerer til de digitale dataene. Metoden bruker et krypteringssystem basert på en offentlig og en privat, abstrakt nøkkel. Den som signerer bruker en privat nøkkel for å generere den digitale

Metode	Fordeler	Ulemper
Sjekksum	• Lett å beregne • Bruker liten datalagringsplass • Bra å bruke for å oppdage tilfeldige feil • Rask	• Lav sikkerhet mot ondsinnede angrep • Enkelt å lage nye data med samme sjekksum • Må vedlikeholde sikker lagring av sjekksum • Binder ikke identitet med data • Binder ikke tid med data
En-veis hash-algoritme	• Lett å beregne • Kan oppdage både tilfeldige feil og onsinnede endringer	• Må opprettholde sikker lagring av hashverdier • Binder ikke identitet med data • Binder ikke tid med data
Digital signatur	• Binder identitet til integritetsoperasjoner • Forhindrer uautorisert regenerering av signatur med midre privat nøkkel er offentliggjort.	• Langsom • Må beskytte den private nøkkelen • Binder ikke tid med data • Hvis nøklene kompromitteres eller sertifikat går ut kan signaturen bli ubrukelig.

Tabell 3.1: Ulike metoder for å sikre digital integritet.

signaturen, og enhver kan validere signaturen ved å bruke den offentlige nøkkelen.

Å bevise integriteten til digitale bevis blir vanskeligere når man tar med tidsperspektivet. Ved å bruke den beste metoden som fins i dag, digital signatur, klarer vi å koble sammen elementene hvem og hva – signaturen og dataene. Hosmer legger vekt på at bruk av digital signatur likevel ikke gir oss svar på to viktige spørsmål:

1. Når ble det digitale beviset signert? Hvor lang tid etter at beviset ble sikret ble integriteten beskyttet?

2. Hvor lenge kan vi bevise integriteten til det digitale beviset vi signerte?

For å svare på disse spørsmålene er tid en kritisk faktor. Man må bestemme hvordan man kan knytte et tidspunkt til en handling. En pålitelig kilde for tidsstempelet må dessuten identifiseres og standardiseres. For å bevise når en hendelse har inntruffet i den digitale verdenen,

må man bli enige om hvordan tidsstempler brukes. Dette krever videre at det etableres nye standarder for sikre og kontrollerbare tidsstempler som representeres digitalt. Når man har laget et tidsstempel som er motstandsdyktig mot manipulasjon og gir en sporbar tidsrekke, kan man digitalt knytte disse tidsstemplene til digitale bevis slik at de kan bli verifiserbare.

Det ideelle tidsstempel

Et ideelt sikkert og sporbart tidsstempel bør i følge Hosmer ha følgende egenskaper:

Nøyaktighet Tiden som presenteres er fra en autoritær kilde og er så nøyaktig med hensyn til presisjon som det kreves: Dags-, times- eller millisekunds-presisjon.

Attestering Tidskilden er attestert til et "nasjonalt måleinstitutt" (eng: National Measurement Institute) slik at en tredjepart kan verifisere nøyaktigheten og presisjonen.

Integritet Tiden skal være sikker og ikke kunne være utsatt for manipulering gjennom vanlig behandling. Dersom den blir utsatt for dette, enten forsettlig eller uforsettlig, skal manipuleringen være synlig for en tredjepart.

Ikke-tilbakeviselig En hendelse skal være knyttet til et tidsstempel slik at koblingen mellom hendelsen og tidstempelet ikke kan tilbakevises på et senere tidspunkt.

Ansvarlighet Prosessen det er å innhente tidsstempel, legge til attestering og integritet og å knytte det til den aktuelle hendelsen skal være ansvarlig slik at en tredjepart kan avgjøre at passende prosess ble brukt og at manipulering ikke har forekommet.

Hosmer mener altså at man må bedre nøyaktigheten og påliteligheten til digitale tidsstempler drastisk. Tidsstemplene må regelmessig bindes til digitale data. Rutinene og prosessene i den forbindelse må gjøres utvetydige og standardiseres i hele den digitale verden, og digitale tidsstempler må være mulig å verifisere mot en rettslig gyldig tidskilde.

3.3 Cybex

Cybex er et spansk firma som har spesialisert seg på konsulenttjenester knyttet opp mot elektroniske[1] bevis. De er med i alle deler av prosessen rundt slike bevis – fra planlegging og innhenting til analysering. De stiller også i retten for å presentere elektroniske bevis og er sakkyndige i saker der slike bevis fremlegges. På deres nettsider hevder de å være ledende på markedet både nasjonalt og internasjonalt [40].

Cybex kommer med en rekke tekniske råd når det gjelder håndtering av elektronisk bevismateriale. Det fokuseres konkret på å kartlegge verdikjeden til elementene man finner i tillegg til å kartlegge rekken av hendelser som har intruffet disse elementene (eng: chain of events) – opprettelsestidspunkt, sist skrevet ut og lignende. Å ivareta validiteten, autentisiteten og integriteten nevnes også som en middel for å sikre at de elektroniske sporene kan brukes som bevis.

Det pågikk i 2006 et omfattende prosjekt i regi av Cybex, "The admissibility of electronic evidence in court: fighting against high-tech crime" (AEEC), der representanter fra 15 EU-land samt Romania deltok. Hovedmålet med prosjektet var å uvikle en undersøkelse som sammenligner den nåværende status innen jurisdiksjon og digital etterforskning i disse landene. Før dette prosjektet startet var det ikke gjennomført undersøkelser av denne størrelsen på tvers av landegrensene i Europa.

3.3.1 The admissibility of electronic evidence in court: Fighting against high-tech crime

Prosjektet forsøkte å gi svar på spørsmål som: Hva er elektroniske bevis? Er elektroniske bevis regulert i europeiske jurisdiksjoner? Hvilke problemer møter europeiske etterforskere, advokater og dommere som er involvert i sikring, analyse, evaluering og presentering av elektronisk bevis. Hvordan arbeider de i de ulike fasene?

Første steg i prosjektet var å analysere den eksisterende lovgivningen omkring elektroniske bevis i hvert av deltakerlandene. Neste steg var å gjennomføre 8 personlige intervjuer med sivile personer, kriminelle personer, offentlige og private advokater, ledere i advokatforeninger og høyeste autoritet innen påtalemakten i hvert land.

Gjennomgangen av lovgivningen i de ulike landene viste at det verken fins en eksplisitt referanse til *elektroniske bevis* eller at uttrykket er spesifikt definert. Det ble imidlertid funnet at det i alle landene forekommer reguleringer med forskrifter som på en eller annen måte refererer til

[1] *Elektroniske* bevis er en videre definisjon enn *digitale* bevis siden elektroniske apparater og systemer i tillegg kan prosessere annen teknikk enn den digitale, for eksempel analog teknikk [15].

elektroniske bevis. Ingen av landene har en konkret definisjon av hva elektroniske bevis er, men det ble funnet referanser som er mer eller mindre konkrete for tradisjonelle bevistyper og som også inkluderer elektroniske bevis.

Analyse av innholdet i lovgivningen viser at elektroniske bevis håndteres som tilsvarende tradisjonelle bevis i alle deltakerlandene. Det ble funnet tre typer av ekvivalens mellom bevistypene:

- Likhet mellom elektronisk dokument og korresponderende papirutskrift
- Likhet mellom elektronisk og håndskrevet signatur
- Likhet mellom e-post og tradisjonell post

Et stort flertall av europeiske dommere ser elektroniske bevis som ekvivalente til de tradisjonelle bevistypene. Videre anser flertallet elektroniske bevis for å være ekvivalente med dokumentbevis, selv om enkelte av dommerne som er intervjuet i prosjektet ikke har dette synet og ikke tillegger dem samme beviskraft som tradisjonelle bevistyper. Videre gjengis det forskerne skriver i sin rapport om hva intervjugruppene tenker om fordeler og ulemper ved elektroniske bevis [12].

Fordeler ved elektroniske bevis

De intervjuede dommerne som deltok i Cybex' prosjekt tolker fordeler og ulemper ved elektroniske bevis ulikt når det gjelder bevisenes troverdighet (jamfør tabell 3.2). Noen dommere mener at deres objektivitet og nøyaktighet i håndteringen av elektroniske bevis gjør bruken mer riktig og mener slike bevis med fordel kan brukes. Andre dommere mener imidlertid at vanskelighetene forbundet med å verifisere autentisiteten til elektroniske bevis gjør bruken mindre troverdig enn tradisjonelle bevis. Disse mener derfor at det er en ulempe å benytte seg av slike bevis.

Blant fordelene dataeksperter og teknikere nevner er at elektroniske bevis tilbyr informasjon som er nøyaktig, komplett, sann og objektiv – gitt at beviset kommer fra en elektronisk innretning som det ikke er knyttet noen form for subjektivitet til, i motsetning til for eksempel vitneutsagn. Disse intervjuobjektene viser videre til at elektroniske medier gir dem tilgang til informasjon som har vært umulig å oppdrive tidligere.

Ulemper ved elektroniske bevis

Advokater og dommere uttrykker en frykt for sårbarhet og hvor enkelt det er å manipulere elektroniske bevis. Noen mener at elektroniske bevis er for tekniske til at de forstås av dommere og påtalemyndigheter, og det uttrykkes derfor en motvilje mot å bruke denne typen bevis i retten.

Fordeler	
Informasjon	Nøyaktig, komplett, klar, presis, sann, objektiv og nøytral
Bevis	Solid, nyttig, troverdig, levedyktig, vesentlig for å bevise typer kriminalitet som tidligere ikke har kunnet bli bevist
Enkelt	Innsamling, bruk, lagring og sikring

Tabell 3.2: Fordeler ved elektroniske bevis i AEEC-undersøkelsen

De anser også problemene forbundet med bevaring og mangelen på standarder og regelverk for sikring av bevisene som ulemper (jamfør tabell 3.3).

Dataeksperter påpeker mangelfull lovgivning og fraværende sertifiseringsmodeller som ulemper ved elektroniske bevis. De mener det er vanskeligere å få gjennomslag for elektroniske bevis i retten fordi dommere ofte ber om større garantier for validitet enn det tradisjonelle bevis møter. Denne gruppen mener mangelen på forståelse og kunnskap blant lovgiverne og -utøverne utgjør en ulempe. De ser på digital etterforskning som en tids- og arbeidskrevende prosess. Store kostnader knyttet til prosessen kan forsinke eller forhindre bruk av elektroniske bevis.

Elektronisk materiale som elektronisk bevis

Cybex' forskningsprosjekt har undersøkt lovgivningen i de deltakende landene og sett på hvilke regler som gjelder for å godta eller avskjære elektroniske bevis. Gjennom intervjuer ble det funnet at elektronisk materiale ikke kan ekskluderes av rettsinstansen uten gyldig motiv.

I Europa er det to modeller for hvilke krav som stilles til bevis som legges frem i retten. Den ene gruppen land har det til felles at den rettslige tradisjonen gir veldig vide krav for å tillate bevis. Her baseres bevisfremleggingen på at dommeren står fritt til å vurdere bevis ettersom de legges frem. I denne gruppen finner vi land som Østerrike, Danmark, Sverige og Finland, og det er i denne gruppen Norge ville befunnet seg dersom forskningsprosjektet også hadde omfattet Norge.

Den andre gruppen av land har en mer restriktiv lovgivning, med en rekke krav til beviset og til hensikten med beviset. Ett av kravene er respekt for fundamentale rettigheter som personvern og rettigheter ansatte har. Andre krav dommeren vil stille for å avgjøre om et elektroniske bevis skal tillates eller avskjæres er troverdighet, relevans og at det er det best tilgjengelige beviset på et gitt tidspunkt.

Ulemper
Mangel på passende og systematisk regulering
Lite rettsvitenskap på området
Ukjent og meget teknisk materiale – få eksperter
Krever spesifikk kunnskap
Vanskelig å presentere i retten på en lettfattelig måte
Vanskeligere å få aksept av retten: Dommere ber om flere garantier enn ved annet bevismateriale
Høye kostnader for å undersøke og tolke informasjon
Vanskelig å vite hvordan data skal tolkes og hvordan spesifikke prosesslover skal tolkes
Vanskelig å bevise autentisitet, troverdighet og opprinnelsen til data
Vanskelig å bevare, sikre og lagre
Enkelt å manipulere
Vanskelig å etablere rettsverdi for beviset
Mangel på rettslig støtte og sertifiseringsmodeller

Tabell 3.3: Ulemper ved elektroniske bevis i AEEC-undersøkelsen

Forbedringsforslag

Gjennom forskningsprosjektet til Cybex har dommere, advokater og tekniske eksperter fått uttale seg om hvilke forbedringer de mener vil hjelpe etterforskningen der elektroniske bevis er relevant. Spesielt advokater mener det vil være til stor hjelp med bedre nasjonale retningslinjer for elektronisk bevis. Spesielt prosedyrer for bevaring, sikring, analyse og presentasjon av slike bevis slik at de kan legges frem i rettsaker på lik linje med tradisjonelle bevis er etterspurt. Det uttrykkes også et behov for internasjonale direktiver og prosedyrer som kan sikre samarbeid mellom stater når det gjelder innsamling og bevaring.

Eksperter innen dataetterforskning mener også at spesifikke reguleringer på et nasjonalt nivå er svært nyttig. Enkelte etterlyser også innføring av retningslinjer for å beskytte fundamentale personvernrettigheter ved innsamling, bevaring og presentering. I likhet med advokatene som er intervjuet, mener de med teknisk bakgrunn at selv et minimum av internasjonale retningslinjer vil forbedre samarbeidet på tvers av landegrensene.

3.4 Defining Standards in Digital Forensics

Prosjektet Defining Standards in Digital Forensics (DESDIFOR) er et prosjekt i regi av Norsk Regnesentral og Institutt for Rettsinformatikk ved Universitetet i Oslo med det som formål å øke bevisstheten om digitale bevis i norske virksomheter. Prosjektet startet i 2004 og utarbeider en standard for å tilrettelegge datasystemer for innsamling og oppbevaring av potensielle digitale bevis slik at det er i overensstemmelse med norsk lovgivning [30].

DESDIFOR jobber med *digital forensic readiness* (DFR): Å fokusere på å tilrettelegge og konfigurere organisasjoners systemer til å proaktivt og strukturert samle og bevare potensielle digitale bevis før en tvis oppstår. DFR fokuserer på en rekke forretningsrisiki ved å avdekke bevismateriale og å motvirke kriminalitet av typen svindel eller informasjontyveri. Ved å jobbe proaktivt øker man tilgjengeligheten og kvaliteten på materiale som trengs for å etterforske hendelser i etterkant. Analysefasen i digital etterforskning er avhengig av det arbeidet som er gjort på forhånd ved hjelp av DFR. Analysearbeidet vil også få lavere kostnader dersom man jobber med potensielle bevis på forhånd. Å samle inn og bevare data på forhånd beskytter integritet og metadata [2].

Proaktiv digital etterforskning reiser en rekke spørsmål med hensyn til personvern. Forskerne som jobber med DESDIFOR er seg bevisst disse spørsmålene. Utsetting av tjenester (eng: outsourcing) kan i tillegg føre til flere utfordringer, og dette blir stadig mer aktuelt etter som stadig flere sikkerhetstjenester også settes ut.

DESDIFORs hovedmål er å definere et rammeverk for proaktiv sikring av digitale bevis. Generelle krav til en strukturert fremgangsmåte for å gjennomføre DFR blir listet opp her: [2]:

- En analyse av rettslige krav til innsamling og bevaring av potensielle digitale bevis.
- En metode for å analysere organisasjoners behov for digitale bevis.
- Identifisering og klassifisering av potensielle beviskilder.
- Utarbeiding av retningslinjer for å bevare digitale bevis, inkludert prosesser, prosedyrer og forslag til hvordan teknologiske løsninger kan brukes.
- Veiledning om når og hvordan hendelser bør rapporteres til myndighetene, inkludert innhold og format på rapportene samt kriterier for rapportering. Standardisering av kommunikasjon og samarbeid mellom berørte parter og myndighetene.

Digital forensic readiness er delt inn i to faser – analyse og implementasjon. Analysefasen starter med en analyse av kriminalitet og potensielle risiki. Mulige beviskilder identifiseres før krav til innsamling av bevis analyseres. Så gjennomføres en analyse av rettslige krav til slik innsamling og organisasjonelle aspekter ved disse kravene. I implementasjonsfasen gjennomføres selve innsamlingen og bevaringen av potensielle digitale bevis.

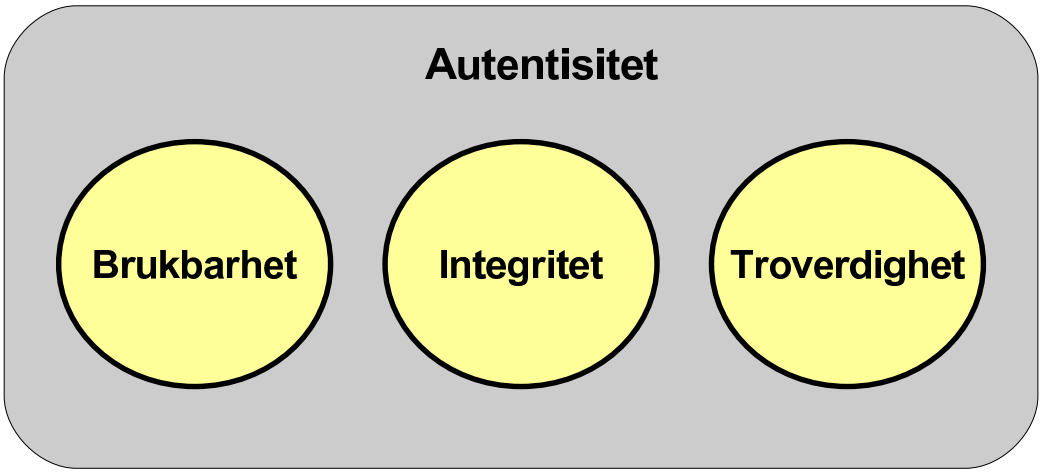
3.5 ARMA International Educational Foundation

ARMA International Educational Foundation er en non-profit organisasjon som jobber for utveksling av kunnskap om informasjonshåndtering. Organisasjonen finansierer forskning og utdanning innen dette området og er et forum der organisering, vedlikehod, gjenhenting, bevaring og sletting av informasjon diskuteres og utvikles [36]. Organisasjonen deler inn sine aktiviteter i tre områder:

- Støtte til forskning som tar for seg kritiske problemer innen informasjonshåndtering
- Økonomisk støtte til undervisning og utdanning for å øke kunnskap og evner til fagpersoner som jobber innenfor feltet
- Fremme og utvikle utdanningsinitiativ som utvikler informasjonshåndtering og treningsprogrammer.

Blant forskningsprosjektene stiftelsen støttet i 2006 finner vi prosjektet ”Proof of the Integrity of a Digital Document Introduced as Evidence” som Stephen Mason ledet. I rapporten basert på prosjektet diskuteres rettslige krav for å fremlegge elektroniske dokumenter som bevis i retten – med vekt på sivilrett i England og Skottland. Mason peker på hvilke betraktninger som bør gjøres dersom en av partene i en sak ikke ønsker beviset avskåret [16]. I Storbritannia er det partene i en sak som fremstiller hver sine sakkyndige. Slik er det ikke i den norske sivilretten. Her kan partene foreslå sakkyndige og protestere mot oppnevnte sakkyndige, men det er i siste instans retten selv som oppnevner dem.

Mason skriver at en rekke tekniske karakteristikker bør kartlegges når digitale bevis legges frem. Forfatteren av et dokumentbevis bør identifiseres. Det bør også fastslås om det fins flere versjoner av dokumentet og hvilken versjon som i så fall skal regnes som original. Det er viktig å bestemme om tidsstempling er brukt, hvilket format som er brukt, og om det er pålitelige tidsstempler. Det bør også tas i betraktning hvor enkelt eller vanskelig det er å sikre, analysere og presentere dataene.



Figur 3.1: Autentisitet

Mason beskriver autentisiteten til et dokument som bygget opp av disse tre hovedelementene: Brukbarhet, integritet og troverdighet (jamfør figur 3.1).

3.6 Digital Evidence Research Programme

Prosjektet "Proof of the Integrity of a Digital Document Introduced as Evidence" som ARMA International Education Foundation støttet, førte til at viktige spørsmål som har både praktiske og rettslige konsekvenser ble reist. Derfor ble prosjektet innlemmet i forskningsprogrammet "Digital Evidence Research Programme", som Mason leder ved British institute of international and comparative law (BIICL) [38, 41].

3.6.1 Forskningsmål

Dette forskningsprosjektet er et langsiktig prosjekt med flere delmål. Tekniske begrensninger for å etablere integritet skal kartlegges, og prosjektet vil se på hvordan dette virker inn på bevisførsel. Metoder utviklet av IT-industrien for å løse problemene som involverer integriteten til digitale dokumenter vil bli undersøkt. I tillegg vil nye problemer som måtte oppstå i forbindelse med nye produkter og programvare som potensielt kan produsere digitale bevis bli undersøkt.

Prosjektet har også som delmål å etablere en forståelse av de rettslige kravene fra de ulike jurisdiksjonene omkring i verden. Fokus vil være på hvordan digitale bevis tillates brukt i retten og hvordan disse blir behandlet. Det vil også undersøkes hvordan lover og prosedyrer tolkes av dommere og hva dette har å si for bruken av digitale bevis.

Det vil bli undersøkt hvor stor kunnskap dommere, advokater og akademikere som underviser i juss har om digitale bevis. Prosjektet søker også å samarbeide med utdanningsinstitusjoner for å dekke manglende kunnskap og foreslå forbedringer i faglig pensum. Forskningsresultatene skal publiseres i akademiske journaler, bøker og på andre passende arenaer. Forskningsfokuset er akademisk, internasjonalt og sammenlignende.

Stephen Mason er i tillegg assosiert med Cybex sitt prosjekt "Admissibility of Electronic Evidence in Court Programme" (jamfør kapittel 3.3) og bidro med resultater til rapporten som dette prosjektet la frem i 2006.

3.6.2 The Admissibility and Disclosure of Electronic Evidence

BIICL jobber også for tiden med et bokprosjekt – "The Admissibility and Disclosure of Electronic Evidence". Dette er en bok som skal ta for seg strafferettslige og sivilrettslige problemstillinger knyttet til fremlegging og avskjæring av elektroniske bevis. Fagfolk fra en rekke land vil bidra, og rettsregler i de ulike deltagerlandene vil bli sammenlignet. Stephen Mason er redaktør for dette bokprosjektet.

3.7 Oppsummering av kapittelet

En rekke organisasjoner og prosjekter har blitt presentert i dette kapittelet. Det er stor enighet i at det er stor mangel på standarder og rammeverk for teorier og teknologi som dreier seg om digital etterforskning og sikring av digitalt materiale. Dette forsøker forskere verden over å gjøre noe med, og i dette kapittelet har vi sett hvordan dette arter seg både i Norge, EU og USA. En oversikt over prosjektene og de enkelte fokusområder listes opp i tabell 3.4 på side 47. I tabell 3.5 på side 48 listes organisasjonene og respektives prosjekter og geografiske tilhørighet opp.

Her i Norge presenterte Datakrimutvalget nylig sitt forslag til ny straffelov. Det ble blant annet foreslått hvordan man kan definere digitale dokumenter i straffeloven. Prosjektet DESDIFOR er nå i avslutningsfasen og har sett på hvordan man proaktivt kan sikre digitalt materiale før eventuelle rettstvister oppstår.

Samtidig har Stephen Mason, med støtte fra den amerikanske organisasjonen ARMA, sett på hvilke rettslige krav for å legge frem digitale bevis som eksisterer. Han beskriver autentisiteten til et dokument som bygget opp elementene av brukbarhet, integritet og troverdighet (jamfør figur 3.1). Masons forskning har blitt utvidet til også å omfatte tekniske begrensninger, forståelse av rettslige krav og kartlegging av kunnskap blant dommere og advokater om digitale bevis. Mason

Forskningsprosjekt	Fokus
Time stamps in Digital Forensics	Tidsstempling
The admissibility of electronic evidence in court: Fighting against high-tech crime	Rettslig regulering av elektroniske bevis
Proof of the Integrity of a Digital Document Introduced as Evidence	Rettslige krav, integritet og autentisitet
Digital Evidence Research Programme	Rettslige krav, integritet og autentisitet
Defining Standards in Digital Forensics	Standard for håndtering av digitale bevis

Tabell 3.4: Forskningsprosjekter og fokus

samarbeidet også med prosjektet den spanske organisasjonen Cybex nettopp har avsluttet. Det har hadde fokus på å kartlegge lover, regler og prosedyrer for digitale bevis i flere EU-land.

Tekniske begrensninger er også fokus for Svein Y. Willassen og prosjektet han leder ved NTNU. I samarbeid med noen av norges ledende aktører innenfor digital etterforskning jobber han for å kartlegge bruken av tidsstempling og tidsreferansesystemer. Prosjektet vil også utvikle metoder, verktøy og programvare for å avdekke feilaktige og falske tidsstempler.

Prosjektene som er presentert i dette kapittelet er interessante og relevante for problemstillingen i denne oppgaven nettopp fordi de tar for seg de tekniske utfordringene knyttet til digitalt bevismateriale og fordi de er enige om at lovverket ofte ikke er oppdatert for denne bevistypen.

Type	Navn	Område
Statlig organ	Datakrimutvalget	Norge
	Teknologirådet	Norge
Forskningsprosjekt ved NTNU	TID – Time stamps in digital forensics	Norge/ internasjonalt
Forskningsgruppe ved Interpol	Interpol European Working Party on Information Technology Crime	EU
Forskningsprosjekt ved Cybex	The admissibility of electronic evidence in court: Fighting against high-tech crime	Spania/EU
Forskningsprosjekt ved Purdue University	CERIAS	USA
Forskningsprosjekt ved Norsk Regnesentral	Digital Evidence Research Programme (DESDIFOR)	Norge/ Sverige
Stiftelse	ARMA International educational foundation	USA
Konferanse	Digital forensic research workshop	USA
Arbeidsgruppe	Scientific Workgroup on Digital Evidence	USA
Organisasjon	International Organisation on Computer Evidence	Internasjonalt
	European Network of Forensic Science Institutes	EU

Tabell 3.5: Organisasjoner, prosjekter og tilhørighet

Kapittel 4

Metode og verktøy

I dette kapittelet beskrives ulike metoder og verktøy jeg benytter i denne oppgaven for å svare på problemstillingen beskrevet i avsnitt 1.2. I avsnitt 4.1 presenterer jeg forskningsmetodene. Her beskrives generering av testdata og fremgangsmåter som skal brukes for å finne svar på de ulike underproblemstillingene. I avsnitt 4.2 beskrives ulike typer verktøy som kan brukes for analysering og manipulering av digitale dokumenter. Både verktøy som skal benyttes i denne oppgaven og verktøy som generelt blir brukt av offentlige og private etterforskningsorganer presenteres.

4.1 Metode

Problemstillingen i denne oppgaven er definert slik:

P: Kan metadata sikre validiteten til digitale dokumentbevis?

For å besvare problemstillingen vil jeg undersøke underspørsmålene definert i avsnitt 1.2 ved å gjennomføre testene beskrevet i dette kapittelet.

4.1.1 Generering av testdata

Det er nyttig å gjøre ulike forsøk og analyser for å undersøke hvordan metadata oppfører seg når man jobber med dokumenter opprettet og behandlet i Microsoft Word 2003. Det vil derfor bli opprettet to testdokumenter i denne programvaren. Fokus vil være på metadata og hvordan denne informasjonen behandles. Målet med testene er å undersøke om en vanlig bruker uten videre innsikt i MS Word kan manipulere digitale dokumenters metadata. Testene vil ikke gi svar på hvordan avanserte hackere eventuelt kan manipulere digitalt materiale. Derfor vil det kunne finnes andre måter å manipulere dokumenter på enn de metodene som blir forsøkt brukt i denne oppgaven.

Testdokument *A*

I MS Word 2003 blir det opprettet et dokument, *A*, som ikke inneholder informasjon utover eventuelle metadata programvaren selv genererer. Dokumentet inneholder ingen tekst og en eventuell utskrift ville vært en tom side.

Testdokument *B*

Det blir også opprettet et annet testdokument, *B*, i MS Word 2003. Testdokument *B* vil inneholde tekstinformasjon som man finner i typiske dokumentbevis.

Følgende data skrives ut dersom man velger default utskriftsformat:

```
Tittel: Testdokument B
Opprettet: 14. februar 2007
Forfatter: Astri Ek Larsen
Sist endret: 02. april 2007
Sist skrevet ut: ikke skrevet ut.
```

4.1.2 Metadatafunn

P_1: Hvilke metadata lagres?

Delproblemstilling P_1 skal besvares ved hjelp av flere metoder:

Microsofts dokumentasjon om MS Word 2003 vil kunne gi relevante opplysninger. Det fins både bruker- og utviklerdokumentasjon, og det av denne informasjonen som er tilgjengelig vil bli analysert. MS Word 2003 har funksjonalitet for lagring på XML-format. Når dokumenter lagres på XML-format brukes ett eller flere skjema for å beskrive strukturen på XML-dokumentet. Skjemaene kan gi svar på hvilke metadata som lagres. Et dokuments egenskaper (eng: properties) vises i MS Word 2003 ved å følge menyvalget *File -> Properties*. Informasjonen som vises for de to testdokumentene vil brukes for å gi svar på hvilke metadata som lagres.

Metadata Analyzer vil benyttes for å se hvilke metadata som lagres (jamfør kapittel 4.2.1).

4.1.3 Generering av metadata

P_2: Hvilke aktører genererer metadata?

Dokumentasjon av hvordan MS Word 2003-dokumenter lagres på XML-format vil undersøkes for å gi svar på denne delproblemstillingen. Testdokumentene vil så lagres på XML-format og analyseres i en XML-editor.

4.1.4 Manipulering av metadata

P_3: Hvordan kan metadata manipuleres?

P_4: Hvilke aktører kan manipulere metadata?

For å finne svar på hvordan og hvem som kan manipulere metadata vil det gjøres tester med testdokumentene og deretter gjøres en analyse av resultatene.

Enkelte metadata-elementer har åpne felter som kan fylles ut i MS Word 2003s egenskaper-visning som følger av menyvalget
File -> Properties. Det vil bli testet hvilke elementer som kan manipuleres på denne måten. Dersom de manipulerte elementene innholder den nye informasjonen etter at programvaren har lagret, lukket og åpnet dokumentet på nytt, er metadataene manipulert av forfatteren.

Metadata som er generert i testdokumentene skal forsøkes manipulert ved å bruke funksjonaliteten i MS Word 2003 som gir lagring på XML-format. Ved å åpne et Word-dokument lagret i XML i en XML-editor skal metadata endres før dokumentet lagres og åpnes i MS Word 2003s vanlige programvare. Så skal metadataene som eksisterer på dette tidspunktet sammenlignes med metadataene som kom frem ved analysen i avsnitt 4.1.2.

Pc'en som testene og analysene gjøres på har et operativsystem med dato/klokke-funksjon. Denne skal manipuleres for så å analysere hvordan dette påvirker metadata lagret i de to testdokumentene. Det vil også undersøkes hvordan metadataene påvirkes av at de lokale versjonene av testdokumentene midlertidig fjernes fra pc'en. Testdokumentene vil midlertidig lagres på et annet medium, før de fjernes fra testpc'en og så flyttes tilbake til pc'en fra det midlertidige mediet.

4.2 Verktøy

Alle tester vil bli gjort med fokus på metadata generert i Microsoft Word 2003, som er en del av Microsoft Office Professional Edition 2003. En engelsk versjon benyttes, og det medfører at informasjonen fra programvaren, også metadata, vil være på engelsk. Testene vil bli gjort i operativsystemet Microsoft Windows XP Professional Version 5.1.2600 Service Pack 2 Build 2600. Pc'en forsøkene vil bli utført på tilhører selskapet daVinci Consulting AS. For midlertidig lagring av testdokumenter på annet medium benyttes en 1. generasjons Apple iPod Nano.

Aktuelle verktøy for dokument-analyse og -manipulering listes opp i tabell 4.1.

Type	Verktøy
Analyse	MS Word 2003
	Metadata Analyzer
	EnCase Forensic
Manipulering	MS Word 2003
	Exchanger XML Editor

Tabell 4.1: Aktuelle verktøy for dokumentanalyse og -manipulering

4.2.1 Analyseverktøy

MS Word 2003

Digitale dokumenter som er generert i MS Word 2003 er fokus i denne oppgaven. Denne programvaren kan brukes for å analysere hvilke metadata som kan genereres, hvor de genereres og om de kan manipuleres. Programvaren har funksjonalitet for å lagre dokumenter på XML-format, for å lese dokumenter på dette formatet og for å vise metadata som er lagret. Dette gjør applikasjonen til en relevant førstehåndskilde for å svare på oppgavens problemstilling.

Metadata Analyzer

Metadata Analyzer er en programvare som analyserer metadata i MS Office-dokumenter. Ved hjelp av et grafisk grensesnitt viser programvaren hvilke elementer og hvilke verdier som er lagret [50]. Metadata Analyzer skal brukes i denne oppgaven for å hente frem potensielle lagrede metadata fra de to testdokumentene.

EnCase Forensic

EnCase Forensic er programvare laget spesielt for å etterforske data-kriminalitet og brukes av offentlige og private etterforskere over hele verden. Den blir også brukt til etterforskning i Norge, blant annet av ØKOKRIM, og programvaren er pensum ved Politihøyskolen.

I EnCase Forensic fins funksjonalitet for å automatisere tidkrevende etterforskningsoppgaver som speilkopiering, søk etter og analyse av sikret bevismateriale. Programvaren har i tillegg funksjonalitet som lar brukeren selv programmere skript og makroer tilpasset spesielle behov. Programmeringsspråket som brukes heter Enscript programming og er basert på metoder i Java og C++. Programvaren kan brukes på ulike plattformer og støttes både av Windows, DOS og Linux. Et bredt spekter av filsystemer og

e-postsystemer kan undersøkes av programvaren – også nettleserbaserte e-postsystemer.

I EnCase Legal Journal beskrives det hvordan EnCase Forensic fungerer [23]. Når programvaren brukes for å sikre og analysere elektronisk informasjon, starter først en boot-prosess som sikrer at materialet på datamaskinen som skal undersøkes ikke endres ved sikring. Etter at boot-prosessen har startet, lages en speilkopi av disken som skal undersøkes. Kopien er en komplett sektor-for-sektor-kopi av alle dataene som fins på den undersøkte disken – også informasjon om slettede filer.

For å sikre integriteten til speilkopien, brukes som regel programvare som genererer matematiske verdier kalt MD5-hashingverdi. Verdien baseres på det nøyaktige innholdet i speilkopien. Det genereres også hashing-verdier av originalmaterialet. Verifisering av at speilkopien er sunn og komplett gjøres så ved å sammenligne de to hashing-verdiene. Hvis ett databit i kopien på en eller annet måte endres, det vil si at så lite som en bokstav eller et mellomrom endrer seg, endrer hashing-verdien seg. Dersom hashing-verdien fra originaldisken og det kopierte bildet er den samme, er det sikkert at de elektroniske dataene ikke er blitt endret ved kopiering. I tabell 3.1 i kapittel 3.2.2 forklarte Hosmer fordeler og ulemper ved denne metoden å sikre digital integritet på.

Mange tror at når man sletter digital informasjon fjernes den fra datamaskinen for alltid. Det stemmer ikke nødvendigvis. Slettet informasjon fins fortsatt på datamaskinen, men diskområdet der den er lagret er nå merket for å kunne overskrives. Det betyr at informasjonen befinner seg på disken helt til annen informasjon blir lagret på samme sted. Dersom informasjonen som er slettet ikke er skrevet over med andre data, kan den rekonstrueres. EnCase Forensic har funksjonalitet for dette i tillegg til funksjonalitet for å speilkopiere, generere og sjekke MD5-hashing-verdier, samt å analysere det sikrede materialet.

De fleste tekst- eller bildefiler inneholder en veldefinert filsignatur som er unik for den aktuelle filtypen. Det gjør det mulig å gjenkjenne filtypen uavhengig av filendelsen. Å endre filendelsen til en fil medfører ikke at filsignaturen automatisk endres. Den vil fortsatt tilsvare den opprinnelige filtypen. Dersom filendelsen og filsignaturen ikke stemmer overens, har man noe som kalles filsignatur-mismatch. EnCase Forensic har funksjonalitet som kan identifisere slike filsignatur-mismatcher.

EnCase Forensic er et nyttig verktøy for å avdekke manipulasjon av dokumenter og metadata. Til dette prosjektet har det imidlertid blitt brukt en demo-versjon av programvaren. Det har medført at det kun har vært mulig å se eksempler på EnCase Forensics funksjonalitet og ikke gjøre tester på testdokumentene. Fordi EnCase Forensics brukes i stor skala over hele verden og også både privat og offentlig i Norge, anses det likevel som relevant for lesere av denne oppgaven at de kjenner til programvarens funksjonalitet.

4.2.2 Manipuleringsverktøy

MS Word 2003

I avsnitt 4.2.1 så vi at MS Word kan brukes for å se på og analysere hvilke metadata som lagres. I tillegg kan programvaren brukes til å undersøke hvordan metadata genereres og manipuleres. I menyvalget *File -> Properties* er det muligheter for endring av dokumentets egenskaper, og denne funksjonaliteten testes ut i denne oppgaven.

Exchanger XML Editor

En XML-editor er en applikasjon som man kan åpne, lese, redigere og endre filer på XML-format i. I tillegg til støtte for å redigere XML, er det vanlig at slike applikasjoner også har funksjonalitet for endring av relaterte filtyper, som XML-skjema, DTD, XPath og XSLT. Dette er teknologi som brukes sammen med XML-dokumentene.

I denne oppgaven vil Exchanger XML Editor Lite versjon 3.2 benyttes både for å analysere og manipulere testdokumentene. Dette er en gratis og ikke-kommersiell javabasert XML-editor og som kan brukes på flere plattformer [57].

4.2.3 Åpen eller lukket kildekode

I kapittel 2 så vi hvilke rettslige krav som gjelder når digitalt materiale skal beslaglegges. Krav til verktøy som brukes for sikring og analyse er også viktige. I "Lov og rett i cyberspace" påpeker I. M. Sunde imidlertid at det i norsk strafferett ikke stilles krav til kunnskap om teknisk programvare eller etterforskeres tekniske kompetanse [27].

Programvare kan enten ha kildekode som er åpen og tilgjengelig for alle eller lukket kildekode som kun invidde har innsyn i. I det amerikanske rettsapparatet må et verktøy være troverdig og relevant for å kunne fremme et vitenskapelig bevis. Troverdigheten til et bevis blir testet ved å bruke en prosedyre kalt Daubert-prosedyren. Testene gjøres ved å vurdere kilden til et bevis, det vil si programvaren som har produsert beviset, opp mot retningslinjene.

Daubert-prosedyren definerer 4 kategorier retningslinjer:

- Testing:

  Kan prosedyrene verktøyet bruker testes?

  Har prosedyrene blitt testet?

  Er prosedyrene både sunne og komplette?

- Feilrate:

  Har prosedyrene en kjent feilrate?

- Publisering:

  Har prosedyrene blitt publisert?

  Har prosedyrene vært gjenstand for fagfellevurdering?

- Aksept:

  Er prosedyrene akseptert i forskningsmiljøet?

Opphavet til digitale bevis produsert av proprietær programvare blir sjeldnere undersøkt og vurdert opp mot Daubert-retningslinjene enn det programvare med åpen kildekode blir. Det viser en gjennomgang av amerikansk rettspraksis av Erin E. Kenneally. Hun skriver i sin artikkel "Gatekeeping Out Of The Box: Open Source Software As A Mechanism To Assess Reliability For Digital Evidence" at spesielt når et proprietært verktøy er kilden til et digitalt bevis – et bevis som ofte har stor beviskraft dersom det antas – må det stilles spørsmål ved troverdigheten til programvaren [14].

Dette er imidlertid det motsatte av hva som skjer i amerikansk rettspraksis. Proprietære verktøy nyter ofte godt av å være markedsledende og velkjente. Programvare med åpen kildekode blir derimot oftere undersøkt med hensyn til korrekthet og troverdighet. Kenneallys undersøkelser viser at det oftere må innkalles ekspertvitner for å kontrollere kildekode når programvaren har åpen kildekode enn når programvarens kildekode er lukket.

Analyseverktøy for bruk i etterforskning med henholdsvis åpen og lukket kildekode har også blitt vurdert av Brian Carrier, en amerikansk forsker og systemutvikler. I artikkelen "Open Source Digital Forensic Tools - The legal argument" diskuterer han hvordan programvare med åpen kildekode fungerer og håndteres i rettsvesenet i forhold til programvare med proprietært lukket kildekode [4]. Carrier har vurdert programvare med de to typene kildekode opp mot retningslinjene i Daubert-prosedyren. Konklusjonen er at open source-verktøy tydeligere og mer omfattende oppfyller retningslinjene enn det proprietære verktøy gjør. Dette skyldes at kildekoden er åpen og tilgjengelig og dermed kan vurderes av alle. I miljøer der programvare med åpen kildekode utvikles er slik vurdering og testing av andres kode svært vanlig. Der er det et mål i seg selv at kildekoden er korrekt og dermed også troverdig. Dette er en svært viktig motivasjon for slike utviklingsmiljøer.

Kenneally konkluderer i sin artikkel med at troverdigheten til enkelte digitale bevistyper blir nedvurdert. I likhet med Carrier gir også hun uttrykk for at programvare med åpen kildekode i større grad enn proprietær programvare møter kravene i Daubert-prosedyren [4, 14]. Dette er interessant siden dokumentene som undersøkes i denne oppgaven genereres av proprietært lukket programvare.

Kapittel 5

Resultater

I dette kapittelet beskrives resultater fra tester og analyser som har blitt gjennomført for å gi svar på problemstillingen.

P: Kan metadata sikre validiteten til digitale dokumentbevis?

I kapittel 4, Forskningsmetode, ble metode og verktøy for å svare på de enkelte spørsmålene i problemstillingen beskrevet. Resultatene for de ulike spørsmålene legges frem i de videre avsnittene.

5.1 Metadatafunn

I dette avsnittet presenteres resultatene for første spørsmål:

P_1: Hvilke metadata lagres?

5.1.1 Dokumentasjon

Microsoft Word 2003 er programvare som er proprietært lukket. Det vil si at kildekoden ikke offentliggjøres og heller ikke universiteter og læringsinstitusjoner får innsyn. Etter samtaler med en representant fra Microsoft Norge har det blitt klart at heller ikke denne oppgaven er unntatt dette. Derfor har det ikke vært mulig å analysere kildekode for å finne svar på problemstillingen i denne oppgaven.

MS Word 2003 har funksjonalitet for å lagre dokumenter på XML-format. Når noe skal lagres slik defineres struktur og datatyper for hvert element som dokumentet kan inneholde i et XML-skjema. Et XML-dokument bygges opp av tagger[1], og XML-skjemaet inneholder regler for hvilke tagger som fins og hvordan disse ser ut. World Wide Web Consortium (W3C) har definert standarden for slike XML-skjemaer,

[1]En tag kan for eksempel se slik ut: <forfatter>Kari Nordmann</forfatter>. <forfatter> starter taggen, "Kari Nordmann" er verdien taggen er tilordnet og </forfatter> avslutter taggen.

og Microsofts skjemaer for Office-dokumenter er formet etter denne standarden [55]. I "Microsoft Office 2003 Edition XML Schema References" forklares det at det fins mange ulike skjemaer som brukes i MS Office-pakken, men det er spesielt to som brukes for MS Word-dokumenter [18]. "Common Properties Schema" er skjemaet som bestemmer hvilke av et dokuments egenskaper som skal lagres og hvordan informasjonen skal lagres (jamfør figur A.1). Dette skjemaet blir brukt både av MS Word og av MS Excel. "Word Schema" beskriver namespace'ene og de ulike delene en XML-fil som beskriver et Word-dokument består av (jamfør figur A.2). WordprocessingML, som Microsoft kaller skjemaet, lagrer nøyaktig den samme informasjonen som ved lagring på binærformat (.doc). Det gjelder kun for MS Word og ikke de resterende programmene i MS Office-pakken. Dersom man lagrer et Word-dokument på XML-format etter dette skjemaet og legger til taggen
<?mso-application progid="Word.document"?> øverst i XML-filen, vil dokumentet åpne seg i MS Word når man klikker på det. Tekstinnholdet vil være det samme som da dokumentet var på .doc-format.

"Common Properties Schema" brukes som namespace[2] for en delmengde av metadata-elementene i WordprocessingML. I referanse-dokumentet som beskriver XML-skjemaene listes elementene som fins i namespace'et "DocumentProperties" opp. Elementene beskriver typiske egenskaper ved et digitalt dokument og fins gjengitt i tabell 2.1.

I white paperet "Preservation Metadata for Digital Objects: A Review of the State of the Art" skriver OCLC/RLG Working Group on Preservation Metadata at et digitalt objekts metadata kan lagres på tre ulike måter:

1. Innlemmet i det digitale objektet – slik som for html-dokumenter
2. Lagret i en adskilt metadatafil som er inkapslet i objektet
3. Lagret i et separat lagringssted (eng: repository) og bundet til det digitale objektet logisk i stedet for fysisk.

I MS Word 2003 blir metadata lagret i et format kalt OLE-struktur. OLE står for Object Linking and Embedding og tilbyr et strukturert lagringssystem for linking og inkludering av objekter i dokumenter og tilsvarer punkt nummer 2 i listen [58].

[2]Et namespace er en abstrakt omgivelse for en tag. Ved bruk av namespace unngår man flertydighet dersom flere tagger har samme navn. To tagger med like navn må da tilhøre hvert sitt namespace.

Nummer	Element	Nummer	Element
1	Title	14	Created
2	Subject	15	LastSaved
3	Description	16	TotalTime
4	Author	17	Pages
5	Keywords	18	Words
6	Version	19	Characters
7	HyperlinkBase	20	Guid
8	LastAuthor	21	Category
9	Revision	22	PresentationFormat
10	AppName	23	Manager
11	LastPrinted	24	Company
12	Bytes	25	Lines
13	Paragraphs	26	CharactersWithSpaces

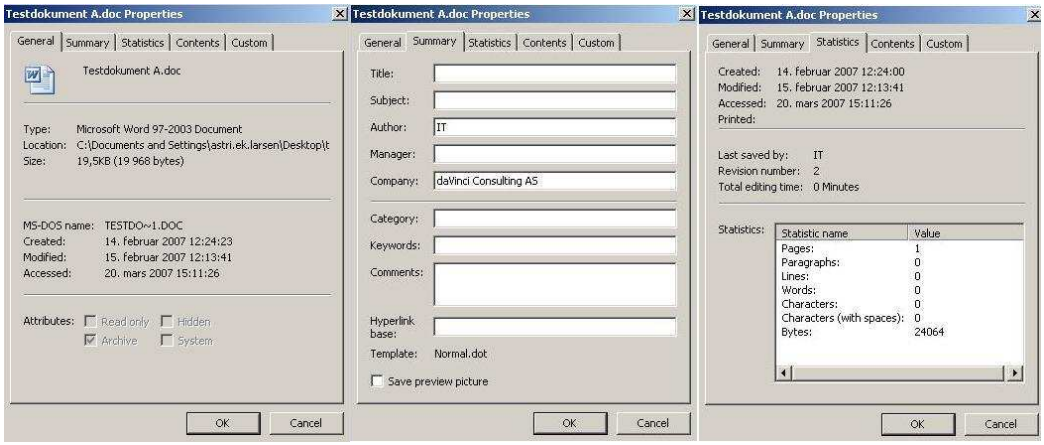
Tabell 5.1: Metadata-elementer i XML-skjema

5.1.2 MS Word - egenskaper

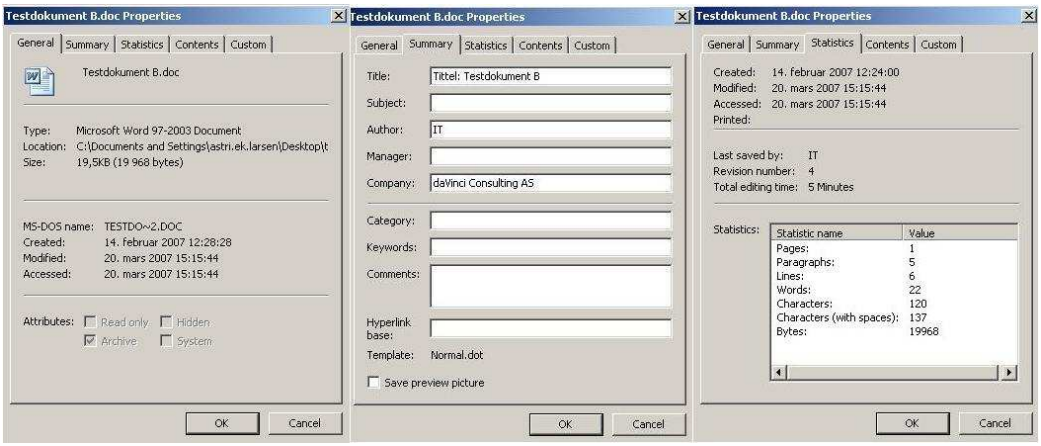
I MS Word 2003 kan man se på egenskapene til et dokument ved å følge menyvalget *File -> Properties*. Der får man opp et nytt vindu med 5 faner: General, Summary, Statistics, Contents og Custom. For de to testdokumentene fantes det kun metadata i de tre første fanene, og det er disse som vises i figurene 5.1 og 5.2.

I første fane, General, vises lokal informasjon om testdokumentene. Type, Location og Size er tre metadata-elementer som sier noe om dokument-versjonen lokalt på pc'en. Videre vises elementene MS-DOS name, Created, Modified og Accessed. Til sist i denne fanen vises 4 elementer i ett, Read-only, Hidden, Archive og System. Disse metadata-elementene vises for begge de to testdokumentene.

Andre fane, Summary, viser typiske metadata-elementer: Title, Subject, Author, Manager, Company, Category, Keywords og Comments. Hyperlink base og Template er også elementer som vises. Det er bare to av disse elementene i testdokument *A* som er tilordnet verdi – Author og Company. For det andre testdokumentet er metadata-elementene Title, Author og Company tilordnet verdi.



Figur 5.1: Testdokument *A*s egenskaper i MS Word 2003



Figur 5.2: Testdokument *B*s egenskaper i MS Word 2003

Til sist i tredje fane, Statistics, vises metadata-elementene Created, Modified, Accessed, Printed, Last saved by, Revision number og Total editing time. De tre første elementene fins også i første fane. Forskjellen er at i første fane inneholder metadata-elementene informasjon om den lokale versjonen av dokumentet mens i tredje fane bærer de informasjon om selve dokumentet. Denne informasjonen følger dokumentet videre dersom dokumentet sendes til en annen pc. Sist i fanen vises statistiske metadata. Elementene for dette er Pages, Paragraphs, Lines, Words, Characters, Characters (with spaces) og Bytes.

Figur 5.3: Testdokument *A* i Metadata Analyzer

Figur 5.4: Testdokument *B* i Metadata Analyzer

5.1.3 Metadata Analyzer

Metadata Analyzer er benyttet for å se hvilke metadata som faktisk er lagret. Som man kan se av skjermbildet i figur 5.3 viser analyseverktøyet for testdokument *A* 7 metadata-elementer: Creation time, Last edited, Author, Last author, Company, Revision number og Total editing time. Alle elementene er tilordnet en verdi.

Figur 5.4 er et skjermbilde av analysen av testdokument *B*. Metadata Analyzer viser 8 metadata-elementer: Title, Creation time, Last edited, Author, Last author, Company, Revision number og Total editing time. Dette dokumentet har altså i motsetning til foregående testdokument et Title-element. Alle de 8 elementene er tilordnet en verdi.

I Metadata Analyzers programvaredokumentasjon angis det at MS Office-dokumenter har metadata-elementene som er listet i tabell 5.2.

Nummer	Element	Nummer	Element
1	Title	16	Creation date
2	Subject	17	Last save time
3	Description	18	Total editing time
4	Author	19	Number of pages
5	Keywords	20	Number of words
6	Comments	21	Number of characters
7	Template	22	Security
8	Last author	23	Category
9	Revision number	24	Format
10	Application name	25	Manager
11	Last print date	26	Company
12	Number of bytes	27	Number of lines
13	Number of paragraphs	28	Number of slides
14	Number of notes	29	Number of hidden slides
15	Number of multi-media clips		

Tabell 5.2: Metadata Analyzers metadata-elementer

```
<?mso-application progid="Word.Document"?>
<w:wordDocument xmlns:w="http://schemas.microsoft.com/office/word/2003/wordml"
  xmlns:v="urn:schemas-microsoft-com:vml"
  xmlns:w10="urn:schemas-microsoft-com:office:word"
  xmlns:sl="http://schemas.microsoft.com/schemaLibrary/2003/core"
  xmlns:aml="http://schemas.microsoft.com/aml/2001/core"
  xmlns:wx="http://schemas.microsoft.com/office/word/2003/auxHint"
  xmlns:o="urn:schemas-microsoft-com:office:office"
  xmlns:dt="uuid:C2F41010-65B3-11d1-A29F-00AA00C14882"
  xmlns:wsp="http://schemas.microsoft.com/office/word/2003/wordml/sp2"
  w:macrosPresent="no" w:embeddedObjPresent="no" w:ocxPresent="no"
  xml:space="preserve">
  <w:ignoreElements
    w:val="http://schemas.microsoft.com/office/word/2003/wordml/sp2"/>
 - <o:DocumentProperties>
    <o:Author>IT</o:Author>
    <o:LastAuthor>IT</o:LastAuthor>
    <o:Revision>2</o:Revision>
    <o:TotalTime>1</o:TotalTime>
    <o:Created>2007-03-20T14:56:00Z</o:Created>
    <o:LastSaved>2007-03-20T14:56:00Z</o:LastSaved>
    <o:Pages>1</o:Pages>
    <o:Words>0</o:Words>
    <o:Characters>0</o:Characters>
    <o:Company>daVinci Consulting AS</o:Company>
    <o:Bytes>24064</o:Bytes>
    <o:Lines>0</o:Lines>
    <o:Paragraphs>0</o:Paragraphs>
    <o:CharactersWithSpaces>0</o:CharactersWithSpaces>
    <o:Version>11.8125</o:Version>
  </o:DocumentProperties>
 - <w:fonts>
    <w:defaultFonts w:ascii="Times New Roman" w:fareast="Times New Roman"
      w:h-ansi="Times New Roman" w:cs="Times New Roman"/>
  </w:fonts>
 - <w:styles>
    <w:versionOfBuiltInStylenames w:val="4"/>
    <w:latentStyles w:defLockedState="off" w:latentStyleCount="156"/>
   + <w:style w:type="paragraph" w:default="on" w:styleId="Normal">
   + <w:style w:type="character" w:default="on" w:styleId="DefaultParagraphFont">
   + <w:style w:type="table" w:default="on" w:styleId="TableNormal">
   + <w:style w:type="list" w:default="on" w:styleId="NoList">
  </w:styles>
 + <w:shapeDefaults>
 + <w:docPr>
 + <w:body>
</w:wordDocument>
```

Figur 5.5: Testdokument *A* i XML

```
<?mso-application progid="Word.Document"?>
- <w:wordDocument xmlns:w="http://schemas.microsoft.com/office/word/2003/wordml"
    xmlns:v="urn:schemas-microsoft-com:vml"
    xmlns:w10="urn:schemas-microsoft-com:office:word"
    xmlns:sl="http://schemas.microsoft.com/schemaLibrary/2003/core"
    xmlns:aml="http://schemas.microsoft.com/aml/2001/core"
    xmlns:wx="http://schemas.microsoft.com/office/word/2003/auxHint"
    xmlns:o="urn:schemas-microsoft-com:office:office"
    xmlns:dt="uuid:C2F41010-65B3-11d1-A29F-00AA00C14882"
    xmlns:wsp="http://schemas.microsoft.com/office/word/2003/wordml/sp2"
    w:macrosPresent="no" w:embeddedObjPresent="no" w:ocxPresent="no"
    xml:space="preserve">
    <w:ignoreElements
      w:val="http://schemas.microsoft.com/office/word/2003/wordml/sp2"/>
  - <o:DocumentProperties>
      <o:Title>Tittel: Testdokument B</o:Title>
      <o:Author>IT</o:Author>
      <o:LastAuthor>IT</o:LastAuthor>
      <o:Revision>2</o:Revision>
      <o:TotalTime>0</o:TotalTime>
      <o:Created>2007-03-20T14:56:00Z</o:Created>
      <o:LastSaved>2007-03-20T14:56:00Z</o:LastSaved>
      <o:Pages>1</o:Pages>
      <o:Words>22</o:Words>
      <o:Characters>120</o:Characters>
      <o:Company>daVinci Consulting AS</o:Company>
      <o:Bytes>19968</o:Bytes>
      <o:Lines>6</o:Lines>
      <o:Paragraphs>5</o:Paragraphs>
      <o:CharactersWithSpaces>137</o:CharactersWithSpaces>
      <o:Version>11.8125</o:Version>
    </o:DocumentProperties>
  - <w:fonts>
      <w:defaultFonts w:ascii="Times New Roman" w:fareast="Times New Roman"
        w:h-ansi="Times New Roman" w:cs="Times New Roman"/>
    </w:fonts>
  - <w:styles>
      <w:versionOfBuiltInStylenames w:val="4"/>
      <w:latentStyles w:defLockedState="off" w:latentStyleCount="156"/>
    + <w:style w:type="paragraph" w:default="on" w:styleId="Normal">
    + <w:style w:type="character" w:default="on" w:styleId="DefaultParagraphFont">
    + <w:style w:type="table" w:default="on" w:styleId="TableNormal">
    + <w:style w:type="list" w:default="on" w:styleId="NoList">
    </w:styles>
  + <w:shapeDefaults>
  + <w:docPr>
  + <w:body>
</w:wordDocument>
```

Figur 5.6: Testdokument *B* i XML

5.2 Generering av metadata

P_2: Hvilke aktører genererer metadata?

5.2.1 Dokumentasjon

I "Microsoft Office 2003 Edition XML Schema References" beskrives et metadata-element som kan tilpasses ved behov. "CustomDocument-Properties" tillater at nye metadata-elementer, som kan tilordnes verdi, defineres og kommer i tillegg til de predefinerte elementene som er listet i tabell 5.1 i kapittel 5.1.1. De nye egendefinerte elementene kan ha en tekststreng, et flyttall, en dato eller en boolsk variabel som verdi.

5.2.2 XML-lagring

Testdokumentene lagret på XML-format inneholder mye informasjon lagret i XML-tagger. Som man kan se av skjermbildene i figur 5.5 og 5.6 er det namespace'et CommonProperties som inneholder egenskaper for dokumentet (<o:DocumentProperties>). Det er bare metadata-elementer som er tilordnet verdi som er med i XML-dokumentet. De to testdokumentene *A* og *B* har lagret de samme metadata-elementene: Author, LastAuthor, Revision, TotalTime, Created, LastSaved, Pages, Words, Characters, Company, Bytes, Lines, Parargraphs, CharactersWithSpaces og Version. Alle elementene inneholder metadata om dokumentene bortsett fra det siste elementet, Version, som forteller hvilken versjon av MS Word 2003 som dokumentene er lagret i. I tillegg til disse elementene har testdokument *B* elementet Title.

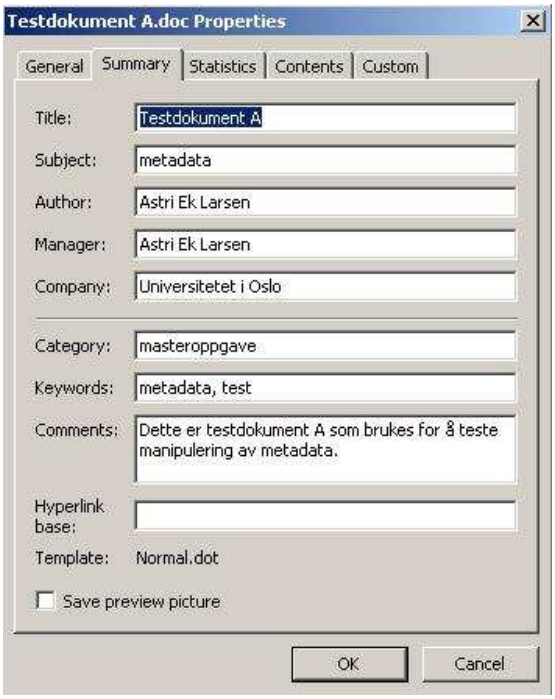
5.3 Manipulering av metadata

P_3: Hvordan kan metadata manipuleres?

P_4: Hvilke aktører kan manipulere metadata?

5.3.1 MS Word - egenskaper

Enkelte metadata-elementer har åpne felter som kan fylles ut i MS Word 2003s egenskaper-visning som følger av menyvalget *File -> Properties*. Dette kan gjøres i andre fane, i Summary-fanen, som vist i figur 5.1 og 5.2. Elementene som ble manipulert på denne måten var Title, Subject, Author, Manager, Company, Category, Keywords og Comments. Figur 5.7 og 5.8 viser at innholdet i de ulike elementene fremdeles eksisterer etter at testdokumentene er lagret, lukket og åpnet igjen.



Figur 5.7: Testdokument *A* manipulert i Egenskaper

Figur 5.8: Testdokument *B* manipulert i Egenskaper

5.3.2 XML-lagring

P_2 ble testet ved å lagre testdokumentene på XML-format. I dette forsøket ble det forsøkt å endre flere av metadata-elementene i namespace'et CommonProperties.

Metadata-elementene Author, Revision, TotalTime og Company i testdokument *A* fikk endret verdi på denne måten. Dette er illustrert i figur 5.9. Legg spesielt merke til at testdokumentet har fått et nytt metadata-element, Title, med tilhørende verdi. Elementet Created som inneholder et tidsstempel har dessuten fått ny verdi. Det var ikke mulig å endre metadata-elementene Pages, Words, Characters, Bytes, Lines, Paragraphs eller CharactersWithSpaces slik at de nye verdiene bevares med testene i denne oppgaven. Dette skyldes at disse elementene får verdi generert av tekstredigeringsapplikasjonen [18].

Metadata-elementene Title, LastAuthor, Created og LastSaved i testdokument *B* har også fått endret verdi i XML-editoren slik det illustreres i figur 5.10. I tillegg har testdokumentet fått et nytt metadata-element, LastPrinted. Legg spesielt merke til at verdien i LastPrinted er et tidspunkt tidligere enn verdien i elementet LastSaved. Verdien i LastSaved er videre et tidspunkt som er tidligere enn verdien i elementet Created. Elementet Author er forøvrig fjernet fra testdokument *B*s metadata ved manipulering i XML-editoren.

XML-fil åpnet i MS Word

Siden både testdokument *A* og *B* inneholder taggen *<?mso-application progid="Word.document"?>* kan de åpnes i MS Word. Når det gjelder metadata i testdokument *A* ser vi av figur 5.11 at alle de fem metadata-elementene som fikk endret verdi i XML-editoren fortsatt har den endrede verdien og at det samme gjelder for det nye elementet Title. For testdokument *B* ser vi i figur 5.12 at tre av de endrede elementene, Title, LastAuthor (Last saved by) og Created fortsatt har de nye verdiene som ble tilordnet. Det fjerde elementet som fikk endret sin verdi, LastSaved (Modified), har derimot ikke beholdt verdien som ble manipulert inn men istedet fått nok en ny verdi. Den nye verdien tilsvarer tidspunktet dokumentet ble åpnet i MS Word for å se på metadataene for denne analysen og vil endres hver gang dokumentet åpnes i denne applikasjonen. Metadata-elementet Author ble fjernet i XML-editoren, og figuren viser at elementet fins, men ikke har noe verdi når man ser på dokumentet i MS Word.

Endret filendelse

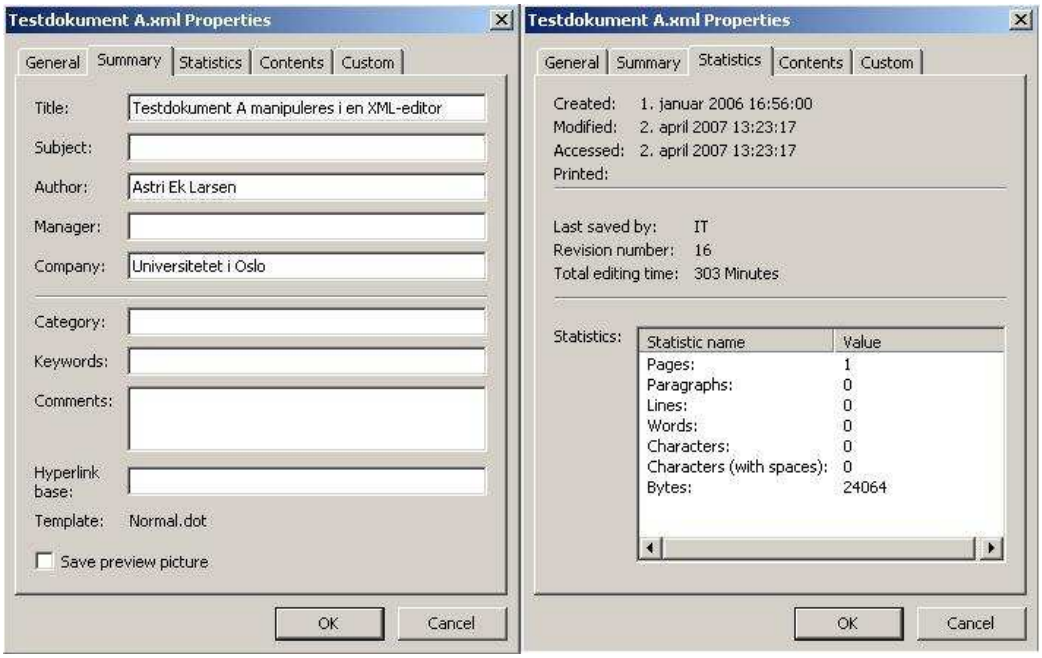
Begge testdokumentene har fått endret filendelse fra .xml til .doc før de åpnes på nytt i MS Word 2003. Dersom man nå ser på testdokumentenes

```
- <o:DocumentProperties>
    <o:Title>Testdokument A manipuleres i en XML-editor</o:Title>
    <o:Author>Astri Ek Larsen</o:Author>
    <o:LastAuthor>IT</o:LastAuthor>
    <o:Revision>16</o:Revision>
    <o:TotalTime>301</o:TotalTime>
    <o:Created>2006-01-01T14:56:00Z</o:Created>
    <o:LastSaved>2007-03-31T19:25:00Z</o:LastSaved>
    <o:Pages>1</o:Pages>
    <o:Words>0</o:Words>
    <o:Characters>0</o:Characters>
    <o:Company>Universitetet i Oslo</o:Company>
    <o:Bytes>24064</o:Bytes>
    <o:Lines>0</o:Lines>
    <o:Paragraphs>0</o:Paragraphs>
    <o:CharactersWithSpaces>0</o:CharactersWithSpaces>
    <o:Version>11.8125</o:Version>
  </o:DocumentProperties>
```

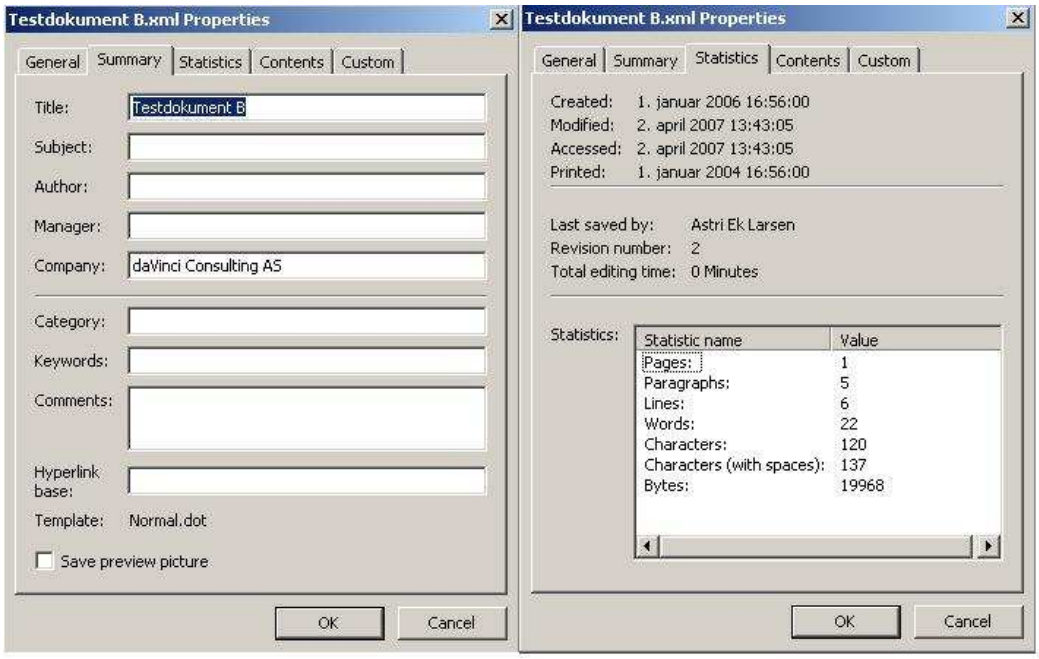
Figur 5.9: Testdokument *A* manipulert i XML-editor

```
- <o:DocumentProperties>
    <o:Title>Testdokument B</o:Title>
    <o:LastAuthor>Astri Ek Larsen</o:LastAuthor>
    <o:Revision>2</o:Revision>
    <o:TotalTime>0</o:TotalTime>
    <o:Created>2006-01-01T14:56:00Z</o:Created>
    <o:LastSaved>2005-01-01T14:56:00Z</o:LastSaved>
    <o:LastPrinted>2004-01-01T14:56:00Z</o:LastPrinted>
    <o:Pages>1</o:Pages>
    <o:Words>22</o:Words>
    <o:Characters>120</o:Characters>
    <o:Company>daVinci Consulting AS</o:Company>
    <o:Bytes>19968</o:Bytes>
    <o:Lines>6</o:Lines>
    <o:Paragraphs>5</o:Paragraphs>
    <o:CharactersWithSpaces>137</o:CharactersWithSpaces>
    <o:Version>11.8125</o:Version>
  </o:DocumentProperties>
```

Figur 5.10: Testdokument *B* manipulert i XML-editor



Figur 5.11: Testdokument *A*s egenskaper etter manipulering i XML-editor



Figur 5.12: Testdokument *B*s egenskaper etter manipulering i XML-editor

egenskaper ved å følge menyvalget *File -> Properties* vil nøyaktig det samme innholdet finnes i metadata-elementene som da testdokumentene hadde filendelsen.xml. Dette illustreres i figurene 5.11 og 5.12.

5.3.3 Operativsystemets dato/klokke

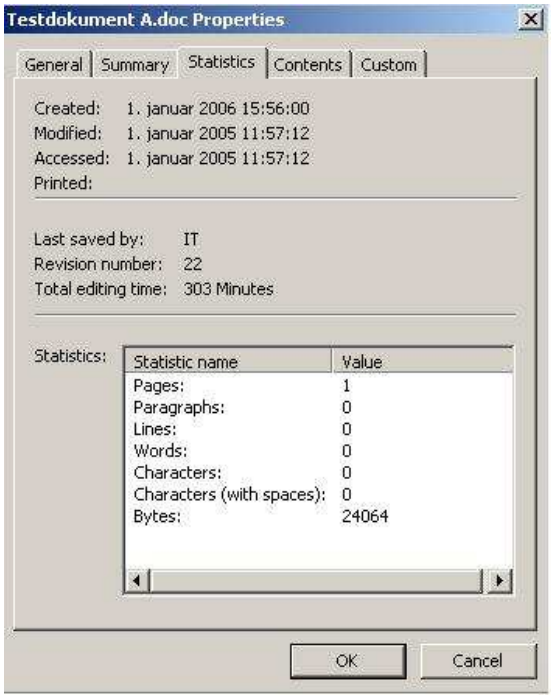
Når MS Word tilordner metadata-elementer som Created, LastSaved og LastPrinted verdi brukes tidsstempler. Disse er basert på operativsystemets dato/klokke-funksjon. Verdien i dato/klokke kan endres av brukeren, og i denne testen har det blitt undersøkt hvilke konsekvenser manipulering av operativsystemets dato/klokke har for tidsstemling av metadata-elementer i MS Word-dokumenter.

Operativsystemets dato ble endret til 1. januar 2005, og de to testdokumentene ble deretter åpnet i MS Word. I figur 5.13 og 5.14 ser vi egenskapene til henholdsvis testdokument *A* og testdokument *B*. For begge dokumentene gjelder det samme – Modified (LastSaved) og Accessed er endret til 1. januar 2005.

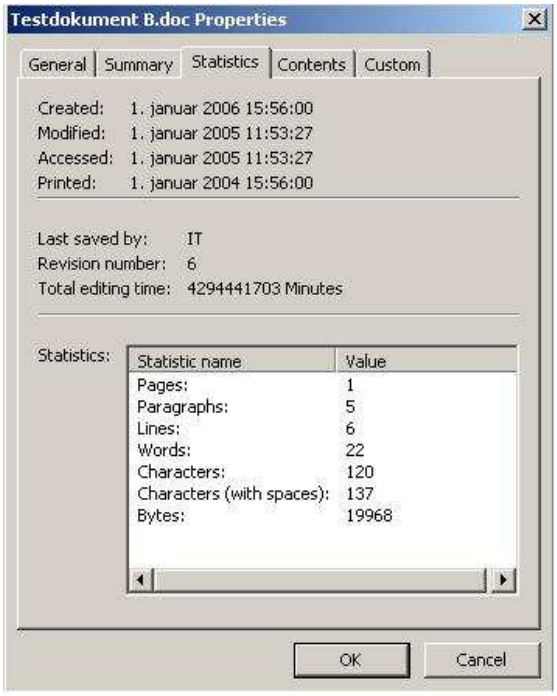
Metadata-elementet Total editing time som beregnes av MS Word har i testdokument *B* fått en interessant verdi. Dette burde være et negativt tall når man ser at tidspunktet dokumentet er opprettet (Created) på er senere i tid enn siste lagringstidspunkt (LastSaved). Istedet er dette tallet satt til 4 294 441 703 minutter som tilsvarer i overkant av 8170 år. En mulig forklaring på dette kan være at MS Word ikke kan håndtere negative verdier i elementet Total editing time. Når verdien her likevel er negativ er det mulig tekstredigeringsapplikasjonen feiltolker denne til et svært høyt positivt tall istedet for et negativt tall.

Hittil har ikke metadataene i første fane, General, som vi så i figur 5.1 og 5.2, blitt endret. Disse metadataene beskriver egenskaper ved den lokale versjonen av testdokumentene og endres ikke ved manipulering via "Egenskaper"-funksjonaliteten i MS Word eller XML-lagring. Metadata-elementene Modified og Accessed som man finner i General-fanen blir imidlertid endret når man endrer operativsystemets dato/klokke slik det har blitt gjort i dette forsøket. Created-elementet som beskriver når de lokale dokumentfilene ble opprettet er ikke påvirket av dette siden dato/klokke-verdien ble endret *etter* at testdokumentene ble opprettet.

For å endre dette metadata-elementet må man opprette en ny lokal versjon av dokumentet mens dato/klokke er endret til verdien man ønsker at Created-elementet i General-fanen skal ha. I denne testen ble testdokumentene midlertidig lagret på et annet medium og den lokale versjonen ble slettet fra pc'en. Deretter ble testdokumentene kopiert tilbake til pc'en. I figur 5.15 og 5.16 illustreres metadata-elementene etter at dokumentene er flyttet tilbake og åpnet i MS Word. Nå er også metadata-elementet Created i General-fanen endret til den manipulerte datoen, 1. januar 2005.

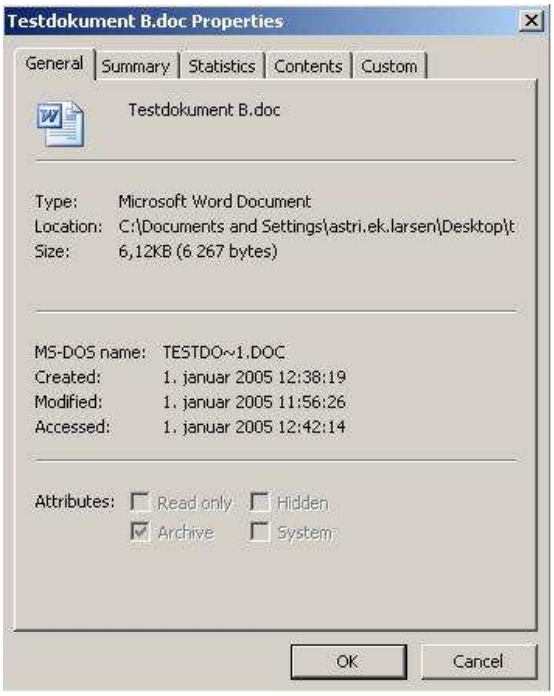


Figur 5.13: Testdokument *A*s egenskaper etter manipulert dato/klokke



Figur 5.14: Testdokument *B*s egenskaper etter manipulert dato/klokke

Figur 5.15: Testdokument *A*s egenskaper etter midlertidig lagring på annet medium



Figur 5.16: Testdokument *B*s egenskaper etter midlertidig lagring på annet medium

Kapittel 6

Analyse og diskusjon

I avsnitt 6.1 analyserer og diskuterer jeg resultatene jeg fant i kapittel 5. I avsnitt 6.2 drøftes tiltak for å sikre validiteten til digitale dokumenter, og i avsnitt 6.3 diskuterer jeg mitt bidrag til det tverrfaglige forskningsmiljøet mellom juss og teknologi.

6.1 Diskusjon over resultatene

P: Kan metadata sikre validiteten til digitale dokumentbevis?

For å finne svaret på problemstillingen blir resultatene fra testene som ble beskrevet i kapittel 4 analysert og diskutert. I de følgende avsnittene besvares underproblemstillingene som ble definert i avsnitt 1.2. Svarene brukes deretter i diskusjonen av hovedproblemstillingen i avsnitt 6.1.4

6.1.1 Metadatafunn

P_1: Hvilke metadata lagres?

Resultatene dokumentasjonen ga oss i avsnitt 5.1.1 viste at det er XML-skjemaene som definerer hvilke metadata-elementer som kan lagres når MS Word-dokumenter opprettes. Elementene som det er mulig å opprette listes opp i tabell 6.1 (også tabell 5.1). Dette er 26 typiske egenskaper ved innholdet i et dokument som tittel, forfatter, nøkkelord og antall sider. Noen av elementene inneholder informasjon om når en hendelse har inntruffet dokumentet, og verdien i disse er tidsstempler. Slike metadata beskriver for eksempel når dokumentet ble opprettet, sist endret og sist skrevet ut. Det er viktig å merke seg punktet i dokumentasjonen som påpeker at denne mengden metadata er felles for alle programmene i MS Office-pakken og derfor også gjelder for eksempel for regneark-programmet MS Excel 2003.

Metadata Analyzer ga resultater om hvilke metadata som lagres på to ulike måter: Ved dokumentasjon og gjennom testing. Programvarens

Nummer	Element	Nummer	Element
1	Title	14	Created
2	Subject	15	LastSaved
3	Description	16	TotalTime
4	Author	17	Pages
5	Keywords	18	Words
6	Version	19	Characters
7	HyperlinkBase	20	Guid
8	LastAuthor	21	Category
9	Revision	22	PresentationFormat
10	AppName	23	Manager
11	LastPrinted	24	Company
12	Bytes	25	Lines
13	Paragraphs	26	CharactersWithSpaces

Tabell 6.1: Metadata-elementer i XML-skjema

dokumentasjonen ramser opp 29 ulike elementer som kan lagre informasjon om et dokument i tabell 6.2 (også tabell 5.2). Også disse elementene gjelder for alle programmene i MS Office-pakken.

Det er mange like elementer i de to tabellene 6.1 og 6.2. Typisk informasjon som tittel, emne, forfatter, beskrivelse og nøkkelord fins i begge listene. Det samme gjelder elementer som inneholder tidsstempler som dato for opprettelse, sist lagret og sist skrevet ut. Også elementene for statistisk informasjon som antall bytes, sider, ord og bokstaver fins i begge listene.

XML-skjemaet som beskriver hvilke elementer som kan lagres i MS Word-dokumenter inneholder imidlertid 4 elementer som ikke fins i Metadata Analyzers liste. Dette er elementene Version, HyperlinkBase, Guid og CharactersWithSpaces. Metadata Analyzers liste over metadata har på sin side sju elementer som ikke fins i listen fra XML-skjemaet. Dette er elementene Comments, Template, Security, Number of notes, Number of multimedia clips, Number of slides og Number of hidden slides. Selv om begge tabellene skal gjengi hvilke metadata-elementer som kan lagres, er det altså ikke fullstendig samsvar dem imellom.

Skjermbildene av de to testdokumentenes metadata i de to ulike verktøyene MS Word 2003 og Metadata Analyzer viser heller ikke de

Nummer	Element	Nummer	Element
1	Title	16	Creation date
2	Subject	17	Last save time
3	Description	18	Total editing time
4	Author	19	Number of pages
5	Keywords	20	Number of words
6	Comments	21	Number of characters
7	Template	22	Security
8	Last author	23	Category
9	Revision number	24	Format
10	Application name	25	Manager
11	Last print date	26	Company
12	Number of bytes	27	Number of lines
13	Number of paragraphs	28	Number of slides
14	Number of notes	29	Number of hidden slides
15	Number of multi-media clips		

Tabell 6.2: Metadata Analyzers metadata-elementer

samme metadata-elementene. I Metadata Analyzer er de 8 elementene Title (kun i testdokument *B*), Creation time, Last edited, Author, Last author, Company, Revision number og Total editing time tilordnet verdi. MS Word 2003s egenskaper-funksjonalitet viser alle disse elementene, men viser også at det er enda flere elementer som er lagret. Disse elementene er statistiske metadata som Pages, Paragraphs, Lines, Words, Characters, Characters (with spaces) og Bytes, i tillegg til tidsstempelet for når dokumentet sist er åpnet, Accessed. Metadata-elementene i tabell 6.3 svarer på spørsmålet om hvilke metadata som lagres default når man oppretter dokumenter henholdsvis med eller uten innhold.

Metadata som lagres i MS Word 2003-dokumenter er ikke identiske med de 22 kjerne-elementene som Dublin Core Metadata Initiative (DCMI) definerer i tabell 6.4 (også tabell 2.1). Elementene som fins i de to listene 6.1 og 6.2 passer med definisjonen, men det er kun 4 elementer

Nummer	Element	Nummer	Element
1	Title (kun i testdokument *B*)	9	Characters (with spaces)
2	Creation time	10	Characters
3	Last edited	11	Pages
4	Author	12	Words
5	Last author	13	Lines
6	Company	14	Paragraphs
7	Revision number	15	Bytes
8	Total editing time	16	Accessed

Tabell 6.3: Metadata-elementer generert default

Nummer	Element	Nummer	Element
1	Title	12	Date
2	Subject	13	Format
3	Description	14	Identifier
4	Type	15	Language
5	Source	16	Audience
6	Relation	17	Provenance
7	Coverage	18	RightsHolder
8	Creator	19	InstructionalMethod
9	Publisher	20	AccrualMethod
10	Contributor	21	AccrualPeriodicity
11	Rights	22	AccrualPolicy

Tabell 6.4: Kjerne-elementer i DCMIs definisjon av metadata

i DCMIs liste som også fins i disse listene. Enhver programvareprodusent står fritt til å definere sine egne metadata-elementer, også Microsoft. Kjerne-elementene til DCMI er èn mulig mengde elementer man kan lagres og er dessuten mer rettet mot strukturering, bevaring og langtidslagring av digitalt materiale enn mengden metadata som lagres i MS Word 2003-dokumenter.

6.1.2 Generering av metadata

P_2: Hvilke aktører genererer metadata?

Følgende aktører kan generere metadata i MS Word-dokumenter:

- Programvare:

  Tekstredigeringsapplikasjon

  XML-editor

  Operativsystem

- Personrolle: Forfatter

Det er først og fremst programvaren de digitale dokumentene er opprettet i som genererer metadata. I avsnitt 5.2 så vi på metadataene ved å lagre testdokumentene på XML-format og deretter åpne dem i en XML-editor. Det kom frem en rekke metadata-elementer som var tilordnet verdi. Ved sammenligning var elementene de samme som i resultatene i P_1. Innholdet i metadata-elementene ble ikke fylt inn av forfatter ved opprettelse av testdokumentene, og derfor må verdiene ha blitt generert av programvare. Testdokumentene ble opprettet i MS Word 2003, og denne programvare-aktøren har generert metadata-elementene og deres innhold.

Operativsystemet spiller også en rolle når metadata genereres. Enkelte av elementene har tidsstempler som innhold. Kilden til tidsstemplene er operativsystemets dato/klokke-funksjon. MS Word 2003 bruker innstillingen i denne funksjonaliteten når elementer med tidsstempler tilordnes verdi. Tekstredigeringsapplikasjonen benytter seg dermed av en annen aktør for å generere metadata, nemlig operativsystemet.

XML-editoren og forfatteren er to andre aktører som sammen kan generere metadata. Forfatteren kan opprette MS Word-dokumenter direkte i XML-format, og dette gjøres i en XML-editor. I avsnitt 5.1.1 så vi at XML-skjemaer definerer hvilke elementer som dokumentet kan inneholde. Forfatteren kan derfor velge hvilke metadata som skal lagres for å beskrive dokumentets egenskaper ut fra Microsofts XML-skjemaer. Dokumentasjonen av XML-skjemaene som MS Word 2003 bruker, ga i avsnitt 5.2 svarene på hvilke elementer dette er. I tillegg til disse elementene beskrev dokumentasjonen et annet element,

"CustomDocumentProperties". Den gjør det mulig for forfatteren å generere valgfrie metadata som lagres innenfor dette ene elementet. Disse metadataene kan ha innhold med en tekststreng, et flyttall, en dato eller en boolsk variabel som verdi. Det er forfatteren, som personrolle-aktør, og XML-editoren, som programvare-aktør, som genererer metadata på denne måten.

6.1.3 Manipulering av metadata

P_3: Hvordan kan metadata manipuleres?

P_4: Hvilke aktører kan manipulere metadata?

I følgende programvare kan disse aktørene manipulere metadata:

- Tekstredigeringsapplikasjon

  Av personrolle: Forfatter

  Av programvare: Tekstredigeringsapplikasjon

- XML-editor

  Av personrolle: Forfatter som manipulerer verdi

  Av personrolle: Forfatter som legger til elementer

- Filsystem

  Av personrolle: Forfatter som endrer filendelse

- Operativsystem

  Av personrolle: Forfatter som endrer operativsystemets dato/klokke

Både programvare og personrolle kan manipulere metadata. Manipulasjonen foregår i ulike typer programvare. Den mest opplagte måten å manipulere metadata på er å bruke programvaren et dokument er opprettet i. I avsnitt 5.3.1 så vi at man i filvalget *File -> Properties* åpner et nytt vindu med dokumentets egenskaper (jamfør figur 5.1 og 5.2). Her har flere av metadata-elementene felter der forfatteren kan skrive inn nye verdier i elementene. Dermed er forfatteren en aktør som kan manipulere metadata ved bruk av programvaren MS Word 2003.

Denne programvaren opptrer selv også som aktør som kan manipulere metadata. I undersøkelsene ble flere statistiske verdier i metadata-elementer som Pages, Words, Characters, Bytes, Lines, Paragraphs og CharactersWithSpaces forsøkt endret ved hjelp av en XML-editor. Det viste seg at disse elementene ikke fikk endret verdi på denne måten. Dokumentasjonen viste at dette ikke er mulig fordi MS Word 2003 selv beregner disse verdiene når dokumenter åpnes i applikasjonen.

I tillegg viste resultatene at MS Word 2003 også beregner andre verdier når dokumenter åpnes. Som ved førstegangs generering av verdier bruker MS Word 2003 operativsystemets dato/klokke-funksjon for å tilordne elementer verdi, og noen av elementene som har tidsstempler som verdi, fikk endret verdi på denne måten. Slik opptrer tekstredigeringsapplikasjonen som aktør som kan manipulere metadata.

XML-editoren er også en programvare-aktør som kan manipulere metadata. Resultatene av testene utført ved hjelp av XML-editoren gir et klart bilde av hvilke metadata som kan manipuleres på denne måten. Alle elementene som beskriver dokumentegenskapene kan endres av forfatter-aktøren i editoren. Nye elementer, hentet fra mengden elementer som blir definert i XML-skjemaer, kan legges til og tilordnes verdi. I testresultatene så vi at ikke alle elementene bevarte verdiene sine når testdokumentene senere ble åpnet i MS Word 2003, men alle endrede verdier ble bevart ved å åpne dokumentene på nytt i editoren. Det betyr at forfatteren og XML-editoren har manipulert metadataene.

Pc'ens filsystem er den tredje programvare-aktøren. Resultatene viser at det fins en enkel måte å bevare manipulerte metadata på – også når dokumenter åpnes i MS Word 2003. Forfatteren kan endre dokumenters filendelse fra .xml til .doc. Da bevares alle metadata som har innhold som *ikke* genereres av MS Word 2003 ved åpning av dokumenter (statistiske metadata og elementet Last Accessed). Manipuleringen skjer i filsystemet, og dermed blir også dette en aktør som manipulerer metadata.

Operativsystemet er delvis en aktør som manipulerer metadata fordi MS Word 2003 bruker dato/klokke-funksjonen som kilde når programvaren tidsstempler metadata-elementer. Resultatene i avsnitt 5.3.3 viser at forfatteren kan manipulere hvilke verdier MS Word 2003 bruker. Nettopp fordi det er tekstredigeringsapplikasjonen som direkte manipulerer metadata-elementene i dokumentet, regnes operativsystemet som delvis aktør i dette tilfellet. Denne aktøren bidrar bare med verdi og gjennomfører ikke selve manipulasjonen.

6.1.4 Sikring av validiteten til digitale dokumentbevis

P: Kan metadata sikre validiteten til digitale dokumentbevis?

Hvorfor er det viktig at digitale dokumentbevis har validitet?

Påtalemyndigheten sitter med bevisbyrden i straffesaker og legger ofte frem mange bevis for å belyse en sak best mulig. Deretter er det opp til dommeren å godta eller avskjære de fremlagte bevisene. Dommeren må ha kunnskap om bevisene, og bevisene må ha validitet for at de skal bli godtatt. De må dessuten være autentiske og ekte. Forskeren Stephen Mason skriver at autentisiteten til et bevis er bygget opp av bevisets troverdighet, integritet og brukbarhet (jamfør figur 3.1).

Et digitalt bevis må analyseres før validiteten kan bestemmes. Digital etterforskning foregår ved at etterforskerne først sikrer bevis ved å lage en speilkopi av det digitale mediet som etterforskes. I analysefasen sorteres og gjennomgås de ulike bevisene som er sikret. Etterforskerne kan analysere det sikrede materialet ved hjelp av programvare utviklet nettopp for etterforskning av digitalt materiale. Dersom det digitale materialet skal legges frem som bevis, må etterforskerne i evalueringsfasen avgjøre hva bevisene kan fortelle om bruken av det digitale mediet de var lagret på og hvilke handlinger som har blitt foretatt.

Handlinger som har blitt gjort på det digitale mediet virker inn på bevisets verdikjede. Verdikjeden til digitale bevis beskriver historien eller livsløpet til et bevis. Den sier noe om hva som hender et bevis, hvilke aktører som får hendelsene til å inntreffe og når de inntreffer (jamfør tabell 2.2).

Når hendelser inntreffer et digitalt dokument kan informasjon om dette lagres i dokumentets metadata. I avsnitt 6.1.1, som diskuterer problemstilling P_1, så vi at Microsoft kun benytter seg av 4 av DCMIs 22 predefinerte kjerne-elementer når metadata om MS Word 2003-dokumenter skal lagres (jamfør tabell 6.4). I tabell 6.1 så vi at programvarens XML-skjema for denne dokumenttypen tillater å lagre 26 ulike elementer. Testene i denne oppgaven viser at de 16 (15 i testdokument *A*) elementene i tabell 6.3 er elementene som faktisk blir generert default når et MS Word 2003-dokument (hhv. med og uten tekstinnhold) opprettes.

Hvilke aktører som genererer metadata er vesentlig i verdikjeden og kunnskap om disse er viktig for etterforskere som undersøker digitale bevis. I avsnitt 6.1.2 diskuteres oppgavens andre delproblemstilling, P_2. Det viser seg at det først og fremst er programvare som genererer metadata, men at også personrolle-aktører, som forfatter av dokumentet, kan generere metadata.

Kan metadata sikre validiteten til digitale dokumentbevis?

For at metadata skal sikre validiteten til digitale dokumentbevis må man være sikker på at informasjonen som er lagret er korrekt, ekte og ikke er manipulert med.

I avsnitt 6.1.3 ble delproblemstilling P_3 diskutert. Spørsmålet om hvordan metadata kan manipuleres har flere svar. Metadata-elementenes verdi endres i programvare. Testene i denne oppgaven viser at slike endringer kan skje både i en tekstredigeringsapplikasjon, en XML-editor, et filsystem og et operativsystem.

Flere ulike aktører kan manipulere metadata. Det viser testene som er gjennomført for å besvare delproblemstilling P_4. Tekstredigeringsapplikasjonen MS Word 2003 kan endre flere av metadataverdiene i testdokumentene og det kan også forfatteren av dokumentene. Det er

fullt mulig å endre elementer som inneholder informasjon om forfatter eller tittel i et digitalt dokument. Metadata-elementer som inneholder informasjon om opprettelsesdato, sist endret eller sist skrevet ut kan også endres av forfatteren. Dermed ser vi at alle metadata-elementer i MS Word-dokumenter kan manipuleres av ulike aktører og spesielt av personrolle-aktører.

Påstanden om at metadata kan sikre validiteten til digitale dokumentbevis avhenger av at metadataene er korrekte og ekte. Testene foretatt i denne oppgaven viser at informasjonen i metadata-elementene kan manipuleres. Manipuleringen kan ikke oppdages med det blotte øye. Man kan derfor ikke hevde at metadata kan sikre validiteten til digitale dokumentbevis.

Kan *ikke* metadata sikre validiteten til digitale dokumentbevis?

Spørsmålet om metadata kan sikre validiteten til digitale dokumentbevis ble altså besvart med et *nei*. La oss derfor undersøke hvor sikkert det er at metadata *ikke* kan sikre validiteten til digitale dokumentbevis.

Tidsstempler Tidsstempler brukes for å lagre informasjon om tidspunktet en hendelse fant sted. Denne informasjonen anses ofte som sikker og korrekt. MS Word 2003 tidsstempler flere av metadata-elementene som ble funnet i P_1. Elementene som beskriver tidspunkt for opprettelse av dokumentet (Created), tidspunkt for siste lagring (LastSaved) og tidspunkt for sist dokumentet ble skrevet ut (LastPrinted) har alle tidsstempler som verdi.

En av årsakene til at tidsstempler har stor troverdighet er at verdien i disse elementene genereres av tekstredigeringsapplikasjonen. Den genererte verdien samsvarer med tidspunktet hendelsene skjer på, og tilordningen skjer "bak kullissene" inne i programvaren. De fleste brukere av programvaren, dette gjelder også dommere og advokater, har ikke detaljkunnskap om hvordan den virker. I mange tilfeller har man større tiltro til programvaren enn den strengt tatt fortjener. Særlig programvare med proprietær lukket kildekode nyter godt av denne mulige misforståelsen.

Det er imidlertid stor usikkerhet omkring tidsstempler. Willassen og Mjølsnes påpeker i "Digital Forensics Research" at det er lite systematisk dokumentasjon av eksisterende formater av tidsstempler [34]. Det er store variasjoner i den faktiske bruken av tidsstempler i forhold til ulike kalender-systemer og tidssoner. Det fins ingen standardiserte retningslinjer for hvordan tidsstempling skal forholde seg til sommer- og vintertid. Operativsystemer tidsstempler hendelser, og det er også fra operativsystemet at tekstredigeringsapplikasjoner, som MS Word 2003, får verdien som de tidsstempler dokumentenes metadata-elementer

med. Desverre er også systematisk dokumentasjon av håndtering av tidsstempler innenfor ulike operativsystemer mangelfull.

Metadata-elementer, som får verdi fra tidsstempler, kan manipuleres. Testene som er gjennomført i denne oppgaven for å besvare P_3 og P_4 viser at alle slike elementer som er lagret i MS Word 2003-dokumenter kan endres. Testene ble gjort ved å manipulere kilden som MS Word bruker for å tidsstemple – dato/klokke-funksjonen i operativsystemet. Dette viser at i tillegg til mangelfull dokumentasjon og kunnskap omkring bruken av tidsstempler, kan også metadata-elementer med tidsstempler manipuleres. Dermed kan tidsstempler gi falsk informasjon om når en hendelse har inntruffet.

Sakkyndige Det er opp til dommeren å avgjøre om bevis skal godtas eller avskjæres. Dersom rettssakens tema eller fremlagte bevis er utenfor dommerens generelle kunnskapsområde, kan sakkyndige kalles inn og avgi vitnemål. Sakkkyndige er personer som forventes å ha god kunnskap på et område, men det er ikke noen formelle krav til den som vitner som sakkyndig.

En sakkyndig som skal uttale seg om et digitalt dokumentbevis, må ha særs god kjennskap til programvaren som dokumentet har blitt opprettet i. Sakkyndige vil ofte være en som til daglig arbeider med den aktuelle programvaren. Dersom dokumentbeviset er opprettet i programvare med proprietær lukket kildekode, vil sakkyndige måtte ha tilgang til kildekode. Dersom kildekoden er proprietært lukket, må sakkyndige enten få innsynsrett eller være ansatt i firmaet som utvikler programvaren. Den sakkyndige vil da måtte gi en objektiv forklaring om et emne og en aktør hun har et avhengighets- og/eller maktforhold til. Det kan stilles spørsmål ved hvor objektiv en slik ekspertuttalelse kan være. Det er heller ikke gitt at en teknisk sakkyndig vil kunne oppdage metadata som er manipulert.

Egenskaper ved digitale dokumentbevis Det er mange fordeler med digitale dokumentbevis. Organisasjonen Cybex fant gjennom prosjektet "The admissibility of electronic evidence in court: Fighting against high-tech crime" en rekke fordeler ved denne bevistypen (jamfør tabell 3.2). Tekniske fagpersoner uttaler i undersøkelsen at slike bevis er vesentlige for bevisførsel ved nye former for kriminalitet. Bevismessig er digital informasjon svært solid og har stor troverdighet. Informasjonen er ofte nøyaktig, komplett og sann.

Digitale dokumentbevis har imidlertid kun disse fordelene dersom man kan garantere at bevisene er autentiske, troverdige og har sikker validitet. Cybex' prosjekt nevner mange ulemper ved digitale data som gjør dem vanskelige å bruke som bevismateriale (jamfør tabell 3.3). Det er i dagens europeiske jurisdiksjoner stor mangel på systematisk regulering

av digitalt materiale. Digital informasjon krever spesiell kunnskap, og det er få eksperter på området. Kunnskapsnivået hos advokater og dommere er heller ikke på et tilfredsstillende nivå. Selv om Cybex' prosjekt har fokusert på situasjonen i EU-land er det ingen grunn til å anta at situasjonen er vesentlig annerledes i Norge.

Det er også flere tekniske ulemper ved bruk av digitale dokumentbevis. Det kan være vanskelig å fastslå hvem som faktisk har hatt befatning med det digitale materialet. Det er svært krevende å sikre, analysere og evaluere digitalt materiale fordi omfanget og mengden ofte er enorm i forhold til resursene man har tilgjengelig. Den største ulempen er at det er vanskelig å bevise autentisitet og troverdighet, samt at et digitalt dokuments validitet ikke er sikkker fordi det er enkelt å manipulere dem. Dette gjør det vanskelig å etablere beviskraft for digitale bevis.

Metadata sikrer *ikke* validiteten til digitale dokumentbevis. Metadata er enkele å manipulere. Elementer som ofte har stor tillit, som tidsstempler, kan også manipuleres og er derfor ikke med på å sikre validiteten til digitale dokumentbevis. Sakkyndige kan ikke oppdage manipulasjon, og selv om digitale bevis har mange fordeler, er problemene rundt autentisitet, ekthet og validitet vesentlige ulemper ved denne bevistypen.

6.2 Mulige tiltak

6.2.1 Undersøkelse av lagrede metadata

Når man skriver ut digitale dokumenter er det default at bare tekstinnholdet i dokumentet vises på utskriften. Dermed forblir metadataene skjulte – selv om de er en del av dokumentet. I sammenhenger der disse dataene er vesentlige, for eksempel innenfor rettsvesenet, vil utskriften dermed bare vise èn del av dokumentet.

Første tiltak for å oppdage at digitale dokumenter er manipulerte er å se på og undersøke metadataene. Ved å gjøre det kan man oppdage ulogiske sammenhenger. Ved å undersøke metadata-elementene som inneholder tidsstempler vil man raskt se om disse inneholder logisk riktig informasjon.

Det fins en rekke måter å se på metadata på. MS Word 2003 åpner et nytt vindu der metadataene med tilhørende verdier presenteres når man følger menyvalget *File -> Properties*. Programvaren har også funksjonalitet som gjør det mulig å skrive ut metadata sammen med tekstinnholdet i dokumentet. XML-editoren Exchanger XML Lite og Metadata Analyzer har også funksjonalitet for å vise metadata som er lagret i et MS Word 2003-dokument. Testene i denne oppgaven har blitt gjennomført ved hjelp av disse to applikasjonene.

Testene som er gjennomført for å besvare P_1, viser at elementer som Created, LastSaved og LastPrinted inneholder tidsstempler. Hendelsene

som disse tre elementene korresponderer med må komme i en gitt rekkefølge for å gi logisk riktig mening. Et dokument må være opprettet før det blir skrevet ut, og det må være skrevet ut før det er lagret siste gang. Testresultatene viser at det er mulig å manipulere verdiene slik at de ikke har logisk riktige verdier (jamfør figur 5.14).

Man kan også undersøke den totale tiden som er brukt på å redigere et dokument. ”Total editing time” er et metadata-element som tilordnes verdi av MS Word. Verdien i dette elementet beregnes ut fra tidsstemplene i elementene nevnt i forrige avsnitt. Dersom dette elementet har en unormal verdi er det grunn til å gå resten av metadata-elementene nærmere etter i sømmene. Ett av resultatene av testene i denne oppgaven viser et eksempel på en slik ulogisk verdi: Total redigeringstid er satt til 8170 år (jamfør figur 5.14).

I vinduet som MS Word åpner når man følger menyvalget *File -> Properties* får man opp 5 faner med dokumentegenskaper. Både i første fane, General, og tredje fane, Statistics, får man opp metadata-elementene Created, Modified og Accessed. Dersom de tre elementene ikke har de samme verdiene i de to fanene, kan dette også være et tegn på at verdiene er manipulerte. Dette er elementer som henholdsvis beskriver metadata i den lokale versjonen av dokumentet og metadata som lagres i dokumentet. Testene som er gjennomført i denne oppgaven viser at de to gruppene metadata-elementer ikke nødvendigvis endres ved èn og samme manipulering.

Ved digital etterforskning brukes ofte en programvare som gjør det mulig å speilkopiere mediet der digitalt materiale er lagret uten å endre dette materialet. Programvaren har ofte også funksjonalitet som letter analyse- og evalueringsarbeidet. EnCase Forensic er èn slik programvare, og denne er i bruk både av offentlige og private aktører i Norge. En av funksjonene denne programvaren har er å sammenligne filendelser og digitale signaturer som ligger lagret i ulike filtyper. Dersom en filendelse og den korresponderende digitale signaturen ikke samsvarer, er det en filsignatur-mismatch (jamfør avsnitt 4.2.1).

Når man skal analysere og undersøke digitale dokumenter bør man vurdere å bruke et slikt analyseverktøy. Det kan avdekke manipulerte dokumenter. Digitalt materiale er ikke fjernet fra en hard-disk før nye data er lagret på samme sted på disken. Derfor kan slettet digitalt materiale ofte rekonstrueres. Når et digitalt dokumentbevis skal presenteres i retten bør beviset presenteres digitalt og på en slik måte at metadata-delen av dokumentet vises. Utskrifter av beviset kan eventuelt supplere den digitale presentasjonen og brukes som referanseinformasjon slik det anbefales i boken ”Brott och digitala bevis” [15].

Tidsstempling

Det er ikke alltid nok i seg selv å undersøke om metadata-elementene som inneholder tidsstempler har logisk korrekte verdier. Testene i denne oppgaven viser tydelig at det ikke er vanskelig å manipulere slike elementer. Elementene trenger dessuten ikke å være bevisst manipulerte for å inneholde gale verdier.

I forskningsprosjektet "Time stamps in digital forensics" ved NTNU arbeides det for tiden med slike tidsstempler. Det er uttrykt en klar mangel på dokumentasjon og standarder for tidsstempler. Derfor dokumenterer og analyserer prosjektet eksisterende tidsreferansesystemer, tidssoner og formater samt beskriver hvordan systemoperasjoner påvirker tidsstempler. Gjennom prosjektet skal det utvikles metoder, programvare og verktøy for å oppdage feilaktige tidsstempler. Programvaren som utvikles vil også bli utsatt for grundig testing.

Dette arbeidet vil kunne bedre troverdigheten til tidsstempler. Man er avhengig av ekte og troverdig informasjon dersom man skal kunne bruke tidsstempler for å avgjøre autentisitet og validitet.

6.2.2 Retningslinjer, lovverk og kunnskapsnivå

For å bedre dagens situasjon er det viktig at man har et tydelig lovverk når det gjelder digitale medier og digitalt materiale. Forskning viser at Europas ulike jurisdiksjoner mangler både lovverk og retningslinjer for bruk av digitale bevis. Det medfører at det i stor grad er opp til advokater og dommere å utøve skjønn når slike bevis fremlegges.

For å bedre situasjonen må lovverket tilrettelegges slik at også digitalt materiale omfattes av dette. I Norge har Datakrimutvalget allerede startet dette arbeidet. I februar presenterte de sitt forslag til ny straffelov. Forslaget går ut på å innføre et nytt kapittel om datakriminalitet i loven. Det skal ta for seg vern av data, databasert informasjon og datasystemer.

I tillegg til å definere straffebud som rammer en rekke kriminelle handlinger der digitalt materiale er i fokus, består lovforslaget også av et omstridt punkt. Utvalget er delt i synet på utkastet til § 76b som omhandler filtrering av utenlandske nettsteder. I dagens lovverk har man bare mulighet til å stenge nettsteder med opprinnelse i Norge. To av utvalgets medlemmer foreslår at det også skal åpnes for at nettleverandører kan pålegges å filtrere tilgangen til enkelte utenlandske nettsteder. Det nye punktet i straffeloven vil gjøre det mulig å filtrere informasjon som kommer fra utenlandske servere, og ikke bare norske, slik dagens lovverk allerede åpner for.

Forslaget til § 76b har møtt stor motstand. Flertallet mener at hensynet til ytringsfriheten taler mot å sensurere utenlandske nettsider for norske brukere og er sterkt imot denne delen av lovforslaget.

I tillegg til mangel på regel- og lovverk viser forskning at kunnskapsnivået blant advokater og dommere om generering, bevaring og presentering av digitalt materiale beklageligvis ikke er høyt nok. Dette er svært betenkelig siden det nettopp er disse personene som skal ta prosessavgjørelser om slike bevis.

Det må være et krav om at både påtalemyndighet, advokater og dommere har en viss grad av kunnskap om digitalt materiale når det blir lagt frem som bevis. Nye lov- og regelverk vil ikke ha videre virkning dersom man ikke har evne til å vurdere bevisenes ekthet og validitet. Kriminelle vil ha store muligheter til å omgå lovverket dersom man ikke kan avdekke manipulering og dokumentfalsk for digitalt materiale. Derfor må man høyne kunnskapsnivået til advokater og dommere.

6.3 Kommentarer

I dette avsnittet gir jeg kommentarer til avgrensningene som er gjort i denne oppgaven. Metodene brukt for å gjennomføre forsøkene i oppgaven kommenteres også, og jeg kommenterer resultatene som kom frem gjennom forsøkene.

6.3.1 Avgrensninger

Denne oppgaven har hatt fokus på dokumenter opprettet i tekstredigeringsapplikasjonen Microsoft Word 2003. Det betyr at andre dokumentformater, som for eksempel er produsert av OpenOffice Writer eller er lagret på pdf-format, ikke har blitt undersøkt og analysert. Denne oppgaven har heller ikke fokus på e-post eller informasjon produsert av operativsystemer og servere. Avgrensningene har også utelukket system- og datalogger samt databaseinnhold.

Disse avgrensningene medfører at størsteparten av alt digitalt materiale ikke har blitt behandlet i denne oppgaven. Det har imidlertid ikke vært forfatterens hensikt å dekke alle digitale datatyper. Målet har vært å sette fokus på egenskaper ved digitalt materiale og hvordan digitale bevis håndteres i rettsvesenet. Oppgaven er en oppmuntring til et tverrfaglig samarbeid mellom det juridiske og det teknologiske fagmiljøet.

6.3.2 Metode og verktøy

Metoden for å svare på oppgavens problemstilling har vært å gjøre tester med to testdokumenter som ble opprettet i MS Word 2003. De to dokumentenes metadata ble analysert og forsøkt manipulert. Målet med testene var å undersøke hvordan en vanlig bruker uten videre innsikt i MS Word kan endre digitale dokumenters metadata. Testene

viser ikke potensielle metoder for hvordan avanserte hackere eventuelt kan manipulere bevis.

Testene ble utført ved hjelp av tekstredigeringsapplikasjonen MS Word 2003, XML-editoren Exchanger XML Lite og metadata-analysereren Metadata Analyzer. Også operativsystemets dato/klokke-funksjon ble benyttet, og filsystem ble benyttet for å endre filendelsene på testdokumentene. Andre programvarer, som scrubbere, ble ikke benyttet for å gjennomføre testene. Det ble heller ikke undersøkt nærmere hvordan fil- og operativsystem påvirker metadata.

Testene har derimot besvart spørsmålene i problemstillingen og på den måten vist hvilke metadata som *kan* lagres. Testresultatene svarte også på problemstillingens spørsmål om hvilke potensielle aktører som kan manipulere metadata og hvordan dette kan gjøres.

6.3.3 Resultater

I problemstillingen ble det stilt spørsmål om metadata kan sikre validiteten til digitale dokumentbevis. Resultatene fra testene ga grunnlag for diskusjonen i avsnitt 7.1. Konklusjonen er at metadata *ikke* sikrer validiteten til digitale dokumentbevis nettopp på grunn av testresultatene.

I digital etterforskning benyttes ofte spesiell programvare for å analysere digitalt materiale. Spesielt speilings- og analyseverktøyet EnCase Forensic er utbredt blant private og offentlige etterforskningsinstitusjoner. Resultatene i denne oppgaven er basert på tester der dette verktøyet *ikke* er brukt. Analyseverktøy som EnCase Forensic kan potensielt oppdage manipulering av den typen som er benyttet i forsøkene beskrevet i kapittel 4. Testresultatene i denne oppgaven viser imidlertid at manipulering er mulig og at man ikke kan oppdage at dokumenter opprettet i MS Word 2003 er manipulert med det blotte øye.

Kapittel 7

Konklusjon

Denne oppgaven har tatt for seg digitale dokumentbevis. I kapittel 2 ble det gitt en introduksjon til norsk straffeprosess, og relevante forskningsprosjekter ble presentert i kapittel 3. Spørsmålene i denne oppgaven har blitt besvart av å utføre tester på digitale dokumenter. Metode og verktøy for dette er beskrevet i kapittel 4, og resultatene ble presentert i kapittel 5. Kapittel 6 inneholder diskusjonen rundt problemstillingen og resultatene, og i dette kapittelet presenteres konklusjonene jeg har kommet frem til.

I avsnitt 7.1 gjennomgår jeg delproblemstillingene og konklusjonene for hvert spørsmål. Konklusjonen for hovedproblemstillingen presenteres sist i avsnitt 7.1. I avsnitt 7.2 foreslår jeg muligheter for videre arbeid med relevante problemstillinger innen digital etterforskning og bruken av digitale bevis.

7.1 Problemstilling

I norsk strafferett er det påtalemyndigheten som sitter med bevisbyrden i straffesaker, og denne har plikt til å belyse en sak best mulig. Derfor fremlegges ofte mange bevis som det deretter er opp til dommeren å godta eller avskjære. Dommeren må ha kunnskap om beviset, og beviset må ha validitet for at det skal bli godtatt. Det må dessuten være autentisk og ekte.

Et digitalt dokumentbevis må analyseres for at validiteten kan bestemmes. Dette kan gjøres ved å se på dokumentets verdikjede. Den beskriver historien til beviset. Hendelser som inntreffer dokumentet blir trigget av en eller flere aktører, og dette kan generere eller endre metadata. Metadata er strukturerte data om det digitale dokumentet. Dataene er strukturert inn i ett og ett metadata-element, og inneholder informasjon om for eksempel forfatteren av dokumentet eller tidspunkt for opprettelse.

Oppgaven omhandler slike metadata og hva de kan fortelle om et digitalt dokument. Spørsmålet som ble reist i kapittel 1 var:

P: Kan metadata sikre validiteten til digitale dokumentbevis?

Dette spørsmålet ble videre konkretisert og delt opp i fire nye delspørsmål:

P_1: Hvilke metadata lagres?

P_2: Hvilke aktører genererer metadata?

P_3: Hvordan kan metadata manipuleres?

P_4: Hvilke aktører kan manipulere metadata?

Oppgaven ble begrenset til å gjelde digitale bevis og spesielt digitale dokumentbevis opprettet i Microsoft Word 2003. I tillegg dreier oppgaven seg hovedsaklig om den strafferettslige prosessen. Spørsmålene i problemstillingen ble besvart med resultater fra en rekke tester og analyser. To testdokumenter ble opprettet, og testdokumentenes metadata ble undersøkt ved hjelp av flere programvareverktøy.

7.1.1 Metadatafunn

P_1: Hvilke metadata lagres?

Resultatet fra testene viste at det i følge Microsofts dokumentasjon kan lagres 26 spesifikke metadata-elementer. Ved opprettelse av test-dokumentene ble henholdsvis 15 og 16 metadata-elementer lagret og tilordnet verdi.

7.1.2 Generering av metadata

P_2: Hvilke aktører genererer metadata?

Resultatene viste at det først og fremst er programvare som genererer metadata. Digitale dokumenter kan opprettes både i XML-editorer og tekstredigeringsapplikasjoner som MS Word 2003, og disse to står for en stor del av genereringen.

Operativsystem kan også bidra til generering av metadata. I testene som ble gjennomført bidro operativsystemet Microsoft Windows XP med verdier til tidsstemplene som ble tilordnet enkelte av metadata-elementene.

Personrolle-aktører som forfatter av dokumentet, kan også generere metadata. Dette er mulig fordi MS Word 2003-dokumenter kan opprettes i en XML-editor. Dersom dette gjøres kan et variabelt metadata-element legges til, CustomDocumentProperties. Dette elementet fungerer som en beholder for andre elementer som forfatteren helt fritt kan definere.

7.1.3 Manipulering av metadata

P_3: Hvordan kan metadata manipuleres?

P_4: Hvilke aktører kan manipulere metadata?

Slik som metadata kan genereres i flere ulike programvarer, kan de også manipuleres av disse. I testene ble metadata manipulert både ved hjelp av tekstredigeringsapplikasjonen MS Word 2003 og XML-editoren Exchanger XML Lite. Filsystem og operativsystem bidro også til å manipulere innholdet i flere av elementene.

Tydeligst av alle aktørene som manipulerte metadata i testene var personrolle-aktøren forfatter. Som forfatter av testdokumentene var det mulig å manipulere de aller fleste metadata-elementene. Ved hjelp av filsystem og operativsystem ble også elementer som var tidsstemplet manipulert.

Elementene som inneholdt statistisk informasjon, som antall ord eller antall sider, blir ikke manipulert av personrolleaktører fordi elementene inneholder verdier som beregnes og dermed manipuleres av tekstredigeringsapplikasjonen.

7.1.4 Sikring av validiteten til digitale dokumentbevis

P: Kan metadata sikre validiteten til digitale dokumentbevis?

For at metadata skal sikre validiteten til digitale dokumentbevis må man være sikker på at metadataene er korrekt, ekte og ikke manipulert med. Testene foretatt i denne oppgaven viser at metadata-elementene kan manipuleres, men at manipuleringen ikke kan oppdages med det blotte øye. Derfor kan ikke metadata sikre validiteten til digitale dokumentbevis.

Metadata sikrer *ikke* validiteten til digitale dokumentbevis. Metadata kan enkelt manipuleres. Selv elementer som man ofte har stor tillit til, som tidsstempler, kan manipuleres og er derfor ikke med på sikre validiteten til digitale dokumentbevis. Sakkyndige kan heller ikke uten videre oppdage manipulasjon. Til tross for at digitale bevis har mange fordeler, er problemene knyttet til autentisitet, ekthet og validitet store ulemper ved denne bevistypen.

7.2 Videre arbeid

I dette avsnittet gir jeg forslag til videre arbeid innenfor dette feltet som inkluderer både juridiske og teknologiske fagfelt. Det blir gitt forslag til formater, metoder og verktøy som kan benyttes til fremtidige undersøkelser.

7.2.1 Formater

Dokumenter opprettet i tekstredigeringsapplikasjonen MS Word 2003 har blitt undersøkt i denne oppgaven. Microsoft har kommet med en ny versjon av programvaren, MS Word 2007. Denne nyeste versjonen er basert på XML og bygget opp på en annen måte enn 2003-versjonen av programvaren. Det vil derfor kunne være svært interessant og aktuelt å se på 2007-versjonens håndtering av metadata og mulighetene for manipulering av disse som en forlengelse av denne oppgaven.

Samtidig vil det kunne være interessant å sammenligne metadata-håndteringen i Microsofts programvare med for eksempel programvare fra OpenOffice. Er det stor forskjell i håndteringen, og kan metadata manipuleres også i tekstredigeringsapplikasjonen OpenOffice Writer?

Det er også essensielt å undersøke hvordan validiteten til digitalt materiale på andre formater påvirkes av metadata. Hvordan håndteres metadata i ulike e-postformater? Hvordan påvirker fil- og operativ-systemer metadata? Kan systemlogger som inneholder metadata manipuleres? Svarene på disse spørsmålene vil utgjøre en del av dokument-asjonen av hvilke metadata som lagres i ulike digitale formater. Svarene vil også gi informasjon om potensielle metoder for manipulering av metadata og i hvilken grad metadataene kan sikre validiteten for digitale formater som faller utenfor fokuset for denne oppgaven.

7.2.2 Verktøy

Testene i denne oppgaven gjør kun bruk av tre verktøy foruten fil- og operativsystem. Det vil derfor kunne være interessant å se hvordan ulike scrubbere kan manipulere verktøy. Videre vil det være svært interessant og viktig å undersøke hvor mye av manipuleringen som kan oppdages ved hjelp av et analyseverktøy som for eksempel EnCase Forensic.

Bibliografi

[1] Convention on Cybercrime. november 2001. http://conventions.coe.int/Treaty/EN/Treaties/Html/185.htm Aksessert 30. april 2007.

[2] Habtamu Abie, Jerker Danielsen, Arne-Kristian Groven, Ingvar Tjøstheim og Åsmund Skomedal. A structured approach to digital forensic readiness. Rapport, Norwegian Computing Center.

[3] Johs. Andenæs. *Norsk Straffeprosess*. Universitetsforlaget, tredje utgave, 2000.

[4] Brian Carrier. Open source digital forensic tools - the legal argument. *At stake*, september 2003.

[5] Line Coll. Rettslige spørsmål knyttet til inssamling og bruk av digitale bevis. 2004.

[6] DCMIs 22 kjernemetadataelementer. http://dublincore.org/documents/usageguide/elements.shtmlAksesset februar 2007.

[7] DCMIs definisjon av "metadata". http://www.dublincore.org/resources/faq/#whatismetadata Aksesset januar 2007.

[8] *Fremmedordbok*. Kunnskapsforlaget, femtende utgave, 1986.

[9] Arve Føyen. Digitale bevis. juni 2006.

[10] John D. Gregory. Authentication rules and electronic records. *The Canadian Bar Review*, 81:529–562, 2002.

[11] Chat Hosmer. Proving the integrity of digital evidence with time. *International Journal of Digital Evidence*, 1(1), våren 2002.

[12] Fredesvinda Insa. The admissibility of electronic evidence in court: Fighting against high-tech crime. Rapport, Cybex, 2006.

[13] Justisdepartementet. Kartleggingsprosjektet - kartlegging av bestemmelser i lover, forskrifter og instrukser som kan hindre elektronisk kommunikasjon. Rapport, juni 2000.

[14] Erin E. Kenneally. Gatekeeping out of the box: Open source software as a mechanism to assess reliability for digital evidence. *Virginia journal of law and technology*, 2001.

[15] Stefan Kronqvist. *Brott och digitala bevis – En handledning*. Norstedts Juridik AB, 2003.

[16] Stephen Mason. Proof of the authenticity of a document in electronic format introduced as evidence. Rapport, ARMA International Educational Foundation, oktober 2006.

[17] Alfred J. Menezes, Paul C. van Oorschot og Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, oktober 1996.

[18] Microsoft Corporation. *Microsoft Office 2003 Edition XML Schema References*, oktober 2004.

[19] United Nations. Uncitral model law on electronic commerce with guide to enactment 1996. 1999.

[20] NOU 2003:27 – Lovtiltak mot datakriminalitet – Delutredning I.

[21] NOU 2007:2 – Lovtiltak mot datakriminalitet – Delutredning II.

[22] National Information Standards Organization. Understanding metadata. NISO Press, 2004.

[23] John Patzakis og Victor Limongelli. Encase legal journal. desember 2004.

[24] Justis og politideparetementet. Høring: Tilrettelegging for elektronisk kommunikasjon med domstolene - forslag til endirnger i rettspleielovene. 2001.

[25] Rt 1991. side 616.

[26] Rt 1992. side 904.

[27] Inger Marie Sunde. *Lov og rett i cyberspace*. Nr 16 i ØKOKRIMs skriftserie. Fagbokforlaget Vigmostad og Bjørke AS, 2006.

[28] Lars Christian Sunde. Elektronisk dokumentfalsk. juni 2004.

[29] SWGDE. Data integrity within computer forensics. april 2006.

[30] Ingvar Tjøstheim og Jerker Danielson. Defining Standards in Digital Forensics. august 2005. http://www.nr.no/pages/dart/projects_security_desdifor Aksessert 01. mars 2007.

[31] Christine Ulrichsen. Kommentar til: Høyesterett - kjennelse (gatekjøkkenkjennelsen). *Rt 1991*, 1991. http://www.personvern.uio.no/pvpn/avgjorelser/kommentarer/Ulrichsen.html Aksessert 12. mars 2007.

[32] Wikipedias definisjon av "digital". http://en.wikipedia.org/wiki/Digital Aksesset januar 2007.

[33] Svein Y. Willassen. Time-stamps, digital traces and forensic evidence (TID). Rapport, juni 2004.

[34] Svein Yngvar Willassen og Stig Frode Mjølsnes. Digital Forensics Research. *Telektronikk*, (1), 2005.

[35] ANSIs nettsted: http://www.ansi.org.

[36] ARMAs nettsted: http://www.armaedfoundation.org/.

[37] Advokat Arve Føyens nettsted: http://www.foyen.no/templates/CV____265.aspx.

[38] British institute of international and comparative laws nettsted: http://www.biicl.org/.

[39] CARTs nettsted: http://www.fbi.gov/hq/lab/org/cart.htm.

[40] Cybex' nettsted: http://www.cybex.es/defaulten.aspx.

[41] Digital Evidence Research Programmes nettsted: http://www.biicl.org/digitalevidence/.

[42] Digital Forensic Research Workshop (DFRWS)s nettsted: http://www.dfrws.org/index.html.

[43] Dublin Core Metadata Initiatives nettsted: http://www.dublincore.org.

[44] European Network of Forensic Science Institutes' nettsted: http://www.enfsi.org/.

[45] FBIs nettsted: http://www.fbi.gov.

[46] International Journal of Digital Evidences nettsted: http://www.ijde.org/.

[47] Interpols nettsted: http://www.interpol.int.

[48] International Organisation on Computer Evidences nettsted: http://www.ioce.org.

[49] Økokrims nettsted: http://www.okokrim.no.

[50] Metadata Analyzers nettsted: http://www.smartpctools.com/metadata/index1.html.

[51] Microsofts nettsted: http://www.microsoft.com/en/us/default.aspx.

[52] NISOs nettsted: http://www.niso.org.

[53] Om RDF: http://www.w3.org/RDF/.

[54] SWIGDEs nettsted: http://www.ncfs.org/swigde/aboutUs.html.

[55] W3Cs nettsted: http://www.w3.org/.

[56] Svein Y. Willassens nettsted: http://www.willassen.no/.

[57] Exchanger XML Editors nettsted: http://www.freexmleditor.com/.

[58] Brian D. Zall. Metadata: Hidden information in Microsoft Word documents and its ethical implications. *The Colorado Lawyer*, 53(10), oktober 2004.

Tillegg A

Microsoft Word 2003 XML-skjema

Detaljer om Microsoft Word 2003 XML-skjemaer omtalt i denne oppgaven beskrives i figur A.1 og A.2 på henholdsvis side 98 og 99.

Common Properties Schema Overview

The elements in these schemas are shared by Microsoft Office Word 2003 and Microsoft Office Excel 2003.

Namespace	urn:schemas-microsoft-com:office:office
Description	Shared office schema. Document properties collection (author, createdate, etc.) and document statistics (word count, etc.).
Prefix	o
Friendly Name	Common Properties

Figur A.1: Skjemaet "Common Properties"

Word Schema Overview

The following schemas are used with WordProcessingML, an XML dialect useful for extracting information from Microsoft Office Office Word 2003. These are important to document developers and to application developers whose programs will read and write WordprocessingML documents.

Namespace	http://schemas.microsoft.com/office/word/2003/wordml
Description	This namespace defines the core elements and attributes used to describe documents in Microsoft Office Word 2003.
Prefix	w
Friendly Name	XML Document 2003

Namespace	http://schemas.microsoft.com/schemaLibrary/2003/core
Description	Contains two elements, SchemaLibrary and Schema, for the schemas referenced by a Word document.
Prefix	sl
Friendly Name	Schema Library

Namespace	http://schemas.microsoft.com/aml/2001/core
Description	Elements used to describe tracked changes (insertions, deletions, and formatting changes), comments, and bookmarks in Word documents only.
Prefix	aml
Friendly Name	Annotation Markup Language

Namespace	http://schemas.microsoft.com/office/word/2003/auxHint
Description	Supplemental XML information added in 2003 that serves as "hints" for transformation of a Word document to HTML format; for example, font, color, border, and absolute positioning information.
Prefix	wx
Friendly Name	Auxiliary XML Document 2003

Namespace	urn:schemas-microsoft-com:office:word
Description	Supplemental XML information added to VML output that serves as a container for additional properties used by Word: for example: relative/absolute positioning, anchoring, text wrapping.
Prefix	w10
Friendly Name	Auxiliary XML Document 2002

Namespace	urn:schemas-microsoft-com:office:vml
Description	VML is an application of Extensible Markup Language (XML) 1.0 which defines a format for the encoding of vector information together with additional markup to describe how that information may be displayed and edited.
Prefix	v
Friendly Name	Vector Markup Language

Figur A.2: Skjemaet "Word"

Tillegg B

SWGDE Disclaimer

Figur 2.3.2 i avsnitt 2.1 illustrerer fem steg for å sikre dataintegriteten til digitale bevis. Figuren er hentet fra Scientific Working Group on Digital Evidence, og denne gruppen krever derfor at informasjonen på side 103 blir gitt i denne oppgaven.

Scientific Working Group on
Digital Evidence

Data Integrity Within Computer Forensics

Disclaimer:

As a condition to the use of this document and the information contained therein, the SWGDE requests notification by e-mail before or contemporaneous to the introduction of this document, or any portion thereof, as a marked exhibit offered for or moved into evidence in any judicial, administrative, legislative or adjudicatory hearing or other proceeding (including discovery proceedings) in the United States or any Foreign country. Such notification shall include: 1) The formal name of the proceeding, including docket number or similar identifier; 2) the name and location of the body conducting the hearing or proceeding; 3) subsequent to the use of this document in a formal proceeding please notify SWGDE as to its use and outcome; 4) the name, mailing address (if available) and contact information of the party offering or moving the document into evidence. Notifications should be sent to swgde@mail.ucf.edu

Redistribution Policy:

SWGDE grants permission for redistribution and use of all publicly posted documents created by SWGDE, provided that the following conditions are met:

1. Redistributions of documents or parts of documents must retain the SWGDE cover page containing the disclaimer.
2. Neither the name of SWGDE nor the names of contributors may be used to endorse or promote products derived from its documents.
3. Any reference or quote from a SWGDE document must include the version number (or create date) of the document and mention if the document is in a draft status.