

Om speilkopiering av data i sivile tvister

av Svein Y. Willassen

Bevis som finnes i datalagret informasjon kan ha stor betydning i mange saker. Moderne kommunikasjon foregår elektronisk og på en slik måte at mesteparten av den informasjon som produseres kun finnes datalagret. Eksempler på slikt materiale er epost, elektroniske dokumenter, herunder avtaler, presentasjoner, produktbeskrivelser og liknende. Videre er det slik at moderne produksjon i de fleste bedrifter er styrt av datasystemer, og forskning og utvikling av produkter og tjenester skjer for det meste ved hjelp av datasystemer. For svært mange arbeidstagere i dag består store deler av arbeidsdagen av bruk av datasystemer og samhandling med andre ved hjelp av datasystemer. Det er derfor åpenbart at datasystemer inneholder informasjon som kan ha betydning som bevis i praktisk talt alle typer rettslige tvister, og at dette er bevis som ikke kan finnes eller dokumenteres på noen annen måte enn å analysere datasystemer.

Samtidig er det slik at bevis som er lagret i datasystemer er svært sårbare. Datafiler kan kopieres, flyttes, manipuleres eller slettes ved hjelp av få museklikk. Enhver vanlig bruk av et datasystem vil også medføre at dataene som er lagret på datasystemet endres. For eksempel vil registreringen av tidspunktet for når et dokument sist er lest bli oppdatert dersom man leser et dokument. Denne egenskapen medfører at man bør bruke spesielle fremgangsmåter ved sikring og analyse av bevis som er lagret i et datasystem. Når man står overfor en datamaskin som potensielt inneholder bevis av betydning er det ikke en tilfredsstillende fremgangsmåte å starte opp datamaskinen og begynne å lete etter bevis på den. Denne

fremgangsmåten vil for det *første* endre data som ligger lagret på datamaskinen slik at man ikke etterpå kan gjenske hvilke data som var der da søket begynte. Det innebærer risiko for at det i ettertid kan stilles spørsmål ved troverdigheten av de analyser som er gjort. For det *andre* blir det ved en slik fremgangsmåte vanskelig å foreta en systematisk gjennomleding i datamengden for å finne data av betydning som bevis. Det blir derfor vanskelig å konkludere sikkert om hvilken mulighet man har hatt til å finne relevante bevis. For det *tredje* er det ved en slik fremgangsmåte ikke mulig å lete i data som ikke fremstår som lesbare filer, for eksempel rester av slettede filer. Løsningen som er blitt utviklet som svar på denne utfordringen er å bruke en teknikk som kalles *speilkopiering*.

Speilkopiering går i korthet ut på å bruke et spesielt utviklet dataprogram til å kopiere alle data som finnes på et medium over til et annet medium på en slik måte at innholdet ikke endres under kopieringen. Ordet *medium* brukes i denne sammenheng om en hvilken som helst bærer av digital informasjon; for eksempel en harddisk, en diskett, en CD-ROM eller en mobiltelefon. Ved analyse av datamaskiner er det normalt harddisker som analyseres. For å være sikker på å ha kontroll med at speilkopieringsprosessen ikke endrer data på harddisken, er det vanlig å benytte en egen datamaskin til selve speilkopieringen. Når man utfører speilkopiering foretar man derfor normalt en demontering av datamaskinen som inneholder bevis. Harddisken tas ut og monteres i speilkopieringsmaskinen. Denne maskinen startes og programmet som utfører selve speilkopieringen



kjøres. Dette programmet kopierer alle data fra kildeharddisken over til en tom harddisk. Når speilkopieringen er ferdig er innholdet på måldisken identisk lik innholdet på kildedisken. Hvor lang tid speilkopieringen tar avhenger av harddiskens størrelse. Med størrelser som er vanlig i dagens PC-er (50-100 Gb) tar kopieringen vanligvis fra 30 til 120 minutter per datamaskin. Når speilkopiering er gjennomført kan man levere tilbake originalmaskinen, og foreta analyse av kopien.

På alle datasystemer er datalagret informasjon organisert i filer, som igjen er katalogisert i et filsystem. Hver fil har et filnavn, en filtype og *tidsstempler* som viser når filen er opprettet, oppdatert og sist lest. På samme måte har kataloger også navn og tidsstempler. Alle kataloger og filer er lenket inn i et filsystem som består av en trestruktur. Brukeren vedlikeholder selv denne trestrukturen ved å gi navn til kataloger og filer, mens tidsstemplene på filer og kataloger settes automatisk fra datamaskinens klokke når filer eller kata-

loger omfattes av brukerens handlinger.

Et naturlig start ved analyse av databevis er derfor å ta utgangspunkt i filsystemet slik det fremstår på sikringstidspunktet. Dette gjøres ved hjelp av spesiell programvare som henter frem filsystemet fra en speilkopi uten å forandre på innholdet i filene. Ved å gå gjennom katalognavn, filnavn og tidsreferanser kan man i de fleste saker relativt raskt få et overblikk over hvilken type informasjon som finnes lagret i filsystemet. Når man så skal vurdere hvilke filer som kan inneholde informasjon som har betydning som bevis i den aktuelle saken har man flere metoder tilgjengelig:

- sortere alle filene etter navn, se etter navn som er relevant for saken
- sortere filene etter filtype, se etter filtyper som er relevant, for eksempel regneark i en sak om regnskapsovertredelse eller bilder i en sak om seksuelle overgrep
- sortere filene etter tidspunkt, se etter filer som er opprettet eller lagret i et tidsrom som er spesielt aktuelt i saken

Det er verdt å merke seg at ovennevnte analyse også vil inkludere en stor mengde slettede filer, idet de fleste dataetterforskningsverktøy automatisk gjenfinner referanser til slettede filer.

I den videre analyse er det mest vanlig å benytte maskinelle søk. Dette er en funksjon som er bygget inn i de fleste dataetterforskningsprogrammer. Ved et maskinelt søk definerer man et sett med søkekriterier, for eksempel 8-10 ord som ansees å være spesielt relevant for saken. Datamaskinen søker så gjennom alle data i speilkopien etter disse ordene. Med en 50 Gb harddisk vil dette ta 1-2 timer. Som resultat får man opp alle treff på de aktuelle ordene i

eksisterende filer, slettede filer samt rester av slettede filer som er delvis overskrevet av nye data. Man kan altså få treff i dataområder som man ikke hadde mulighet til å avdekke i den innledende analysen hvor man tok utgangspunkt i filene i filsystemet.

Antall treff avhenger av hvor mye data som finnes på harddisken og hvilke søkeord som er valgt. Definisjon av gode søkeord ut fra en konkret sak er derfor viktig for å unngå at man bruker svært lang tid på å gå gjennom irrelevante søketreff. Samtidig må søkeordene, og kombinasjoner av søkeord, ikke velges så snevert og spesifikt at man ved søket risikerer å forbigå relevant og potensielt avgjørende materiale. Ved godt valg av søkeord viser det seg at denne metoden er svært effektiv i praksis. Ved bruk av denne metoden har man i mange saker avdekket bevis i informasjon som var slettet og som eieren av datamaskinen trodde var borte for lengst, eller som eieren trodde aldri hadde vært lagret på datamaskinen.

I sivile tvister kan det være et problem at datamengdene kan inneholde materiale som hver av partene legitimt ikke ønsker skal tilflyte den andre parten. Problemet kan løses ved at den som utfører analysen er en kompetent person som er uavhengig i forhold til partene, eller at analysen gjennomføres av partenes representanter i fellesskap. Dersom vedkommende får et klart mandat for analysearbeidet bør det i praksis være uproblematisk å utelukke informasjon som er irrelevant eller taushetsbelagt. Dette forutsetter imidlertid at man får en beskrivelse fra begge parter om hva slags informasjon som ettersøkes og hva slags informasjon som det eventuelt skal sees bort fra. En slik beskrivelse fra begge parter kan naturligvis ikke gis uten at begge parter er klar over bevissikringen. Idet man ønsker å kunne

sikre bevis uten at det gis varsel til motparten kan det derfor være formålstjenlig å dele bevissikringen inn i to faser:

1. Sikring:

Data sikres fra relevante medier ved bruk av speilkopiering. Speilkopier deponeres hos domstolen eller representant for denne. Det gjøres ikke noe analysearbeid i denne fasen.

2. Analyse:

Data analyseres ut fra kriterier som bestemmes av partene eller dommeren.

Mellom de to fasene kan spørsmålet være gjenstand for kontradiktorisk behandling. På denne måten kan man både tilfredsstillende behovet for bevissikring uten at parten som det sikres bevis hos får anledning til å ødelegge bevis, og behovet for at begge parter får være med å spesifisere hva det skal letes etter og hva som ikke skal frembringes.

Svein Y. Willassen (32) er sivilingeniør og PhD-stipendiat innen dataetterforskning ved NTNU. Han er medlem i datakrimutvalget og styremedlem i dataetterforskningssselskapet INFOSEC ProtectIT AS.