

NINA SUNDE / OLAV DAHL

Erfaringer fra Kvalitetskontroll- prosjektet

Kan sjekklisterbasert fagfelle-
vurdering bidra til læring og
økt kvalitet på datatekniske
rapporter?



Erfaringer fra Kvalitetskontrollprosjektet

*Kan sjekklisterbasert fagfellevurdering bidra til
læring og økt kvalitet på datatekniske rapporter?*

Politihøgskolen, 04.12.2023

Nina Sunde

Olav Dahl

PHS Forskning 2023:7

Erfaringer fra Kvalitetskontrollprosjektet. Kan sjekklisterbasert fagfellevurdering bidra til læring og økt kvalitet på datatekniske rapporter?

ISBN 978-82-7808-197-6 (elektronisk utgave)

ISSN 0807-1721

Det må ikke kopieres fra denne boka i strid med åndsverkloven og fotografiloven eller i strid med avtaler om kopiering inngått med Kopinor, interesseorgan for rettighetshavere til åndsverk.

Alle henvendelser kan rettes til:

Politihøgskolen

Slemdalsveien 5

Postboks 2109, Vika

0125 Oslo

Tlf: 23 19 99 00

www.phs.no

Sammendrag

Kvalitetskontrollprosjektet ble gjennomført som et kombinert forsknings- og utviklingsprosjekt av Nina Sunde og Olav Dahl høsten 2022 til høsten 2023. Politidistrikter og særorgan ble invitert til en serie med workshops for å gjøre seg kjent med det forskningsutviklede Kvalitetskontrollhierarkiet – et nivå delt rammeverk med tilhørende sjekklister for fagfelleevaluering og kvalitetskontroll av datatekniske rapporter. Etter invitasjon takket syv politidistrikter/særorgan ja til å delta i Kvalitetskontrollprosjektets pilotfase for å teste ut fagfelleevaluering og kvalitetskontroll av datatekniske rapporter ved hjelp av sjekklister, som er sentrale verktøy i dette rammeverket. Piloteringen foregikk i perioden januar - juni 2023. Deltakerne, som var medarbeidere ved enheter for digitalt politiarbeid tiltenkt rolle som fagfelle og deres ledere, ble så intervjuet om erfaringene med fagfelleevaluering og kvalitetskontroll i perioden, og om deres tanker om kvalitet og rettssikkerhet mer generelt.

Studien viser at en fagfelleevaluering er egnet til å fange opp feil i datatekniske rapporter. Samtidig er det klart at noen feil kan gå under radaren dersom fagfelleevalueringen kun omfatter lesing av rapporten framfor å verifisere funnet i beslaget eller gjøre nye undersøkelser. Erfaringene fra Kvalitetskontrollprosjektets pilotfase viser også at prosessen med fagfelleevaluering og tilbakemelding til rapportskriver gir god læring både for fagfellen som utfører kvalitetskontrollen og for rapportskriver som mottar tilbakemelding. Studien peker tydelig i retning av at ledelse og forankring av kvalitetsarbeidet er viktig for å bidra til å bygge en kvalitetskultur, samt å sørge for at den som gjennomfører fagfelleevaluering har riktig kompetanse og tilstrekkelige rammer for å gjennomføre arbeidet. Deltakerne er samstemte om at kvalitetskontroll av datatekniske rapporter er nødvendig for å styrke rettssikkerheten og at det bør gjennomføres som en obligatorisk og systematisk aktivitet.

Studien viser at digitale spor blir håndtert av mange ulike roller i straffesakskjeden, og at manglende rolleavklaring, kompetansekrav, regulering og styring i prinsippet innebærer at 'alle kan gjøre alt'. Resultatene peker på behov for en større grad av regulering og standardisering av hvordan digitale spor og bevis blir håndtert i straffesakskjeden.

Deltakerne har gitt beskrivelser av en mengde feil og nesten-feil fra etterforskninger i egne politidistrikt, og er samstemte i sin bekymring om at systemet i dag ikke er rigget for å forebygge eller avdekke feil i tilstrekkelig grad. Studien har avdekket et behov for rutiner og systemer når kritiske feil blir oppdaget. Det er behov for avklaring av roller og ansvar for kartlegging av hvilke følger feilene har fått i avsluttede og pågående etterforskninger, og for å hindre at feilen kan påvirke framtidige straffesaker.

Innhold

Sammendrag	3
1. Introduksjon	7
1.1 Prosjektorganisering	7
1.2 Bakgrunn	7
1.2.1 Behov for kvalitetssystemer	11
2. Metode	15
2.1 Innføringsfasen	15
2.1.1 Oversettelse og tilpasning av sjekklister.....	16
2.1.2 Kartlegging	16
2.1.3 Workshop I og II - opplæring i Kvalitetskontrollhierarkiet og tilpasning av sjekklister.....	17
2.2 Pilotfasen	21
2.2.1 Oppstartsmøter	22
2.2.2 Antall fagfellevurderinger og tilbakemeldinger i pilotfasen	24
2.2.3 Framdrift.....	24
2.3 Innsamlings- og analysefasen.....	25
2.3.1 Datainnsamling	25
2.3.2 Analyse	26
2.3.3 Workshop III – formidling av resultatet fra pilotfasen.....	27
2.4 Begrepsavklaring	27
3. Resultat og analyse	30
3.1 DPA-medarbeideres, fagkontakters og etterforskeres rolle i dataetterforskningsprosessen	30
3.1.1 DPA-medarbeiderens rolle i endring.....	30
3.1.2 Digitalt politiarbeid og cyberkriminalitet.....	35
3.1.3 Etterforskerens og fagkontaktens rolle.....	35
3.2 Erfaringer fra Kvalitetskontrollprosjektets pilotfase	39
3.2.1 Tidligere erfaring med kvalitetskontroll	39
3.2.2 DPA-ledernes rolle i pilotfasen.....	40
3.2.3 Sjekkliste og støttedokument.....	41
3.2.4 Tilbakemelding	46

3.2.5 Blind eller åpen fagfellevurdering	48
3.2.6 Nivå i hierarki og fagfellens kompetanse	49
3.2.7 Punkter fra sjekklisten til diskusjon.....	51
3.2.8 Endring av egen og lokal praksis	59
3.3. Eksempler på feil og håndteringen av disse	60
3.3.1 Feil	60
3.3.2 Håndtering og varsling	68
3.4. Veien videre	69
3.4.1 Holdninger til kvalitetskontroll	69
3.4.2 Om å bygge kvalitets- og tilbakemeldingskultur	72
3.4.3 Motstand i retten	73
3.4.4 Tid og ressurs.....	73
3.4.5 Standardisering	75
3.4.6 Kompetansebehov	75
4. Diskusjon	78
4.1 Kan fagfellevurdering og kvalitetskontroll bidra til å heve kvaliteten og avdekke kritiske feil og/eller villedende informasjon?.....	78
4.2 Kan fagfellevurdering og kvalitetskontroll skape læring hos fagfelle og/eller mottaker av fagfellevurdering?	81
4.3 Kan fagfellevurdering og kvalitetskontroll lede til systemforbedring / organisasjonslæring?.....	82
4.4 Kan fagfellevurdering og kvalitetskontroll gi uønskede sideeffekter?	84
5. Oppsummering og anbefaling.....	85
5.1 Oppsummering	85
5.2 Kortsiktige anbefalinger	86
5.3 Langsiktige anbefalinger.....	87
Referanser	89
Vedlegg.....	96

1. Introduksjon¹

1.1 Prosjektorganisering

Kvalitetskontrollprosjektet ble gjennomført høsten 2022-høsten 2023. Det ble ledet av politioverbetjent (PhD) Nina Sunde (Politihøgskolen)² i samarbeid med seksjonsleder og PhD stipendiat Olav Dahl (Brann- og redningsskolen/NTNU) - som i det videre omtales som 'prosjektgruppen'. Student og spesialletterforsker Rune Kenneth Bauge fra Vest Politidistrikt bidro med datainnsamlingen i forkant av pilotprosjektet i forbindelse med sin masteravhandling ved NTNU (Bauge, 2023).

1.2 Bakgrunn

Digitale spor og bevis inngår i dag i så å si alle straffesaker. Dette har trolig sammenheng med de digitale sporenes tilgjengelighet kombinert med deres egnethet til å belyse ulike informasjonsbehov og bevistema i etterforskning. Samtidig har dataetterforskningsområdet blitt stilt overfor utfordringer som store datamengder, kompleksitet, ny teknologi og nye brukermønstre (Cervantes Mori et al., 2021; Reedy, 2019, 2022; I. M. Sunde, 2021 s. 21; Vince, 2016). Dette har ført til et vedvarende press på økt kompetanse og evne til å sikre og analysere digitale spor slik at de kan anvendes som ledetråder i etterforskning eller bevis i straffesaken.

Utviklingen har krevd et høyt fokus på effektivitet i håndteringen av digitale spor og bevis, noe som kan ha gått på bekostning av oppmerksomhet på struktur og kvalitet. Digitale bevis håndteres av mange ulike roller i straffesakskjeden. Samtidig mangler norsk

¹ Takk til fagfelle Silje A. Bakken for grundig lesing og konstruktive forslag til forbedring av rapporten.

² Professor Johanne Yttri Dahl v. Politihøgskolen var formell leder på prosjektet inntil Nina Sunde hadde oppnådd PhD graden. Fra og med 25.11.22 var Nina Sunde formell leder for prosjektet.

politi en kompetanseprofil for de involverte rollene som tydelig definerer hva de har kompetanse til å gjøre og hvor arbeidet må overlates til en med mer spesialisert kompetanse. Selv om mye godt arbeid er på gang i fagforvaltningsapparatet, hersker det fortsatt forvirring rundt begreper, prosesser, roller og beste praksis for utøvelse av dataetterforskningsprosessens ulike faser. Studier viser at det gjøres per i dag ingen systematisk, dokumentert og forpliktene kvalitetskontroll av undersøkelser som gjennomføres av datateknisk etterforsker eller taktisk etterforsker (se f.eks. Bauge, 2023; Jahren, 2020). Dette ble også påpekt av Riksrevisjonen (2021) som ga følgende beskrivelse av situasjonen: «Undersøkelsen viser at flere politidistrikter mener det er behov for retningslinjer og rutiner for det digitale politiarbeidet. Utarbeidelse av rutiner og veiledere blir i for stor grad opp til det enkelte distrikt og det etterlyses tydeligere nasjonale føringer for hvordan digitale bevis skal innhentes, analyseres og etterforskes. Mange enheter for digitalt politiarbeid savner støttesystemer for å utføre oppgavene funksjonen har ansvar for» (pkt. 2.1.3, s. 8). Forskning som i stor grad omfatter norske dataetterforskere (N. Sunde, 2021) og reelle norske straffesaker (Stoykova et al., 2022) har avdekket at det er til dels store mangler i datatekniske rapporter. Forskningen viser at det trekkes slutninger omkring funnene som kan være direkte feil eller villedende med hensyn til databevisets relevans og verdi for straffesaken (N. Sunde, 2022c). Det er også funnet store mangler i dokumentasjonen omkring undersøkelsesprosessen, noe som dermed kan medføre usikkerhet omkring bevisets pålitelighet (Stoykova et al., 2022; N. Sunde, 2021, 2022b).

Overdreven tiltro til påliteligheten av teknologiske systemer og verktøy har ført til flere store rettsskandaler i nyere tid. Teledata-skandalen (se for eksempel Lentz & N. Sunde, 2021; Søre, 2019) og den britiske Post Office Horizon-skandalen (Helm, 2022) er eksempler på hvordan feil i systemer forblir uoppdaget eller bortforklart, og at justisfeil rammer uskyldige over tid. I 2019 ble det oppdaget at et system for

prosessering av innhentede teledata ved Rigspolitiets Telecenter i Danmark genererte feil og ufullstendige data. Umiddelbart førte oppdagelsen av feilen til at Danmarks Riksadvokat beordret full stopp i bruk av teledata i danske rettssaker i to måneder. Følgene av dette var blant annet at 50 straffesaker ble utsatt, og 35 personer ble løslatt fra varetekt. Undersøkelsene avdekket at systemet hadde generert feil siden 2012. Hele 10 700 straffesaker måtte derfor gjennomgås på nytt for å undersøke om feil i systemet som prosesserte innhentede teledata hadde påvirket utfallet av straffesaken. Ifølge daværende justisminister Nick Hækkerup kostet gjennomgangen 73 millioner danske kroner (Justitsministeriet, 2021). Pressen har stilt spørsmål om feilaktige og unøyaktige teledata villedet politiet i etterforskningen av drapet på Emilie Meng. I denne saken førte prosesseringsfeilen ved Rigspolitiets Telecenter til at politiet manglet data for en hel time i tidsrommet hvor Meng forsvant. I tillegg viste det seg at teledataene inneholdt feil opplysninger om basestasjoners faktiske plassering. Undersøkelser har vist til at det var avvik fra oppgitt og faktisk plassering av basestasjoner fra 200 meter til 82 kilometer (Kjær, 2019).

I England ledet avsløringen av feil i regnskapssystemet Horizon (utviklet av det japanske selskapet Fujitsu) til det som omtales som Englands største rettsskandale (Nokes & Moorhead, 2023). Det britiske Post Office tok i bruk Horizon-systemet i 1999. Ifølge en oversiktsartikkel publisert av BBC (Peachy, 2022) begynte systemet å flagge avvik i regnskapet, noe som indikerte at penger var borte. Selv om det ble avdekket mange problemer med Horizon, ble det lagt til grunn at systemet var feilfritt, og i årene 2000 til 2014 ble 736 Post Office-ansatte dømt for økonomisk kriminalitet på grunnlag av data fra regnskapssystemet. Forsvarsadvokater som ba om innsyn i dokumentasjon om Horizon-systemets pålitelighet ble nektet adgang til dette av britiske domstoler. Først i 2019 ble feilen anerkjent av engelsk høyesterett etter to gruppesøksmål fremmet av tidligere Post Office-ansatte (House of Commons, 2022). Dette har ledet til granskning og gjenåpning av flere saker. Per i dag har 93 post-ansatte

(per 2. oktober 2023) fått sine dommer opphevet, og til sammen 21,6 millioner britiske pund er blitt utbetalt (Post Office Corporate, 2023). Flere saker er til behandling eller venter på avgjørelse.

Også i Norge har vi sett eksempler på systematiske feil i systemer og verktøy for analyse av databaseslag. I 2020 ble det kjent at det at en versjon av dataverktøyet Cellebrite feiltolket tidsstempler ved iOS-telefoner som benyttet filsystemet AFIS, og at verktøyet var benyttet i minst 59 straffesaker (Grut, 2020). Riksadvokaten anmodet politidistrikter og særorgan til å undersøke om feilen hadde hatt betydning for utfallet av saken, og konkluderte senere med at det ikke var påvist feil i saker som ikke allerede var blitt henlagt (Riksadvokaten, 2021).

I de senere år har mediene omtalt flere straffesaker hvor digitale bevis har spilt en rolle for feilaktige dommer eller som har medført feil underveis i etterforskningen. For eksempel ble en mann fra Trøndelag siktet for å ha bestilt direkteoverførte overgrep på nett. Media fikk anledning til å være med politiet og dokumentere aksjonen hvor siktede ble pågrepet og senere varetektsfengslet. Etterforskningen viste at siktede var uskyldig, og etter nesten fire uker i varetekt ble han løslatt. Saken ble henlagt som «intet straffbart forhold» med begrunnelse i at bevisene viste med særlig styrke at han ikke hadde begått forholdet han var siktet for (Bjerkan, 2022). I intervju forteller mannen om at politiet hadde påført ham stor personlig belastning og selvmordstanker, blant annet på grunn av mange ubehagelige meldinger fra folk som hadde gjenkjent ham i mediasaken (Stølan, 2022). Et annet eksempel gjelder en feilaktig dom for sovevoldtekt i Trondenes tingrett (TTRES-2017-155163). Anmelder hadde selv overlevert en samtalelogg mellom henne og den angivelige gjerningsmannen til politiet som dokumenterte at hun hadde vært ute av stand til å motsette seg sex og at hun ble voldtatt. Mannen ble dømt i tingretten. Det ble så klart at loggen kun representerte et utvalg av fornærmedes SMS-aktivitet i tidsrommet. Mannen ble frikjent i

Hålogaland lagmannsrett (LH-2018-22766) etter at den fullstendige loggen ble lagt frem, som viste SMS-aktivitet mellom fornærmedes og hennes kjæreste i tidsrommet hvor sove-voldtekten angivelig skulle ha funnet sted (se også Hultgren, 2018).

1.2.1 Behov for kvalitetssystemer

Marit Gjerde var den første norske med politibakgrunn som tok master i digital forensics ved University College Dublin (2007). Hun undersøkte dataetterforskere og hvilke faktorer som hadde betydning for effektiv etterforskning av sedelighetssaker, med spesielt fokus på saker fra Operasjon EMEA. Hun fikk svar fra 34 av de 45 som den gang jobbet med dataetterforskning i politidistriktene. En interessant oppsummering finnes i Gjerdens konklusjon:

«On this foundation the study has found that the digital forensic investigators in the Norwegian police districts need comprehensive more training and education related to each phase of digital forensic investigation. Moreover, the study has found indications of digital forensic investigators who need leaders that show more understanding and thereby in practice meets their professional needs, such as lab facilities, forensic computers, equipment, tools and. This is probably related to limited budgets and finite resources, however, the findings indicate that there are insufficient communication routines between digital forensic investigators and the leaders concerning long-term planning related to their professional requirements. These mentioned findings might have caused contributions to the digital forensic investigators' requests concerning how they want the digital forensic field organized in the Norwegian Police. The patterns found indicate that the digital forensic investigators need a stronger and more central regulated and governed digital forensic field in the Norwegian Police.»

Siden den gang har mange studier undersøkt ulike aspekter omkring kvaliteten på dataetterforskning og digitale bevis i norsk straffesakssammenheng (Andreassen & Andresen, 2020; Bauge, 2023; Borhaug, 2019; Erlandsen, 2019; Friheim, 2016; Haraldseid, 2021; Heitmann, 2019; Jahren, 2020; Stoykova et al., 2022; N. Sunde, 2017). Jahren skriver følgende i sitt sammendrag:

«Studien viste at det var lite involvering og kvalitetsstyring fra ledelsen i fagområdet dataetterforskning i de deltagende distrikter, og at dette kan skyldes manglende digital kompetanse. Eventuelle kvalitetssikringstiltak i arbeidet måtte iverksettes av den enkelte dataetterforsker» (Jahren, 2020, s. 3).

Bauge skriver dette i sitt sammendrag:

«Studien avdekket at det i liten grad gjennomføres kvalitetskontroll av rapporter som dokumenterer arbeidet som blir utført, bortsett fra enkel grammatikksjekk og meningskontroll. Tid og bemanning ble identifisert som de største hindringene for implementeringen av gode systemer for kvalitetskontroll. Videre ble det klart at mange spesialistoppgaver i dag blir utført av personer som ikke er spesialister på dataetterforskning, uten at det finnes et system som sikrer kvaliteten på arbeidet som utføres. Funnene i studien gir grunn til bekymring når det gjelder forvaltningen av digital kriminalteknikk som fagområde, samt rettsikkerheten til de involverte partene» (Bauge, 2023 s. v).

Sitatet fra Gjerdes konklusjon sammenholdt med funnene i Jahres og Bauges studier indikerer at de systemiske utfordringene påpekt ovenfor ikke er nyoppståtte - og at de heller har vedvart over lengre tid uten at store endringer har skjedd.

I artikkelen *A review of quality procedures in the UK forensic sciences: What can the field of digital forensics learn?* (Page et al., 2019) ble kvalitetssystemene for digitale bevis sammenholdt med analyse av

fingeravtrykk, kroppsvæsker og DNA. Artikkelforfatterne konkluderte med at dataetterforskning ligger langt etter disiplinene det ble sammenlignet med, og at de eksisterende kvalitetssystemene i all hovedsak er rettet mot laboratoriemiljøer, og dermed i liten grad favner produksjonskjeden av digitale bevis internt i politiet. Verktøy eller systemer for å vurdere kvaliteten i arbeidet til den enkelte etterforsker, eller det digitale beviset som resultat av slikt arbeid, manglet. Dette ledet til to artikler i det vitenskapelige tidsskriftet *Forensic Science International: Digital Investigation* (Horsman & N. Sunde, 2020; N. Sunde & Horsman, 2021). I artikkelserien blir behovet for fagfellevurdering og kvalitetskontroll for digitale bevis ble diskutert, og 'The Phase-Oriented Advice and Review Structure' for kvalitetsarbeid i datatekniske undersøkelser blir presentert. Dette er en seks-steps modell for kvalitetssikring og kvalitetskontroll i datateknisk etterforskning, hvor Kvalitetskontrollhierarkiet, som er utgangspunktet for dette forsknings- og utviklingsprosjektet, inngår.

Kvalitetskontrollhierarkiet er et fleksibelt rammeverk for fagfellevurdering og kvalitetskontroll av datatekniske undersøkelser. Her deles fagfellevurdering inn i sju ulike nivå, fra det som kun omfatter en administrativ sjekk av at rapportene formelt er i orden, underskrevet og at alle vedlegg er med, via en omfattende vurdering av rapportens innhold både hva gjelder dokumentasjon av prosess og resultat, til verifikasjon av funnene og en ny og uavhengig analyse av beslaget (se fig. 1). Arbeidsprosessen tilhørende hvert nivå er beskrevet, og sjekklister følger for hvert nivå (N. Sunde & Horsman, 2021). I dette rammeverket inngår tre viktige roller fra enheten for digitalt politiarbeid (DPA-enheten): Rapport skriver, fagfelle og DPA-leder. Rapport skriver gjennomfører datateknisk undersøkelse og skriver rapport. DPA-medarbeidere utpekt til rollen som fagfelle leser og vurderer rapporten, og gir tilbakemelding til rapport skriver. DPA-leder koordinerer, velger ut saker for fagfellevurdering og peker ut fagfelle.

DPA-enheten ved Agder politidistrikt har i de senere år testet ut et system for sjekklisterbasert kvalitetskontroll av datatekniske rapporter i eget distrikt. De har benyttet egenutviklede sjekklister, som inneholder mange av de av de samme elementene som sjekklisterne i Kvalitetskontrollhierarkiet. Agder presenterte sine erfaringer ved den nasjonale konferansen rettet mot dataetterforskere, Ctrl Alt, i 2021, og beskrev at de hadde god erfaring med dette arbeidet og at det førte til avdekking av feil. De framhevet at de gjerne skulle ønske at en sjekklisterbasert kvalitetskontroll ble tatt i bruk av flere, slik at kvalitetskontroll kunne gjennomføres på tvers av politidistriktene.

Med inspirasjon fra Agders erfaringer med sjekklisterbasert kvalitetskontroll ble Kvalitetskontrollprosjektet startet høsten 2022. Målet med prosjektet var todelt:

- Å øke kunnskapen om hvorvidt og hvordan kvalitetssikring og kvalitetskontroll av datatekniske undersøkelser og rapporter i politidistrikter og særorgan ble gjennomført
- Å skape et kunnskapsgrunnlag gjennom systematisk testing av om fagfellevurdering og kvalitetskontroll, ved hjelp av verktøy fra Kvalitetskontrollhierarkiet, var egnet til å øke kvaliteten på datatekniske rapporter og bidra til læring.

I del 2 vil vi presentere og diskutere forskningsdesign og metoder brukt i Kvalitetskontrollprosjektet. Analyse av resultatene presenteres i del 3, og resultatene diskuteres i del 4.

2. Metode

Kvalitetskontrollprosjektet ble startet opp med formål om å teste ut om fagfellevurdering og tilbakemelding ved bruk av sjekklister fra Kvalitetskontrollhierarkiet var egnet til å øke kvaliteten på datatekniske rapporter og til læring for de involverte.

Følgende forskningsspørsmål dannet utgangspunkt for Kvalitetskontrollprosjektet:

Kan (og i tilfelle hvordan) fagfellevurdering og kvalitetskontroll i henhold til Kvalitetskontrollhierarkiets nivå 4 – Innholdskontroll

- bidra til å heve kvaliteten på datatekniske rapporter gjennom å avdekke kritiske feil og/eller villedende informasjon?
- skape læring hos fagfelle og/eller rapportskriver (mottaker av fagfellevurdering)?
- lede til systemforbedring/organisasjonslæring?
- gi uønskede sideeffekter?

Kvalitetskontrollprosjektet ble gjennomført fra september 2022 til oktober 2023, og kan deles i tre faser: Innføringsfasen (2.1), Pilotfasen (2.2) og Innsamlings- og analysefasen (2.3), som presenteres nedenfor. I del 2.4 gjøres det rede for begreper som brukes i denne rapporten.

2.1 Innføringsfasen

Målet med innføringsfasen var at DPA-medarbeidere og DPA-ledere skulle bli kjent med Kvalitetskontrollhierarkiet som konsept for fagfellevurdering og kvalitetskontroll av datatekniske rapporter.

2.1.1 Oversettelse og tilpasning av sjekklister

Høsten 2023 oversatte Nina Sunde sjekklister fra artikkelen *Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations* (N. Sunde & Horsman, 2021) til norsk, og justerte disse for norsk straffesaks kontekst (vedlegg 1). Det ble videre utarbeidet et støttedokument som forklarte formålet med hvert sjekkpunkt og ga eksempler på hvordan informasjonen kunne presenteres i rapport (vedlegg 2). For å tilgjengeliggjøre kunnskapen om Kvalitetskontrollhierarkiet ble det skrevet et sammendrag av de to vitenskapelige artiklene (Horsman & N. Sunde, 2020; N. Sunde & Horsman, 2021) på norsk (vedlegg 3).

2.1.2 Kartlegging

Rune Kenneth Bauge, som til daglig arbeider som spesialletterforsker i Vest PD, gjennomførte en kombinert kvalitativ og kvantitativ undersøkelse blant DPA-ledere og ansatte som et ledd i sin masteravhandling ved NTNU. Formålet var å kartlegge i hvilket omfang og på hvilket nivå kvalitetskontroll ble gjennomført før Kvalitetskontrollprosjektets pilotfase. I tillegg undersøkte han holdninger til kvalitet og kvalitetskontroll, samt opplevd risiko for feil blant DPA-ledere. Resultatene er presentert i Bauges masteravhandling (Bauge, 2023).

Figur 1: Kvalitetskontrollhierarkiet (eng: Peer Review Hierarchy for Digital Forensics, først presentert i Horsman & N. Sunde, 2020)



2.1.3 Workshop I og II - opplæring i Kvalitetskontrollhierarkiet og tilpasning av sjekklister

Høsten 2022 inviterte prosjektgruppen politidistrikter og særorgan til en serie på to workshops. Det ble bedt om at DPA-leder og DPA-medarbeidere som kunne inneha rollen som fagfeller i framtiden møtte. Formålet var å presentere Kvalitetskontrollhierarkiet som rammeverket for kvalitetskontroll, samt verktøyet i form av sjekklister og støttedokument for gjennomføring av kvalitetskontroll.

Workshop I ble arrangert på Teams 27. September, med ca. 35 deltakere fra landets politidistrikter, samt PST og Kripos. Her presenterte Nina Sunde Kvalitetskontrollhierarkiet, sjekklister og støttedokument. Andreas Røed fra Agder PD presenterte erfaringer

med kvalitetskontroll og sjekklister fra eget distrikt. Olav Dahl holdt et foredrag om tilbakemelding og læring.

Deltakerne ble gitt konkrete oppgaver fram til neste workshop. DPA-medarbeidere som var tiltenkt en framtidig rolle som fagfeller ble gitt følgende oppgave fram til Workshop II:

- Bli kjent med grunnlagsdokumentene (oppsummering av artikkel 1 og 2, sjekklister og støttedokument til sjekklistene)
- Velg ut 3-4 rapporter, og gjennomfør fagfelleevaluering på nivå 4 - Innholdskontroll (som også omfatter nivå 1, 2 og 3)
- Noter ned:
 - Hvordan fungerer sjekklstens punkter?
 - Er det sjekkpunkter som ikke bør være der?
 - Er det sjekkpunkter som bør inn?
 - Gir støttedokumentet tilstrekkelig veiledning? Hvor er det behov for oppklaring/utdyping/eksempler?

DPA-lederne ble bedt om å tenke gjennom følgende punkter i tiden fram til Workshop II:

- Hvordan vil du velge ut rapporter til kvalitetskontroll?
- Hvordan kan du bidra til at fagfellen kan gjennomføre blind fagfelleevaluering, ved å
 - Anonymisere hvem rapportskriver er før oversending til fagfelle?
 - Maskere konklusjon, slik at fagfelle først kan formulere denne på bakgrunn av funnene i rapporten?
 - Samarbeide om kvalitetskontroll med andre distrikter for å hindre 'statuseffekt'?
- Hvordan tenker du at fagfelleevalueringprosessen skal dokumenteres?
- Når systematiske feil blir avdekket – hvordan skal dette håndteres i din enhet?

- Hvordan ser du for deg din rolle i løsningen av konflikt dersom rapportskriver og fagfelle ikke kommer til enighet?
- Hvordan kan kvalitetskontroll bli en kunnskapskilde for å øke kvaliteten i framtidige dataetterforskninger?

Workshop II ble arrangert av prosjektgruppen den 25. -26. Oktober, med ca. 25 deltakere.

Første dag var satt av til diskusjonsgrupper med henholdsvis DPA-medarbeidere og DPA-ledere, hvor de ble bedt om å diskutere punktene i oppgavene som var gitt mellom Workshop I og II. Første dag ble rundet av med en diskusjon i plenum om erfaringer, samt innspill til endringer og justeringer av sjekklisten og støttedokument.

Andre dag startet med en oppsummering av første dag. Så presenterte IT-etterforsker Lars Wenstrup fra NC3/Danmark erfaringer med systematisk kvalitetskontroll. En sak som omhandlet brannstiftelse fra 2014 hadde vært en direkte opptakt til økt kvalitetssikringsfokus ved enheten. På mistenktes datamaskin hadde man funnet dokumenter som handlet om oppføring av et nytt hus, og dokumentene var datert før brannen fant sted. Ved en tilfeldighet ble det oppdaget at dataetterforsker hadde misforstått det aktuelle filsystemets datering av filer, og at riktig opprettsdato faktisk var etter tidspunktet for brannen. Dette førte til at de måtte gjennomgå 500 saker for å se om lignende feil hadde blitt gjort. Wenstrup hadde på tidspunktet for presentasjonen vært medlem av en intern prosjektgruppe for kvalitetssikring i 4 år. Gruppen har jobbet med å definere ansvar og roller, beskrivelse av standardmetoder, innføring av dokumenthierarki, samt nye rutiner for datasikring og håndtering av mobiltelefoner.

Kvalitetssikringsrollen, med kvalifikasjonskrav og oppgavebeskrivelse var nå implementert ved NC3/Danmark. Erfaringene etter innføring av systematisk kvalitetskontroll viste at de nå skrev færre, men bedre rapporter, og de hadde et tettere samarbeid med etterforskerne. Kvalitetssikringen av rapporter hadde ført til jevnlig avdekking

av graverende feil, og at det var uoverensstemmelse mellom de forskjellige undersøkelsesmetodene som ble brukt. De ansatte hadde dermed gått fra en generell skepsis til systematisk kvalitetssikring, til å verdsette denne ved at feil de selv hadde begått ble oppdaget i tide. De hadde også etablert et teknisk kvalitetsteam med ansvar for å sjekke og verifisere funn i tekniske analyser (dette tilsvarer nivå 5-6 i Kvalitetskontrollhierarkiet).

Wenstrup anbefalte på bakgrunn av erfaringene ved NC3/Danmark:

- Å gjennomføre kvalitetskontroll for å avdekke feil og mangler ved datatekniske undersøkelser
- Å omtale kvalitetskontroll som et samarbeid med mål om å levere det beste produktet
- Å kombinere både skriftlig og muntlig tilbakemelding
- Å dokumentere kvalitetssikringen, slik at det ikke oppstår tvil på et senere tidspunkt om kvalitetskontroll har blitt gjennomført.
- Å være bevisst på at fagfellen ikke alltid har rett, og at det fagfellen oppfatter som feil kan være en misforståelse av rapporten. Det er derfor viktige å stille spørsmål framfor å kun gi anvisninger om rettelser. Om man som fagfelle ikke er sikker - søk kunnskap hos noen som kan mer.
- Holde regelmessige møter mellom kvalitetssikringspersonalet for å sikre en ensartet praksis.

Deretter presenterte tidligere Forensic Science Regulator i England og Wales, Gillian Tully, erfaringer med innføring av systematisk kvalitetskontroll og krav til akkreditering for eksterne interne laboratorier som leverte kriminaltekniske undersøkelser (inkl. digital kriminalteknikk) til politiet. Tully framhevet at en stor del av arbeidet på veien mot akkreditering etter standarden ISO 17025³ besto i å kartlegge, beskrive og validere metoder og delprosesser i undersøkelsesprosessen. I dataetterforskning, som i stor grad er en verktøyavhengig disiplin, var det også essensielt å ha et robust system

³ ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories

for validering av verktøy, slik at f.eks. nye versjoner av verktøy ble gjennomgått og testet før de ble tatt i bruk.

Workshop II ble avrundet med informasjon fra prosjektgruppen om den planlagte pilotfasen av Kvalitetskontrollprosjektet. Deltakerne ble informert om at det ville bli sendt ut en invitasjon til deltakelse i pilotfasen som ville starte i januar 2023 og avsluttes i mai/juni 2023, og at det ville foreligge en justert versjon av sjekklister og støttedokument innen oppstart, basert innspillene fra Workshop I og II.

2.2 Pilotfasen

Målet med pilotfasen var å teste ut hvordan fagfellevurdering og kvalitetskontroll av datatekniske rapporter etter Kvalitetskontrollhierarkiets prinsipper og verktøy fungerer i praksis.

Invitasjoner ble sendt ut til politidistrikter og særorgan den 07.11.23. Prosjektet ble meldt til NSD/SIKT, og vurdering ble mottatt 10.12.22 (ref. 324498). Troms PD, Nordland PD, Vest PD, Øst PD, Oslo PD, Trøndelag PD og Kripos meldte at de ønsket å delta i pilotfasen.

Politidistriktene/særorgan ble invitert til å gjennomføre kvalitetskontroll i tråd med Kvalitetskontrollhierarkiets *nivå 4, Innholdskontroll* av datatekniske rapporter i perioden januar til mai 2023. Det ble presisert at datatekniske rapporter her skulle forstås som rapporter som beskriver resultatet av tekniske analyse og evt. innholdsanalyse utført av datateknisk etterforsker. Det ble ikke gitt noen føringer på antall rapporter som skulle gjennomgå fagfellevurdering i perioden. Følgende oppgaver ble skissert:

DPA-ledere:

- velger ut rapporter til kvalitetskontroll
- tilrettelegger for at kvalitetskontrollen av rapporten i størst mulig grad kan gjennomføres 'blindt' (dette gjelder ikke den

etterfølgende tilbakemeldingen til rapportskriver, som gjerne skjer ansikt til ansikt)

- bidrar til å løse uenighet mellom fagfelle og rapportskriver dersom dette skulle oppstå
- noterer ned sine erfaringer og refleksjoner i forbindelse med hver fagfellevurdering som blir organisert
- samler inn sjekklister, anonymiserer og sender en kopi til Nina Sunde v/Politihøgskolen. Sjekklister skal ikke inneholde saksopplysninger/taushetspliktig informasjon eller personopplysninger. Det anføres kun informasjon om sakstype (f.eks. 'drap', 'seksuallovbrudd')

DPA-medarbeidere utpekt til rollen som fagfelle:

- gjennomfører kvalitetskontroll i henhold til sjekklisten for kvalitetskontroll og tilhørende støttedokument
- gir tilbakemelding til rapportskriver i henhold til sjekklisten
- noterer ned sine erfaringer og refleksjoner i forbindelse med hver gjennomføring
- leverer sjekklister til leder etter at tilbakemeldingen er gjennomført

2.2.1 Oppstartsmøter

Det ble holdt oppstartsmøter i januar, hvor DPA-leder og DPA-medarbeidere som var gitt rollen som fagfeller deltok. Møtet med Trøndelag PD ble holdt i februar på grunn av senere påmelding til prosjektet.

I oppstartsmøtene informerte prosjektgruppen om rammer for prosjektet, de siste justeringene av sjekklister og støttedokument, samt leders og fagfelles oppgaver i pilotfasen. Deltakerne ble informert om at datainnsamlingen ville foregå gjennom intervjuer, samt gjennomgang av anonymiserte sjekklister. Fagfeller og DPA-ledere ble oppfordret til å bruke en refleksjonslogg for hver gjennomføring av

fagfelleevaluering. Formålet med dette var å understøtte egen læring i perioden, samt å huske bedre når de ble intervjuet mot slutten av pilotfasen.

Deltakerne fikk mulighet til å stille spørsmål, og det ledet til noen viktige avklaringer. Det ble stilt spørsmål om det var et minimumskrav til hvor mange saker som skulle fagfellevurderes i pilotfasen, om gamle saker kunne brukes, om det var bestemte kriterier for utvelgelse av saker som skulle brukes, hvordan tilbakemeldinger til rapportskriver skulle gis, og om rapportskriver burde spørres om de samtykket til å delta i prosjektet. Prosjektgruppen informerte om at det ikke var satt noen minimumskrav til antall eller omfang, og at dette må skje innenfor rammene av hva politidistriktet fikk til i sin daglige drift. Videre ble det informert om at rapporter i saker som ble etterforsket før pilotfasen i Kvalitetskontrollprosjektet kunne brukes, men at dette måtte avklares med etterforskningsledelsen på den enkelte sak. Prosjektgruppen understreket at det ikke ble satt noe krav til type eller størrelse på straffesak eller rapport, og at kvalitetskontrollen like godt kunne omfatte små som store og omfattende saker, og like godt korte som lange rapporter. Deltakerne ble orientert om at prosjektgruppen bevisst ikke hadde lagt noen føringer på hvordan tilbakemelding skulle gis, siden det var ønskelig å samle erfaringer fra hvordan dette ble løst av deltakerne selv. Deltakerne ble oppfordret til å bruke skjønn og legge vekt på hva de antok ville gi god kvalitet og læring. Deltakerne ble orientert om at dersom det ble aktuelt å intervju rapportskrivere, så ville de på samme måte som øvrige deltakere motta informasjonsskriv, og bli informert om frivillig deltakelse i intervju, og om hvordan personopplysninger ville bli håndtert. Prosjektgruppen oppfordret også DPA-ledere til å informere rapportskrivere om at rapporten var valgt ut for kvalitetskontroll, slik at ikke dette kom som en overraskelse når fagfelle tok kontakt for å gi tilbakemelding.

Etter oppstartsmøtet ble en dokumentpakke sendt til DPA-leder og DPA-medarbeidere som deltok på oppstartsmøtet. Denne inneholdt

sjekklister – versjon 1.2 (vedlegg 1), støttedokument - versjon 1.1 (vedlegg 2), sammendrag av de vitenskapelige artiklene hvor Kvalitetskontrollhierarkiet var beskrevet (vedlegg 3) mal for fagfelles refleksjonslogg (vedlegg 4), mal for leders refleksjonslogg (vedlegg 5) og kopi av presentasjonen som ble vist under oppstartsmøtet.

2.2.2 Antall fagfellevurderinger og tilbakemeldinger i pilotfasen

I pilotfasen ble det gjennomført fagfellevurderinger i fem av de syv enhetene som hadde meldt seg på prosjektet. Totalt ble 28 rapporter vurdert. 5 hadde både vært fagfelle og fått tilbakemelding på sin rapport, 6 hadde kun vært fagfeller, og 1 hadde kun fått tilbakemelding. I tillegg ble en DPA-medarbeider intervjuet om erfaringer fra fagfellevurderinger som ble gjennomført i eget distrikt før Kvalitetskontrollprosjektet, og antallet vurderinger er ikke medregnet her. De fleste hadde kun vurdert eller fått tilbakemelding på 1-2 rapporter, mens noen få hadde vurdert 5 rapporter. I DPA-enhetene hvor det ikke var gjennomført fagfellevurderinger i pilotfasen ble kun DPA-leder intervjuet.

Prosjektgruppen mottok totalt 11 sjekklister. Antallet avviker fra antallet vurderte rapporter, siden sjekklister ble brukt på ulike måter, og ikke alltid 'fylt ut' (se nærmere beskrivelse i punkt 3.2.5).

DPA-medarbeider	1	2	3	4	5	6	7	8	9	10	11	12	13	Totalt
Fagfellevurdert	1	3	1	0	2	3	3	3	4	2	5	1	-	28
Fått tilbakemelding	1	0	1	1	3	0	0	0	4	0	0	1	-	11

2.2.3 Framdrift

I begynnelsen av mars ble det sendt mail til DPA-lederne med spørsmål om hvordan framdriften var og hvordan pilotfasen forløp. Noen få meldte tilbake at de var godt i gang, mens flere fortalte at de av ulike grunner var på etterskudd. Flere av DPA-enhetene hadde ikke fått gjennomført noen fagfellevurderinger ennå. I mai ble DPA-lederne på

nytt kontaktet for status. Da ba flere om også å få bruke mai og deler av juni for å komme i mål. To DPA-enheter fortalte at de ikke hadde fått gjennomført fagfelle vurderinger i perioden.

2.3 Innsamlings- og analysefasen

2.3.1 Datainnsamling

Datainnsamlingen ble gjennomført i løpet av mai til august 2023, og besto i intervjuer med deltakerne og innsamling av anonymiserte sjekklister.

Intervjuer

DPA-ledere (n=7) og DPA-medarbeidere (n=13) ble intervjuet på Teams. Det ble gjennomført til sammen 20 semistrukturerte intervjuer, av henholdsvis 7 DPA-ledere og 13 DPA-medarbeidere. DPA-medarbeiderne i prosjektet hadde blandet bakgrunn. 7 hadde politibakgrunn og 6 hadde sivil bakgrunn. Blant de med sivil bakgrunn hadde de fleste datateknisk utdanning, mens en var utdannet jurist. I forkant av intervjuet fikk deltakerne tilsendt et informasjonsskriv med samtykkeerklæring (vedlegg 6), som ble underskrevet og returnert til Prosjektgruppen. I noen tilfeller ble samtykket gitt muntlig ved oppstart av intervjuet. Nina Sunde gjennomførte samtlige intervjuer, og Olav Dahl deltok i 4 intervjuer (DPA-medarbeidere).

Intervjuene handlet om tre overordnede tema:

- Erfaringer fra kvalitetskontroll gjennomført i pilotfasen
- Egne erfaringer med eller kjennskap til feil og nesten-feil i saker med digitale spor fra egne distrikt/særorgan
- Holdninger til eller meninger om digitale bevis, kvalitet og rettssikkerhet

Det ble presisert overfor samtlige deltakere at i den grad de brukte konkrete straffesaker som utgangspunkt i intervjuet så måtte de ha særskilt oppmerksomhet på at de ikke gikk på tvers av sin taushetsplikt.

Sjekklistor

Anonymiserte sjekklistor (n=11) brukt i pilotfasen ble oversendt til prosjektgruppen fra DPA-lederne. Formålet med innsamling av sjekklistor var at disse skulle danne grunnlag for en kombinert kvalitativ og kvantitativ dokumentanalyse. Siden antallet sjekklistor ble nokså lavt, ble grunnlaget vurdert til å være utilstrekkelig for analysen.

2.3.2 Analyse

Intervjuene ble transkribert, og analysert med metoden Tematisk analyse (Braun & Clarke, 2006). Intervjuene av DPA-ledere og DPA-medarbeidere (fagfeller og rapportskrivere) ble analysert hver for seg. Resultatene ble så sammenstilt og presenteres i del 3. Sitater fra deltakerne presenteres i kursiv skrift, er markert med anførselstegn før og etter, og er merket med kode for deltaker ('M for DPA-medarbeider og 'L' for DPA-leder).

Etter sammenstilling og diskusjon av resultatene i rapporten, ble sitatene og noe av konteksten det var plassert i sendt til deltakerne. Formålet var å avklare om vi hadde misforstått sitatet eller konteksten det var relatert til. En deltaker ønsker ikke at sitatet skulle stå slik som det var oppført fordi deler av uttalelsen var for stor grad basert på det hen karakteriserte som sin 'synsing'. Sitatet ble tatt ut, og uttalelsen ble omskrevet til løpende tekst uten 'synsingen', siden dette heller ikke var det prosjektgruppen ønsket å framheve med bruken av sitatet.

2.3.3 Workshop III – formidling av resultatet fra pilotfasen

Workshop III ble arrangert 13. oktober 2023 med ca. 30 deltakere. Her presenterte Rune Kenneth Bauge resultatene fra sin masteravhandling, som omhandlet kartleggingen av kvalitetskontroll blant DPA-ledere og DPA-medarbeidere i forkant av Kvalitetskontrollprosjektets pilotfase. Deretter presenterte Nina Sunde og Olav Dahl resultatene fra datainnsamlingen etter pilotfasen. Etter innledende kommentarer fra Rune Rustad Rudi (Politidirektoratet) og Odin Heitmann (NC3, Kripos) ble det åpnet for spørsmål og diskusjon av resultatene.

2.4 Begrepsavklaring

Nedenfor finnes en utdypende forklaring på hvordan noen sentrale begreper skal forstås i denne rapportens del 1,2, 4 og 5. Unntaket er sitater fra deltakerne i rapportens del 3, som noen ganger bruker andre begreper eller forstår begrepene annerledes enn slik de forklares nedenfor.

Databærer: Datasystemer (inkludert skysystemer), brukerkonti og frakoblede lagringsmedier (I. M. Sunde, 2021, s. 64).

Dataetterforsker: DPA-medarbeidere (sivilt ansatte eller politiutdannede) med ansvar for sikring og analyse av digitale spor/bevis (også omtalt som digital kriminalteknikk) som sin hovedoppgave.

Dataetterforskningsprosessen: Prosessen som innebærer å

- Identifisere databærere og sporsteder hvor det antas å finne digitale spor av betydning for straffesaken,
- Sikre data fra databærere og sporsteder,
- Prosessere sikret data og tilrettelegge for analyse,
- Analysere data ved å utføre
 - Innholdsanalyse/gjennomgang for å finne relevante spor som kan opplyse straffesaken

- Teknisk analyse for å finne (flere) relevante spor som kan opplyse straffesaken og/eller verifisere relevans og pålitelighet av spor funnet i gjennomgang/innholdsanalyse
- Vurdering av sporenes verdi (i lys av hypoteser)
- Dokumentasjon og presentasjon av digitale spor/bevis
- Kvalitetskontroll – fagfellevurdering av rapport og evt. verifisering av spor.

Kvalitetssikring bør inngå i hver fase av prosessen (se beskrivelse av systematisk kvalitetssikring som en del av PARS-rammeverket (The Phase-Oriented Advice and Review Structure) som beskrevet i N. Sunde og Horsman (2021).

Digitale spor/bevis: Digitale spor og digitale bevis brukes om digitale objekter som er eller kan være relevante for å opplyse straffesaken som er under etterforskning.

DPA-leder: Ansatt ved DPA-enhet med lederfunksjon. I denne rapporten brukes benevnelsen om deltakere som har hatt et ansvar for å tilrettelegge for fagfellevurdering ved sin DPA-enhet i pilotfasen.

DPA-medarbeider: Medarbeider ved enhet for digitalt politiarbeid (DPA) med rolle som dataetterforsker. I denne rapporten brukes benevnelsen om deltakere som har hatt rollen som fagfelle.

Fagfelle: Medarbeider ved enhet for digitalt politiarbeid (DPA-medarbeider) som er utpekt til å vurdere og gi tilbakemelding på rapporter om digitale spor/bevis.

Fagfellevurdering: En vurdering av en rapport om digitale spor/bevis etter gitte kriterier og faglig skjønn, og en tilbakemelding til rapportskriver.

Fagkontakt: Politibetjent med utvidet rolle og ansvar for mindre avanserte oppgaver innenfor sikring av digitale spor – for eksempel sikring fra mobiltelefoner eller skytjenester.

Kvalitetssikring: Oppgaver og tiltak for å sikre tilstrekkelig grad av kvalitet og forebygge feil underveis i etterforskningen.

Kvalitetskontroll: Oppgaver og tiltak for å avdekke kvalitetsbrister og feil. I dette prosjektet har kvalitetskontrollen vært rettet mot rapporter fra datatekniske undersøkelser.

Kvalitetskontrollhierarkiet: Rammeverket som er utarbeidet med henblikk på fagfellevurdering og kvalitetskontroll av rapporter som omhandler resultatet analysefasen dataetterforskningsprosessen, altså rapporter om digitale spor/bevis. Analysefasen forstås her som fasen hvor gjennomgang/ransaking av datamateriale skjer med formål om å finne relevant informasjon som belyser straffesakens bevistema og øvrige informasjonsbehov, samt tekniske undersøkelser for å undersøke relevansen og påliteligheten (usikkerheten) av slik informasjon. Kvalitetskontrollhierarkiet ble først beskrevet i engelsk versjon (Horsman & N. Sunde, 2020) under navnet The Peer Review Hierarchy for Digital Forensics.

Rapportskriver: Forfatter av rapport om datatekniske undersøkelser og digitale spor/bevis – og som i denne sammenhengen har fått vurdering av sin rapport fra fagfelle.

Sjekkliste: Dokument med punkter som skal vurderes ved fagfellevurdering av rapport om digitale spor/bevis. I dette prosjektet er sjekklisten utformet i tråd med nivåene i Kvalitetskontrollhierarkiet.

Støttedokument: Dokument med utdypende forklaring på og begrunnelse for hva som skal vurderes under punktene i sjekklisten, og med illustrerende eksempler der det er relevant.

3. Resultat og analyse

Nedenfor presenteres resultatene fra intervjuer med deltakerne fra pilotfasen av Kvalitetskontrollprosjektet. Intervjuer med en leder og en ansatt som ikke deltok i pilotfasen, men som har erfaring med fagfellevurdering i sitt distrikt fra tidligere, er også tatt med. Analysen er delt inn i tema, og hvor både DPA-ledere og DPA-medarbeideres uttalelser inngår. Benevnelsen DPA-leder og DPA-medarbeider brukes når det har betydning for hvem som har gitt informasjonen. Flere tema sammenfaller, og når begge gruppene er omtalt, brukes benevnelsen 'deltakerne'. DPA-medarbeiderne har hatt ulike roller, som omtales som 'fagfelle' (de som har vurdert rapporter og gitt tilbakemelding) og 'rapportskriver' (de som har fått sin rapport vurdert og mottatt tilbakemelding). Når DPA-medarbeidere eller ledere omtaler de som er involvert i håndtering av digitale spor og bevis brukes benevnelsen 'dataetterforsker' om DPA-medarbeidere på spesialistnivå, 'fagkontakter' om de som har et utvidet ansvar for sikring av digitale spor som del av sin stilling, og 'etterforskere' om de som gjennomfører innholdsanalyse eller gjennomgang av beslaglagte data for å identifisere relevant informasjon til straffesaken.

3.1. DPA-medarbeideres, fagkontakters og etterforskeres rolle i dataetterforskningsprosessen

3.1.1 DPA-medarbeiderens rolle i endring

Ifølge deltakerne har dataetterforskers rolle gjennomgått en betydelig endring de siste 10-15 årene. Tidligere gjennomførte dataetterforskerne større deler av dataetterforskningsprosessen selv, altså sikring av data fra beslag, analyse og rapportering. I dag er dette mer oppstykket, og langt flere er involvert i prosessen. Vanligvis er det patruljen som tar beslag. Dataetterforsker sikrer og tilrettelegger

data fra datamaskiner og mer avanserte enheter, mens fagkontakter gjerne sikrer data fra mindre avanserte enheter og systemer, som mobiltelefoner og skytjenester. Flere er involvert i analysen av det sikrede datamaterialet, som kan inndeles i tre typer aktiviteter:

- *Innholdsanalyse/gjennomgang*: innebærer å lete i det sikrede og tilrettelagte databeslaget etter relevant informasjon for straffesaken. Dette gjennomføres vanligvis av etterforskere uten datateknisk formalkompetanse. Til dette formålet brukes gjerne verktøy som presenterer data på en oversiktlig måte, som for eksempel Cellebrite eller Magnet Axiom.
- *Verifisering*: innebærer å kvalitetssikre om funn i innholdsanalysen/gjennomgangen er korrekte. Dette gjøres ved å sjekke om resultatet kan verifiseres ved bruk av et annet verktøy, eller ved manuell undersøkelse av datastrukturer i beslaget. Verifisering handler også om å vurdere om funnets betydning er forstått/tolket riktig av etterforsker. Verifisering gjennomføres av dataetterforskere eller fagkontakter med tilstrekkelig teknologisk kompetanse.
- *Teknisk analyse*: innebærer å gjennomføre dypere undersøkelser av datastrukturer for å finne flere spor eller svare ut mer komplekse spørsmål omkring funn som er gjort i innholdsanalysen. Tekniske analyser initieres som regel som følge av konkrete spørsmål og avgrensede oppdrag, som «har bildet blitt slettet eller ikke?». Tekniske analyser krever mer avansert teknologisk kompetanse og gjennomføres av dataetterforskere.

Ifølge en DPA-leder lå det ved overgangen til denne arbeidsdelingen en mer eller mindre uttalt forventning om at dataetterforsker i større grad skulle verifisere og kvalitetssikre funn etter en slik gjennomgang og eventuelt bistå med dypere og mer avanserte tekniske undersøkelser hvis nødvendig. Det ble imidlertid færre slike oppdrag enn forventet. Ifølge DPA-medarbeiderne blir de noen ganger bedt om å kvalitetssikre

eller gjøre dypere tekniske undersøkelser, men langt fra alltid. Ifølge DPA-medarbeiderne bruker de nå en stor andel av tiden på sikring og tilrettelegging for innholdsanalyse eller 'gjennomgang', og langt mindre tid på verifisering og tekniske analyser.

DPA-medarbeiderne fortalte at de ofte ikke hører mer fra etterforsker etter at tilretteleggingen er utført, og det skrives vanligvis heller ingen rapport på tilretteleggingen. Det er altså etterforsker som selv dokumenterer resultatet av undersøkelsene. Det er ingen rutine for å skrive rapport om tilretteleggingen – noe en DPA-medarbeider framhevet som en svakhet siden det kan være avvik mellom verktøy som brukes for tilrettelegging og verktøy som senere brukes for en teknisk analyse. I mange tilfeller sikres, analyseres og dokumenteres beslag, som oftest mobiltelefoner, uten at DPA-medarbeider på noen måte er involvert i prosessen.

Men selv om de fleste som har deltatt i prosjektet har lang erfaring og god datateknisk kompetanse, så beskriver majoriteten av DPA-medarbeiderne at de bruker relativt lite av tiden sin på tekniske analyser. Utviklingen som beskrevet over har ført til at DPA-medarbeidere i stor grad utfører sikringsoppgaver, og at den videre gjennomgangen/analysen og rapporteringen av funn blir gjort av andre med mindre teknisk formalkompetanse. Dataetterforskere skriver derfor langt færre tekniske rapporter enn tidligere; faktisk så få at flere fortalte om vanskeligheter med å finne nok rapporter for vurdering i de 6 månedene pilotfasen varte. Noen få av DPA-medarbeiderne beskriver at de bruker mest tid på tekniske analyser. Dette er gjerne distrikt som har gjort organisatoriske grep, og etablert team som har dette som hovedoppgave.

Som følge den nye arbeidsdelingen i dataetterforskningsprosessen har også rapportene DPA-medarbeiderne skriver endret seg. DPA-medarbeiderne fortalte at de stort sett skriver sikringsrapporter, og svært få rapporter fra tekniske analyser. Videre fortalte de at analyserapportene også hadde endret seg. Fra å tidligere skrive en stor

rapport som omfatter en komplett innholdsanalyse og teknisk analyse, skrives nå heller flere kortere rapporter knyttet til ulike avgrensede undersøkelser av beslaget.

Det faktum at DPA-medarbeiderne gjør færre analyser og skriver færre rapporter har ifølge DPA-lederne imidlertid ikke resultert i at de har fått bedre tid til rådighet. Sakene har større volum, med flere databærere og større datamengder som skal sikres. Dataetterforsker bruker mye tid på administrative oppgaver relatert til sikringene, som å flytte data mellom systemer og å føre samme informasjon manuelt inn i flere systemer. Teknologitvillingen går raskt, noe som resulterer i at kunnskapen fra forrige sak ikke alltid kan gjenbrukes i neste sak. Dermed går det mye tid til å vedlikeholde og skaffe seg tilstrekkelig kunnskap for å være kunne sikre relevant informasjon og gjøre forsvarlige undersøkelser av digitale beslag.

I kraft av sin kompetanse er det høye forventninger til hva dataetterforsker skal bidra til i straffesaksetterforskningen ut over de nevnte oppgavene. Dataetterforsker skal møte i Etterforsknings-KO (koordinerer og leder arbeidet i større straffesaker) og andre interne saksmøter for å gi råd om digitale spor, de skal tilbakeføre filer fra beslag som er inndratt og gjøre uttrekk fra beslag som inneholder beslagsfritt materiale (f.eks. advokatkorrespondanse). En DPA-leder fortalte at det forventes at DPA-enheten skal ha datakrimspesialister, kryptospesialister, spesialister på sikring og analyse av ulike filsystemer, mobiltelefonteknologi og andre multimediaenheter og databærere, samt spesialister i sikring fra skytjenester og åpne og lukkede rom (darkweb) på internett.

Ifølge DPA-lederne har dataetterforskerne mange ulike oppgaver som spiser av tiden til straffesaksarbeid. De drifter fagkontaktordningen i distriktet slik at de har tilstrekkelig kompetanse og utstyr til sikring av digitale beslag, og de driver opplæring i verktøy som fagkontakter og etterforskere i distriktet benytter. De bruker mye tid på tiltaket Obligatorisk årlig opplæring (OÅO), hvor de skal besøke alle

driftsenheter for å delta i paroler og tilbakemeldingsmøter. I tillegg drifter de fleste sine egne beslagsservere og systemer for datatekniske undersøkelser. De monitorerer lisensbruken i distriktet, drifter fagkontakters utstyr, håndterer beslagsrom og vedlikeholder fysisk utstyr. Disse oppgavene faller på noen ganske få medarbeidere. Det er store ulikheter med hensyn til antall dataetterforskere per distrikt, men enkelte distrikter har over tid hatt så lite som 2-3 dataetterforskere som skal ivareta alle oppgavene som beskrives ovenfor.

Samtlige av de intervjuede DPA-lederne var bekymret for denne situasjonen. Flere fortalte om rekruttering av sivil spesialistkompetanse som man hadde ønsket å bruke til de mest teknisk krevende oppgavene. På grunn av den pressede ressursituasjonen ble de nå satt til sikring/tilrettelegging og administrative oppgaver som heller kunne vært utført av noen med mye lavere kompetanse. Flere var bekymret for hvor lenge man ville kunne holde på denne spesialistkompetansen dersom medarbeiderne ikke fikk tid til å fordype seg i faget. Kompetanse som ikke vedlikeholdes forvitrer også som kjent med tiden. Flere etterlyste derfor rolleavklaring, slik at det blir tydeligere hva rollen til dataetterforskere som skal utføre tekniske analyser og 'digital kriminalteknikk' skal inneholde, hva som skal overlates til fagkontakter og hva etterforskere kan utføre.

Ett distrikt hadde tatt noen radikale grep for å bedre denne situasjonen. De hadde ansatt to i Politibetjent 1-3 stillinger til å gjøre sikring og tilrettelegging på fulltid. De hadde fått grundig opplæring, og var samlokalisert med dataetterforskerne, slik at de lett kunne få bistand hvis nødvendig. Dataetterforskerne sikret nå kun det mest teknisk avanserte beslagene, og hadde fått frigjort mer tid til tekniske analyser. DPA-lederen fortalte at de var svært fornøyde med løsningen, og ønsket å videreføre denne.

3.1.2 Digitalt politiarbeid og cyberkriminalitet

Flere DPA-ledere pekte på at begrepet 'digitalt politiarbeid', som ble innført ved nærpolitireformen, kunne være en medvirkende årsak til at dataetterforskers rolle med tiden hadde blitt mer fragmentert og uklar, og at ressursene ikke strakk til. Dette ble begrunnet i at det på papiret kunne framstå som DPA-enheten i distriktet er godt rustet til oppgaven, mens det i praksis var noen ganske få ansatte ved DPA-enheten som hadde kompetanse til å gjennomføre kjerneoppgaver som digital kriminalteknikk og dypere tekniske undersøkelser. Begrepet digitalt politiarbeid omfatter også nettpatruljer og internettetterforskning, og det ble påpekt at det ikke finnes en kompetanseprofil som dekker alt som faller inn under et slikt vidt og uklart begrep. DPA-lederne understreket behovet for tydeligere rolleavklaringer og kompetansekrav til de ulike funksjonene.

Flere DPA-ledere ytret bekymring for at etterforskning av cyberkriminalitet også ble lagt til enheter for digitalt politiarbeid, og at det ikke var dedikerte medarbeidere som arbeidet med dette området. Per i dag tilbyr ikke Politihøgskolen utdanning på dette området, og det ble understreket at en utdanning ville vært til stor hjelp for å få på plass dedikert kompetanse i distriktene til å etterforske denne typen kriminalitet.

3.1.3 Etterforskernes og fagkontaktens rolle

De fleste distriktene har etablert fagkontakter, som har et ekstra ansvar for sikring av mobiltelefoner og tilrettelegging for gjennomgang av innholdet. Disse har enten NCFI Core utdanning fra PHS eller blir gitt lokal opplæring i sikring av mobiltelefoner, og blir tildelt utstyr for å gjennomføre slikt arbeid. De fleste gjør som regel dette som en del av stillingen, i kombinasjon med for eksempel etterforskning eller patruljetjeneste. Mange av deltakerne uttrykte misnøye med etableringen av denne rollen, og pekte på flere utfordringer.

Siden fagkontaktene ikke er administrativt underlagt DPA-enheten, er det linjeleder som bestemmer og prioriterer hvor mye tid fagkontakten får brukt på kompetanseheving og datateknisk arbeid. Ifølge deltakerne utgjøre fagkontaktrollen gjerne kun en andel av stillingen, noe som medfører at de ofte ikke får nok tid til å utføre mobilsikringer. Dermed blir det også vanskelig å opprettholde et tilstrekkelig kompetansenivå for å utføre sikringsoppgavene på en effektiv og forsvarlig måte. Utstyret og lisenser som fagkontaktene behøver er dyrt. Ett distrikt hadde derfor 'reversert' løsningen og trukket tilbake utstyret fra fagkontaktene.

En DPA-medarbeider fortalte:

«Og det er klart at, hadde jeg fått valget mellom tre og en halv million i lisenser eller ansatt to dataetterforskere så hadde jeg helt klart valgt to dataetterforskere og så hadde jeg opprettet et postbud, som da kunne tatt de to tre fire gangene i året, og kjørt med lynets hastighet». (M13)

Det ble framhevet fra flere DPA-ledere at fagkontaktordningen burde evalueres før den blir permanent innført som en funksjon hvert politidistrikt skal inneha.

Som nevnt er det i all hovedsak etterforsker som gjennomgår de digitale beslagene. Dataetterforskerne blir først koblet inn hvis etterforsker finner noe man er usikker på betydningen av, eller hvor etterforsker finner at det trengs en dypere teknisk analyse. Denne utviklingen skjedde trolig som følge av flere forhold. Flere DPA-ledere peker på stort arbeidspress som årsak til denne utviklingen, og at DPA ikke lenger hadde ressurser til å håndtere oppgavene som vokste i omfang. Noen deltakere henviste til en annen faktor, nemlig at etterforsker, med sin førstehåndskunnskap til saken, hadde bedre forutsetninger for å finne relevant informasjon i den aktuelle saken, og at DPA-medarbeiderne ikke hadde de samme forutsetningene.

Deltakerne var jevnt over bekymret for kvaliteten på innholdsanalysene eller 'gjennomgangene', som i hovedsak ble utført av etterforskerne. Grunnen til bekymring var at etterforskerne ofte ikke forsto sin egen begrensning og manglende tekniske kunnskap. Flere DPA-medarbeidere opplevde at etterforskerne stolte på verktøyene i for stor grad, og at de kunne komme med beskrivelser og konklusjoner i rapportene sine som egentlig krevde en helt annen teknisk kompetanse og langt dypere tekniske undersøkelser. Flere deltakere uttrykte bekymring for at det ikke var etablert et system som fanger opp gjennomgangsrapporter som burde vært kvalitetssikret før de ble lagt inn i straffesaken. Vurderingen om hvorvidt kvalitetssikring er nødvendig gjøres av etterforsker selv, som ofte ikke har kunnskap om potensielle fallgruver og usikkerhetsmomenter ved undersøkelsen. Per i dag er det altså ingen regel om at dataetterforsker skal kvalitetssikre undersøkelser eller rapporter fra fagkontakter eller etterforskere, og DPA-medarbeiderne så heller ikke at de hadde tid og ressurs til å kvalitetssjekke alt som ble produsert. En DPA-medarbeider foreslo at en mulig framtidig arbeidsdeling kunne være at fagkontakter fikk ansvar for å kvalitetssjekke etterforskeres arbeid, og at DPA kvalitetssjekk fagkontaktens rapporter.

DPA-lederne så med stor bekymring på hvordan dette nå hadde utartet. Flere understreket at den framtidige justisskandalen ville komme som følge av etterforskeres gjennomgang og manglende kvalitetssikring av dette arbeidet. En DPA-leder uttalte:

«Men det går jo på kompetanse primært utenfor oss hvor de største feilene skjer da, og det sier seg egentlig kanskje selv da kanskje. At det er der, den neste rettskandalen ligger nok i leddet etter oss. Og det ser vi kanskje allerede med den der Baneheia saken, at det er jo tolkning og vurdering av digitale spor i etterkant av at det er hentet inn som er det store problemet». (L6)

I stedet for å be om at funn blir verifisert og kvalitetssikret, blir etterforskeres rapporter lagt rett i straffesaken. En DPA-leder fortalte

at på tross av at mange burde vite bedre, så ble digitale spor fortsatt gjerne betraktet som objektive og 'sanne', og at det derfor ble stilt få spørsmål ved disse videre i etterforskningskjeden. Dette har medført situasjoner som at påtalejuristen, få dager før hovedforhandling, tok kontakt med DPA-enheten og ba om verifisering av spor. Dette er svært uheldig, siden det da ofte ikke var tid til å gjøre gode nok undersøkelser. I tillegg ble det ordinære saksarbeidet forsinket, siden det måtte legges bort inntil hasteoppdraget var løst.

Ifølge deltakerne var det tiltak på gang i flere distrikt med formål om å stramme til praksisen som innebærer at 'alle kan gjøre alt' med digitale beslag. Flere distrikt setter nå krav til at etterforskere som skal gjennomgå beslag skal ha kurs i verktøyet de skal bruke.

Noen DPA-ledere fortalte at de så en tendens til bedring, spesielt etter at ledere og medarbeidere fra DPA-enheten selv har vært aktive på paroler og fagdager og snakket om behovet for kvalitetssikring av funn. For noen DPA-enheter hadde det ført til en økt pågang fra etterforskningsledere og etterforskere for å få verifisert funn fra gjennomganger av digitale beslag. Selv om dette var en positiv utvikling, ble det understreket at de få og pressede dataetterforskerne som finnes i dag ikke har kapasitet til å kvalitetssikre alle gjennomganger, og at fagkontacters rolle i dette arbeidet må tas med i betraktning, slik at arbeidsbelastningen kan fordeles.

For å oppsummere DPA-medarbeideres, fagkontacters og etterforskeres rolle i dataetterforskningsprosessen, viser analysen at dataetterforskningsprosessen er mer fragmentert og involverer flere roller enn tidligere. Deltakerne uttrykker bekymring over denne utviklingen i og med at medarbeidere - uten tilstrekkelig datateknisk kompetanse - nå i større grad enn tidligere utfører analyseoppgaver, uten at disse kvalitetssikres.

3.2. Erfaringer fra Kvalitetskontrollprosjektets pilotfase

3.2.1 Tidligere erfaring med kvalitetskontroll

Ifølge deltakerne var det i flere av distriktene blitt gjort mer eller mindre vellykkede forsøk på å etablere en rutine for kvalitetskontroll av datatekniske rapporter før oppstarten av Kvalitetskontrollprosjektet.

«Ja, altså det er jo en sånn, hva skal vi kalle det, det er ikke noe sånt sjekkpunkt i verken vårt interne journalsystem eller i BL eller i noe som helst, men det er en uskreven regel da på seksjonen at rapporter skal leses. Så vi, på en måte, formidler jo det også ut til andre personer vi jobber med som er utenfor, på en måte, DPA-miljøet og da. Og da er det ofte innholdsanalyser. Vi sier vel i hvert fall at vi anbefaler sterkt at en kollega som har litt kunnskap om det du har sett på ser gjennom rapporten din da». (M3)

Flere hadde blitt inspirert av Agder politidistrikts lokale prosjekt for kvalitetskontroll som ble presentert på Ctrl Alt konferansen i 2021, og hadde fått sjekklister derifra som de hadde forsøkt å ta i bruk. Dette var stort sett sporadiske og 'ildsjel-drevne' initiativ. Mange av deltakerne hadde dermed erfaring med å lese og vurdere andres rapporter. Fokuset i disse kontrollene hadde i stor grad vært språk/grammatikk, klarhet og tilgjengelighet, som tilsvarer Kvalitetskontrollhierarkiets nivå 2 og 3. Mange hadde fått tilbakemeldinger på rapporter fra kollegaer på egen enhet, mens enkelte også hadde fått tilbakemeldinger fra etterforskningsledere i saker de har vært involvert i. Ifølge DPA-ledere hadde de ansattes holdninger til tidligere kvalitetskontrollinitiativ variert fra å ønske systematisk kvalitetskontroll av rapporter velkommen til boikott av tiltaket begrunnet i at det var en form for sensur. Enkelte distrikter hadde en form for kvalitetskontroll på tidspunktet hvor Kvalitetskontrollprosjektet ble introdusert, men ingen brukte sjekklister på en systematisk måte.

3.2.2 DPA-ledernes rolle i pilotfasen

Lederne i DPA-enhetene som gjennomførte kvalitetskontroller i pilotfasen hadde stort sett løst oppdraget nokså likt. De la vekt på å sette rammer, som å bestemme sakstyper eller type analyse, fasilitere og være interessert i framdriften. En leder hadde selv gitt tilbakemelding på to rapporter etter at fagfelle hadde vurdert rapport og fylt ut sjekklisten, med formål om å lære mer og forstå hva dette går i. Under workshop I og II, som ble gjennomført før pilotfasen, uttrykte flere DPA-ledere bekymring for at det ville være vanskelig å finne tid til kvalitetskontroll i en presset hverdag. Ved intervjuene etter pilotfasen uttalte imidlertid flere DPA-ledere at det hadde gått greit siden de selv hadde kunnet styre omfanget.

«Vi så at dette her gikk egentlig så grei at de tok det seg imellom. Så min rolle ble egentlig bare oppstarten på dette, og egentlig fasilitere og komme i gang, og så ble vi enig etter hvert at, vet du hva, vi kjenner hverandre så godt, og vi vet hvem som har skrevet rapportene når vi begynner å lese innholdet her. Og alle var interessert i å få en tilbakemelding på rapportene sine, sånn at min rolle ble egentlig ganske ubetydelig i denne sammenhengen, fordi dette her fikset de selv». (L5)

For de to distriktene som hadde takket ja til deltakelse i Kvalitetskontrollprosjektets pilotfase, men som ikke gjennomførte fagfelle-vurderinger i perioden, ble omorganisering og sykdom oppgitt som årsaker til dette. DPA-lederne understreket at de helt klart så viktigheten av kvalitetskontroll. En DPA-leder fortalte:

«Nei vi har ikke fått brukt det siden vi startet piloten, dessverre. Men jeg tenker jo at sjekklisten er kjempebra, og vi har fått være med på å fikse og ordne litt på de fram mot oppstarten, så vi synes de virker veldig bra da, og jeg tenker at vi ønsker å bruke de framover. Det har vi en ambisjon om». (L1)

3.2.3 Sjekkliste og støttedokument

Hvordan ble sjekklisten brukt?

Sjekklisten består av en veiledning for fargekoding omtalt som ('trafikklys'), og punkter fordelt i henhold til Kvalitetskontrollhierarkiets nivåer. For hvert punkt i sjekklisten er det anledning til å sette kryss i en rute med fargene rød, gul, grønn eller grå. Hvert punkt har også et kommentarfelt, hvor avkrysningen kan forklares eller begrunnes. Til slutt i sjekklisten finnes et punkt for oppsummering, hvor rapportens styrker, utfordringer og forbedringspunkter eller andre kommentarer kan beskrives i fritekst, samt et punkt om Synergi (politiets forbedringssystem), med beskrivelse av forhold som skal meldes inn i dette systemet og hvem som har ansvar for å gjøre det.

Støttedokumentet inneholder en utdypende forklaring og begrunnelse for hvert punkt i sjekklisten, og illustrerende eksempler der det er relevant.

DPA-medarbeidere ved fire av de fem enhetene som deltok hadde brukt sjekklisten aktivt, altså ved å krysse av eller skrive kommentarer, mens DPA-medarbeidere ved en enhet opplyste at de kun hadde hatt sjekklisten i bakgrunnen. De som brukte sjekklisten aktivt, hadde tilnærmet seg denne på litt ulike måter. Noen satte seg godt inn i sjekklister og støttedokument først, og leste så gjennom rapporten. Andre gikk i gang med vurdering av rapporten, og brukte sjekklisten og støttedokumentet underveis. Her var det tydelig at første tilnærming var den mest effektive med tanke på tidsbruk. Siden det i dag ikke finnes en felles rapportmal, varierer rapportenes oppbygning og struktur fra hvordan sjekklisten er organisert. Dermed brukte mange mye tid på å lete fram og tilbake i rapportene for å finne informasjonen som tilsvarte punktet i sjekklisten og vice versa. Flere fortalte at tidsbruken gikk ned etter hvert som man fikk erfaring med sjekklistene. En DPA-medarbeider som hadde vurdert 5 rapporter anslo at tidsbruken var redusert med 50% fra første til siste rapport.

Det varierte også hvordan sjekklstens komponenter, som trafikklys og kommentarfeltet ble brukt. Noen brukte både trafikklys og kommentarfeltet, noen brukte kun trafikklys og tok eventuelle kommentarer muntlig, mens noen få brukte kun kommentarfeltet.

Trafikklysene, med fargene rødt, gult, grønt og grått, ble i stor grad opplevd som logiske, anvendelige og tilstrekkelige for formålet av deltakerne. Ingen så behov for en utvidet eller tallbasert skala. Unntaket var de som hadde avgrenset fagfelle vurderingen til språkvask og klarhet/tilgjengelighet (kvalitetskontrollhierarkiets nivå 3), hvor trafikklysene ble oppfattet som overflødige til dette formålet.

DPA-medarbeiderne sin oppfattelse av meningen med de ulike fargene så ut til å være nokså sammenfallende. De fleste fagfellene fortalte at rød avkrysning ble brukt på det de mente var en kritisk feil i rapporten. De fleste assosierte rødt med at rapporten måtte endres. Gul avkrysning ble brukt når rapporten burde eller kunne forbedres, mens grønn avkrysning ble brukt når det ikke var behov for justeringer. En DPA-medarbeider kommenterte at det var fint at trafikklysene ga rom for skjønn, slik at man kunne krysse av i gult felt, og dermed påpeke at det var ting som kunne forbedres, men at rapporten likevel kunne 'gå gjennom'.

Trafikklysene tjente flere funksjoner. Ifølge DPA-medarbeiderne begrenset de fortolkningsrommet av den tilhørende kommentaren, slik at en feil som var oppfattet som kritisk av fagfelle ikke så lett kunne tones ned av rapportskriver. Trafikklysene ga oversikt over totalinntrykket av rapporten, og en forklarte det med at når de startet møtet, så kunne man si 'ja, her var det mye gult' eller 'her var det noen røde', og at de derfor fungerte som en slags oppsummering.

Noen fagfeller syntes det var vanskelig å gi røde avkryssinger, og at det skulle være veldig graverende før man gikk til dette steget. Flere opplevde at det var et ganske stort sprang fra gult til rødt, og noen fortalte at om de var i tvil, var det lettere å sette kryss i gult felt enn i

det røde. Flere fortalte at de trodde det ville bli enklere å gjøre røde avkryssninger når de fikk litt mer erfaring. De fleste som hadde gitt tilbakemeldinger tidligere hadde satt søkelyset på språkvask og hvor forståelig rapporten var, og flere uttalte at det å gi tilbakemelding på feil og ting som burde være med i rapporten derfor var uvant og ubehagelig. En framhevet at det var viktig å gi en tydelig begrunnelse når man brukte det røde trafikklyset.

Det grå trafikklyset ble brukt på to måter. Det ble brukt når punktet ikke var relevant for den aktuelle undersøkelsen. Det ble også krysset i grått felt når punktet ikke ble vurdert som relevant for den praksisen de hadde på området, noe som trolig henger sammen med at det per i dag ikke finnes en rutine som beskriver 'beste praksis' for datatekniske undersøkelser. Dermed vil praksisene trolig variere mellom de ulike distriktene, og kanskje også innenfor det enkelte distrikt.

Trafikklysene så jevnt over ut til å ha stor nytteverdi for fagfellene. De effektiviserte og forenklet vurderingen og tilbakemeldingen slik at de fikk lagt vekt på de viktigste tingene. Noen opplevde at trafikklysene ansvarliggjorde dem, siden man måtte ta stilling til om det var en grov feil eller ikke. Om vurderingen gikk over flere dager, gjorde trafikklysene det lett å se hva som var vurdert, og hva som gjensto.

Blant de som brukte kommentarfeltet varierte det hvem de skrev sine kommentarer *for*. Noen skrev kommentarer til seg selv, hvor kommentarene ble en huskeliste over hva man skulle si i tilbakemeldingen. Andre skrev kommentarer som var ment til å leses av rapportskriver. Enkelte hadde brukt mye tid på formulering av kommentarene til rapportskriver slik at disse skulle være klare og entydige. En DPA-medarbeider kommenterte at det oppsummerende feltet til slutt i sjekklisten hadde vært fint til å gi en overordnet tilbakemelding, men også for å notere punkter som ikke var omfattet av sjekklisten.

Kommentarene ble ansett som viktige av flere DPA-medarbeidere for å forklare og nyansere trafikklyset. Flere påpekte at kommentarene

lett kunne misforstås, og det ikke var sikkert man skjønnte ut ifra kommentaren alene at feilen var så stor at rapporten må endres. På den annen side mente en DPA-medarbeider at trafikklysene alene kunne gi et skjevt inntrykk av rapporten, for eksempel at den er av lavere kvalitet enn den faktisk er, og at kommentarene derfor var viktige for å forklare og nyansere et slikt inntrykk.

Ble sjekklisten og støttedokumentet opplevd som nyttige verktøy?

De fleste DPA-medarbeiderne syntes at sjekklisten fungerte bra. Den ble beskrevet som god, og at den skapte refleksjon og at den minnet om ting som du ikke hadde tenkt så mye over ved gjennomlesning. En hadde kontrollert rapporten tidligere, og kontrollerte denne på nytt ved bruk av sjekklisten. DPA-medarbeideren fortalte at hen fant mange forbedringspunkter på kvalitetskontrollhierarkiets nivå 1-4 som hen ikke hadde tenkt på tidligere.

Noen opplevde sjekklisten som lite brukervennlig, spesielt av de som hadde hovedfokus på språkvask og klarhet/forståelighet. De opplevde det enklere å skrive rett i rapporten de vurderte. Flere påpekte at om sjekklisten hadde tilsvart oppbygningen av en standardisert rapportmal, ville den også kunne vært brukt mer effektivt ved at man slapp å lete etter hvor man fant innhold som tilsvarte punktet i sjekklisten, og det ville blitt mindre preget av subjektive vurderinger. En DPA-medarbeider uttalte:

«Jeg synes de funker ganske bra. Det jeg savner er jo at det, vi burde jo hatt en mal. Det har vi jo ikke. Så jeg føler jo at det blir veldig, ihvertfall for min så følte jeg at det ble veldig sånn subjektive vurderinger. Hvordan jeg føler rapporten skal se ut. Fordi jeg har ikke noe annet å referere til, følte jeg litt. Bortsett fra det jeg har lest eller liksom hatt erfaring med selv da». (M6)

Opplevelsen av hvor effektiv og brukervennlig sjekklisten var så også ut til å ha sammenheng med hvor godt man hadde satt seg inn i den før de gikk i gang med vurderingen. En sa det slik:

«Jeg begynte, eller, i begynnelsen så begynte jeg bare på toppen. Men det som slo meg først var at det tar jo, det tar litt tid å kunne, du må liksom kunne sjekklisten på en måte. Så når du bare har rapporten, og så har du sjekklisten, og så kan du egentlig ikke sjekklisten så godt, for du har ikke gjort det før, så går det mye tid på å sette seg inn i sjekklisten. For at det blir veldig mye sånn at du ruller opp og ned og frem og tilbake og, om du liksom har oppfylt sjekklisten på det punktet, eller om det kanskje skal på et annet punkt heller.» (M11)

Støttedokumentet ble beskrevet som godt, utfyllende og at det ga god veiledning til punktene i sjekklisten.

Tidsbruk

Rapportene som ble vurdert varierte stort i omfang, fra 1-2 siders rapporter fra avgrensede undersøkelser, til omfattende rapporter på opptil 50 sider. En uttalte at den typiske rapporten lå mellom 5-15 sider. Ifølge DPA-medarbeidernes beskrivelser var tidsbruken omtrent slik:

- 50 sider: 1 dag
- 30-40 sider: ½ dag
- 10 sider: 2 timer
- 1-2 sider: ½ til 1 time

3.2.4 Tilbakemelding

Hvordan ble tilbakemeldingen gitt?

Det ble ikke lagt noen føringer for hvordan fagfellevurderingen skulle gripes an i pilotfasen, og heller ikke hvordan tilbakemeldingen skulle gis. Dette ledet til mange interessante og nyttige erfaringer.

I ett distrikt hadde DPA-leder tatt en mer aktiv rolle i fagfellevurderingen enn de øvrige, og selv gjennomført to tilbakemeldinger til rapportskriver etter at fagfellen hadde gjort sin vurdering og fylt ut sjekklisten. Ett distrikt skilte seg ut ved at de kun ga skriftlige tilbakemeldinger gjennom 'spor endringer' i tekstbehandlingsprogrammet og kommentarer i marginen. En medarbeider beskrev at hen ikke hadde tilstrekkelig teknisk kompetanse, og dermed fokuserte på språkvask og klarhet/forståelighet, som i stor grad tilsvarete nivå 2-3 i kvalitetskontrollhierarkiet. Samtlige deltakere fra dette distriktet beskrev at de hadde hatt sjekklisten i bakhånd, men ikke brukt trafikklys eller kommentarer. Det som var gjennomgående her var at DPA-medarbeiderne hadde vurdert det som sto i rapporten, men i mindre grad hadde vurdert om det var andre elementer fra nivå 4 i sjekklisten som burde vært med.

Ut over dette var den typiske fremgangsmåten at fagfellen først gjorde en vurdering av rapporten ved bruk av sjekklisten og evt. støttedokumentet, og så hadde et møte med rapportskriver hvor de ga tilbakemelding og diskuterte forbedringspotensialet i rapporten. Noen fagfeller sendte sjekklisten til rapportskriver før tilbakemeldingsmøtet, mens andre ga dette etter møtet. De som kun gjennomførte språkvask og kontroll klarhet/tilgjengelighet hadde som regel ingen etterfølgende prat, med unntak av hvis rapportskriver ga uttrykk for at noe var uklart. De fleste som gjennomførte fagfellevurdering på kvalitetskontrollhierarkiets nivå 4 hadde et uformelt møte ansikt til ansikt hvor de gikk gjennom og diskuterte punktene i fellesskap. Flere

beskrev at de la til rette for at tilbakemeldingen skulle foregå som en samtale fremfor en monolog fra fagfelle. En av fagfellene fremhevet viktigheten av gode relasjoner for å få til denne typen samtaler – og vektla at gode relasjoner åpnet både for diskusjoner og å ta til seg tilbakemeldingene.

Hvordan fungerte tilbakemeldingen?

På spørsmål om hvordan tilbakemeldingen fungerte, var det entydige svaret fra rapportskriverne som hadde mottatt tilbakemelding at dette var ønsket og nyttig. Dette var tilfelle selv om det også kunne være litt flaut å bli gjort oppmerksom på mangler i rapporten. En DPA-medarbeider sa det slik:

«Særdeles nyttig, det var litt sånn, du føler jo på det at du blir blottlagt litt når du har glemt av å ta med tidsstempel her og der eller tidssone og sånt. Men, det er litt sånn, du vet at det blir bedre når du får den tilbakemeldingen, og derfor så blir det på en måte sånn at du leverer et produkt som du er mer stolt av enn hva du ble flau av å få tilbakemeldingen». (M4)

Flere rapportskrivere framhevet at omarbeidingen de gjorde som følge av tilbakemeldingen bedret kvaliteten på rapporten, både gjennom at feil ble rettet, at unøyaktigheter ble fjernet og at konklusjoner ble begrunnet bedre.

Verken de som fikk eller de som ga tilbakemelding fortalte om uenigheter, men snarere om at de var enige, og at det var gode diskusjoner om praksis i tilbakemeldingsmøtet. Flere pekte på at støttedokumentet hadde nytte i disse diskusjonene, og at begrunnelsene og eksemplene i dokumentet hadde bidratt med avklaringer og god veiledning.

Smerten ved å gi tilbakemelding så imidlertid noen ganger ut til å være større enn å motta tilbakemelding. Flere DPA-medarbeidere fortalte at det var vanskelig å påpeke til en kollega at noe var direkte

feil i rapporten, og enkelte syntes det å finne feilen i seg selv var 'ekkelig'. De fleste som hadde lest og vurdert rapporter tidligere hadde konsentrert seg om språk og klarhet/tilgjengelighet, og det var derfor uvant for flere å skulle mene noe det som sto der var korrekt, eller om noe manglet. Det ble også fremhevet at nettopp tilbakemeldingsmøtet var viktig for å kunne forklare hvorfor man hadde krysset av slik man gjorde, og dermed ufarliggjøre de røde avkrysningene.

«Erfaringen vår er at det er veldig greit å sette seg ned i sammen. For da, det skapte på en måte en sånn - ja, for du fikk diskutert litt hvorfor vurderingen var gjort, og hva som kunne gjøres for å fikse, og en fikk også på en måte litt ufarliggjort dette røde da, sant. Du får skapt den der felles forståelsen. Ja, den er rød på grunn av sånn og sånn, og da er det lettere for den andre og å akseptere det da. At ja, du har jo faktisk rett i det». (M9)

Flere DPA-medarbeidere la vekt på at fagfellevurderingen fungerte godt fordi de gjorde dette innenfor enheten hvor de kjente hverandre godt. En DPA-medarbeider pekte på at fagfellevurdering burde gjøres internt på enheten siden det bidro til ansvarliggjøring av dem som gruppe med hensyn til kvaliteten på det de leverte:

«Ja, jeg vil jo si at man ønsker jo å framstå best som en gruppe, så det er ikke hver enkelt som leverer et produkt egentlig, vi er jo sammen om dette her.» (M10)

DPA-medarbeideren uttalte at fagfellevurdering helst burde gjennomføres lokalt for å skape og vedlikeholde en god tilbakemeldingskultur. Samtidig anså hen at kalibrering og standardisering mellom fagfeller fra ulike distrikt også var viktig, og at det måtte skje sentralt.

3.2.5 Blind eller åpen fagfellevurdering

Ved oppstart av pilotfasen oppfordret prosjektgruppen DPA-lederne til å forsøke å få til fagfellevurdering på tvers av distriktene i pilotfasen.

Intervjuene viste at ingen hadde tatt slike initiativ. Dette ble forklart med at de ønsket å teste ut dette lokalt med mål om å bygge en intern kvalitets og tilbakemeldingskultur i første omgang. Ett distrikt hadde forsøkt seg på blind fagfelle vurdering internt i DPA enheten tidlig i pilotfasen, hvor DPA-leder hadde anonymisert rapporten før den ble gitt til fagfelle. DPA-medarbeiderne opplevde imidlertid at denne anonymiseringen ikke tjente hensikten, siden de gjenkjente hvem som hadde skrevet rapporten på tross av anonymiseringen. Det ble dermed raskt konkludert med at en slik anonymisering var nyttesløst når kvalitetskontrollen ble gjennomført internt i DPA-enheten, og de fortsatte uten anonymisering.

Flere DPA-medarbeidere framhevet viktigheten av å utnytte og bygge videre på tette og gode relasjoner internt i distriktet for å oppnå en god tilbakemeldingskultur, og at gevinsten for læring og kvalitet ville bli større ved å styrke dette samarbeidet framfor å kreve en blind fagfelle vurdering på tvers av distrikt som hovedregel. Flere DPA-medarbeidere mente imidlertid at blind fagfelle vurdering på tvers av distriktene var viktig å få på plass på sikt, med formål om å bidra til utjevning av ulikheter i praksis mellom distriktene. Det ble samtidig understreket at dette ikke burde erstatte en lokal åpen fagfelle vurdering, men komme i tillegg eller utvidelse når en god lokal praksis var etablert.

3.2.6 Nivå i hierarki og fagfellens kompetanse

Ved oppstarten av prosjektet formidlet prosjektgruppen at man ved gjennomføring av kvalitetskontroll fra og med nivå 4 (innholdskontroll) som hovedregel skulle fagfelle vurderes av en med *lik eller bedre kompetanse*. Deltakernes erfaringer fra pilotfasen brakte imidlertid verdifulle synspunkter og kunnskap omkring dette temaet fram i lyset.

For det første ble det klart at likt eller bedre kompetansenivå er en forutsetning for å vurdere tekniske aspekter, for eksempel hvorvidt

hensiktsmessige verktøy og metoder var benyttet, og om det var sammenheng mellom funn og konklusjon. Samtidig ble det framhevet fra flere deltakere at utfordringen med rapporter ofte handler om å kommunisere tydelig og forståelig. Det som er tydelig og forståelig dataetterforskere imellom er kanskje uforståelig for en etterforsker, påtalejurist eller dommer. En fagfelle med mindre teknisk kompetanse vil dermed kunne ha bedre forutsetninger for å vurdere om et komplekst datateknisk budskap formidles på en forståelig måte for en leser uten kompetanse på området.

For det andre ble det pekt på at det ikke kun er nivået av kompetanse som har betydning, men også *type* kompetanse. En datateknisk spesialist leser rapporten fra et annet perspektiv enn en jurist eller en politifaglig etterforskningsleder, og vil dermed også fange opp andre momenter av betydning for etterforskningen enn en med datateknisk bakgrunn. Deltakere fra ett distrikt fortalte at de forsøkte å få etablert en redaksjon med fagfeller fra ulike fagbakgrunner som kunne vurdere rapporter, slik at man kunne dra nytte av de ulike perspektivene de brakte med inn i fagfellevurderingene.

Flere la også vekt på at om man faktisk skulle finne feil, så måtte man verifisere funnene mot datagrunnlaget, og ikke kun støtte seg på beskrivelsene i rapporten. Dette innebærer en kontroll på nivå 5 (verifisering av utvalgte funn) eller 6 (verifisering av alle funn) i Kvalitetskontrollhierarkiet. En svarte slik på spørsmål om hvor egnet kontroll på nivå 4 ville være til å fange opp feil:

«Ikke nødvendigvis til å fange opp feil, men til at rapporten blir etterrettelig, slik at du faktisk kan, hvis du oppdager feil i ettertid, så kan du ettergå og finne igjen det som var brukt til da. Mens opp til nivå fire, så er det veldig god kontroll på om det er forståelig det du har levert fra deg på et forklaringsnivå». (M4)

Det var et mål for flere å gjennomføre kvalitetskontroll på nivå 5 i framtiden, men det ble understreket at dette ville innebære langt mer

tid og ressurser enn innholdskontroll på nivå 4. En framhevet at en effektiv måte å komme til nivå 5-6 i Kvalitetskontrollhierarkiet ville være å jobbe i par, slik at man underveis kunne verifisere hverandres funn.

3.2.7 Punkter fra sjekklisten til diskusjon

På spørsmål om det var punkter i sjekklisten som var i strid med eller ikke ga gjenklang i praksis, så ble flere punkter løftet fram av DPA-medarbeiderne. Flere understreket imidlertid at selv om punktene ikke stemte overens med nåværende praksis, så var ikke det ensbetydende med nåværende praksis var hensiktsmessig, og at det ikke var behov for endring. Punktene som ble påpekt hadde ført til diskusjoner internt på enhetene om hvordan praksis er og hvordan den bør være, og hva man bør strekke seg mot i framtiden. Noen fortalte at de allerede hadde endret sin praksis i tråd med punkter i sjekklisten i sin etterforskningsmetodikk eller rapportering, se punkt 3.2.8.

Beskrivelse av kompetanse

I sjekklstens nivå 1, Administrativ kontroll står det følgende om egen kompetanse:

1.5 Er egen kompetanse med relevans for det utførte oppdraget beskrevet?

Ingen av de intervjuede pleide å skrive om hvorvidt de hadde kompetanse til å gjennomføre den aktuelle undersøkelsen, så dette punktet i sjekklisten var helt klart et avvik fra praksis. DPA-medarbeiderne fortalte at vanlig praksis var å angi stillingstittel, for eksempel 'spesialletterforsker' eller lignende. En fortalte at den eneste gangen hen hadde beskrevet egen kompetanse var ved innlevering av rapport under NCFI utdanning ved PHS. En utfordring som ble trukket fram var hvordan en slik kompetansebeskrivelse skulle se ut i rapporten, og at det ikke ville være særlig nyttig om DPA-medarbeiderne begynte å lime inn CV-en sin.

Flere opplyste at selv om de ikke skrev dette i rapporten, var de blitt spurt om dette i retten. En DPA-medarbeider beskrev det slik:

«[...] men jeg har jo blitt spurt om det i retten noen ganger da, det har jeg. For da ville de ofte at, eller, når jeg var der sist i november så ville de at jeg skulle starte litt med å fortelle min bakgrunn og hvorfor jeg kunne mene noe om denne rapporten her da». (M3)

Flere DPA-medarbeidere trakk fram argumenter for at en beskrivelse av relevant kompetanse for oppdraget burde beskrives i rapporten. Flere uttalte at det var viktig å besitte riktig og tilstrekkelig kompetanse for å gjennomføre oppgaven og for å forstå funnene. En DPA-medarbeider pekte på at det var en reell sammenheng mellom kompetanse og om man kunne stole på funnet. En annen understreket at selv om man kanskje hadde tillit innenfor det norske rettssystemet, så var det ikke nødvendigvis slik når saken skulle irettesføres i andre land, og at en kompetansebeskrivelse ville være nødvendig for en slik vurdering. En DPA-medarbeider uttalte at krav til å vurdere og dokumentere egen kompetanse kunne lede til at man faktisk begynte å kreve utdanning for å utføre datatekniske undersøkelser, noe hen anså som positivt. DPA-medarbeidernes argumenter mot en kompetansebeskrivelse var at det ville gi forsvarer anledning til å så tvil om man hadde riktig kompetanse. En DPA-medarbeider uttalte:

«Du stoler jo på at de er kompetente da, men jeg synes det er litt problematisk og, for jeg tenker sånn i retten, hvordan skal det stå seg da at det var liksom satt av en med veldig lang kompetanse til terrorsaken og så til en annen litt mindre men like ganske ille sak så var det liksom satt av en nybegynner, ja. At man skulle begynt å, at forsvarer skulle begynt å så tvil rundt det og dommere stille spørsmål. Men kanskje er det riktig og, jeg er faktisk litt sånn usikker akkurat der». (M2)

En annen DPA-medarbeider pekte på at det kunne utløse et sertifiseringsfokus slik at funn ble trukket i tvil i retten dersom man

ikke hadde riktig sertifisering. En fortalte at selv om en beskrivelse kunne høres imponerende og overbevisende ut, var det ikke ensbetydende med at man hadde rett kompetanse for oppgaven, og at det dermed kunne gi en falsk trygghet.

Verktøy og metode

I sjekklstens nivå 1, Administrativ kontroll, står det følgende om verktøy og metoder:

1.8. Er verktøy med versjonsnummer oppgitt, og framgår det tydelig til hvilke deler av undersøkelsen de enkelte verktøy er brukt?

I sjekklstens nivå 4, Innholdskontroll, står det følgende om verktøy og metoder

4.3. Er framgangsmåte, verktøy og metodebruk beskrevet så nøyaktig at undersøkelsen kan gjentas av andre?

4. 6. Er forbehold, usikkerhet og begrensninger med metoder og/eller verktøy beskrevet?

De fleste DPA-medarbeiderne hadde en etablert praksis med å nevne hvilket verktøy og hvilken versjon de hadde brukt til undersøkelsene. Imidlertid så det ut til å være variasjon i praksis omkring beskrivelse av metodene de hadde brukt for å komme fram til funnene som var omtalt i rapporten. En fortalte at hen tidligere skrev mer detaljert omkring metoder, men at beskrivelsene med tiden var blitt mindre detaljerte fordi ingen brydde seg uansett. Flere DPA-medarbeidere trakk fram at de hadde blitt påminnet om eller mer bevisst på viktigheten av metodebeskrivelser i løpet av pilotfasen. En framhevet at uten en detaljert beskrivelse kunne heller ingen andre rekonstruere arbeidsprosessen for verifisering av funnene. En uttrykte bekymring for at detaljerte beskrivelser ville føre til at rapportene ble veldig lange. Hen fortalte at det gjerne ble gjort omfattende testing omkring

generaliserbarheten av artefaktene, og at selv om dette var et viktig for å kunne trekke slutninger omkring artefaktene, så ville detaljerte beskrivelser av dette arbeidet medføre veldig lange rapporter. En framhevet at beskrivelsene av metodene som ble brukt måtte være på et språk slik at en ikke-tekniker kunne forstå det.

Hypoteser

I sjekklstens nivå 4, Innholdskontroll, står det følgende om hypoteser:

- 4.1 Dersom etterforskningen er hypotesedrevet - er de overordnede hypotesene som danner grunnlaget for undersøkelsen beskrevet?
- 4.2 Dersom etterforskningen er hypotesedrevet - er relevante under-hypoteser (aktivitets og kildenivå) utarbeidet på grunnlag av de overordnede hypotesene eller mandatet/oppdraget?
- 4. 13. Dersom undersøkelsen er hypotesedrevet, er resultatet av undersøkelsene beskrevet i lys av minst to konkurrerende hypoteser, derav en som omfatter uskyld/ikke straffbart forhold?

Hypoteser var et punkt som hadde skapt mye diskusjon ved DPA-enhetene som deltok i pilotfasen. Selv om hypoteser tjener en viktig funksjon for ledelse og målstyring av etterforskningen på overordnet nivå, var det ifølge DPA-medarbeiderne helt uvanlig å få oppdrag som hadde tydelig oppheng i eller utspring fra sakens overordnede hypoteser. Dette resulterte i at mange krysset av i grått felt i sjekklsten på punktene om hypoteser. DPA-medarbeiderne fortalte at de gjerne får mer spissede oppdrag eller spørsmål, for eksempel om det har vært aktivitet på telefonen før en trafikkulykke inntraff, for å svare ut spørsmålet om sjåførens bruk av mobilen forårsaket ulykken eller ikke. Noen beskrev at de nok tenkte på hypoteser under analysen, men at disse ikke ble beskrevet i rapporten. En annen uttalte at hypoteser sikkert var mer relevant når det ikke dreide seg om veldig konkrete spørsmål:

«Ja, så var det også dette med hypoteser som vi også snakket litt rundt. For hypoteser er noe som vi på en måte er vant til å bruke i etterforskningen generelt, men ikke mot de konkrete undersøkelsene. Så det er på en måte ikke noe vi har hatt noe tanke rundt sånn formalisert da. Men litt av grunnen til det er at de analysene vi gjør er ofte veldig små og konkrete, så det er kanskje ikke rom for så mange hypoteser. Mens i kanskje litt større analyser, hvor man har mange spørsmål, så er det kanskje mer relevant med hypoteser da. Men i de analysene vi har, så er det på en måte alltid en etterforsker som har gjennomgått beslaget først, og så går vi gjennom og svarer på konkrete spørsmål. Som for eksempel det bildet, var det slettet, og da blir det enten ja eller nei på det da. Men det er på en måte en grei tanke å ha med seg at man skal tenke hypoteser i de undersøkelsene hvor det kan være en fordel, eller hvor det er mulig å tenke hypoteser da, og det ikke bare er sånn ja nei type ting da». (M8)

Funn og fortolkning av funn

I sjekklstens nivå 3, Kontroll av klarhet og tilgjengelighet, står det følgende om funn og fortolkning av funn:

3.3. Gjør rapportskriver det klart om rapporten inneholder en eller flere av disse rapporteringsformene:

- Beskrivelser av funn (deskriptiv)
- Fortolkninger av hva funnene er, og hva de betyr
- Vurdering (foreløpig eller endelig) av funnenes bevisverdi

I sjekklstens nivå 4, Innholdskontroll, står det følgende om funn og fortolkning av funn:

4.7. Skilles det tydelig mellom rene beskrivelser av funn og fortolkning/evaluering av funn?

Det var ingen uenighet å spore blant deltakerne om at beskrivelser av funn hørte hjemme i datatekniske rapporter. Imidlertid så det ut til å

råde større usikkerhet rundt hvorvidt og hvordan dataetterforskere skulle tolke eller vurdere funn.

En fortalte at det har vært et utydelig skille mellom funn og fortolkning av funn tidligere, og at det var et viktig tema å løfte fram:

M7: «Ja, også punkt nummer tre, beskrivelse av funn, fortolkninger, vurderinger. Der tror jeg det er mange som ikke helt forstår forskjellen på de tre».

Intervjuer: *«Ja, hva tenker du?».*

M7: «Den siste, nei, jeg tenker at man kan gå på en liten smell i retten hvis man gjør for mye fortolkninger, hvis man legger for mye meninger i funnene. Men samtidig så er det jo vi som har kompetansen til å gjøre de fortolkningene. Så det er litt sånn to forskjellige skoler her. Noen mener at vi kun skal beskrive resultater og andre mener at vi er nødt til å gjøre, altså, legge mening i funnene. Så jeg har ikke noe mer på det punktet enn at det er mye uklarhet og usikkerhet og at det er noe som bør bli løftet litt opp. Ihvertfall så man er klar over det».

Flere uttrykte usikkerhet på hvor langt man skulle gå i å fortolke funn, spesielt siden dette kunne lede til det de gjerne omtalte som 'trøbbel' eller 'støy' i retten. Dette kan henge sammen med forventninger til digitale spor, og en uttalte:

«Det har jo vært en gjengs oppfatning at det vi leverer er objektivt uansett». (M4)

Konklusjon og sannsynlighet

I sjekklstens nivå 4, Innholds kontroll, står det følgende om konklusjoner:

4.14. Dersom rapporten inneholder en konklusjon

- er resultatene, omstendighetene og resonnementene som konklusjonen bygger på beskrevet?
- er det er samsvar mellom konklusjonens styrke og funnene den er basert på?
- dersom konklusjonsgrader benyttes, er disse forklart eller henvises til et etablert rammeverk med beskrivelse av de ulike konklusjonsgradene?

Konklusjon og sannsynlighet henger også sammen med sjekklstens punkter som nevner vurdering/evaluering av funnenes bevisverdi, altså 3.3, 4.7 og 4.15.

Når det gjelder evaluering/vurderinger av funn, så er det på det rene at det per i dag ikke finnes noen rutine eller veiledning på hvordan dette skal gjøres i politiets kompetansedelingsystem, ei heller hvordan det skal beskrives i rapport. Flere DPA-medarbeidere henviste til at kriminalteknikk har konklusjonsgrader, men at digital kriminalteknikk ikke hadde etablert noen standard på dette. Dette resulterer i at hver enkelt har sine begreper, og at selv om ordlyden sammenfaller, så kan de være uttrykk for ulike grader av sannsynlighet og usikkerhet. En DPA-medarbeider fortalte at hen hadde sett hen til etterretningsdoktrinens sannsynlighetsbegreper når hen skrev egne rapporter, og tatt dette perspektivet med seg når hen vurderte kollegaens rapport. I diskusjonen under tilbakemeldingsmøtet viste det seg at rapportskriver hadde brukt sin egen subjektive forståelse av disse uttrykkene, og at de derfor hadde et annet meningsinnhold.

På spørsmål var det ingen deltakere som kjente til noen metodikk for å vurdere bevisenes verdi eller styrke. Dette innebærer at prosessen fram til en slik vurdering blir preget av den enkelte DPA-medarbeiders praksis og skjønn.

Det ser ut til å være mange ulike forståelser rundt kravet til objektivitet. Noen så ut til å legge fravær av subjektivitet inn i

forståelsen, og at man dermed kun burde rapportere funn, og ingen tolkninger eller vurderinger av funn. I stedet for å synliggjøre sin rolle i vurderingen, fortalte en DPA-medarbeider at hen brukte formuleringer som «dataene viser at...». Andre syntes imidlertid det var viktig at dataetterforskere tolket og vurderte, så lenge de hadde kompetansen til å gjøre det, og at det ble gjort på bakgrunn av grundige undersøkelser. En DPA-medarbeider uttalte at sannsynlighetsbegrepene i dag ble brukt som forbehold framfor aktive vurderinger:

«Ja, også blir det litt sånn hvis man skal skrive en litt lengre oppsummering, så skriver du, ja det er sannsynlig at den ble satt på lader, og det er sannsynlig at det programmet ble åpnet. Da er du, hvis man bruker det for mye også, så har jeg litt sånn opplevelse av at det blir litt sånn fraskrivelse av, kanskje ikke ansvarsfraskrivelse, men at man liksom prøver å fjerne seg litt fra ansvaret i den konklusjonen. At man fjerner seg litt fra det, og egentlig ikke tar noe ordentlig standpunkt. Så man gjør det kanskje litt for å sikre seg selv er kanskje det jeg har litt opplevelse av». (M7)

Etikk og jus

I sjekklstens nivå 4, Innholdskontroll, står det følgende om etikk og jus:

4.16. Er undersøkelsen og rapportskrivningen gjennomført i tråd med gjeldende straffeprosessuelle regler og relevante etiske retningslinjer?

4.17. Framgår det av rapporten hvorvidt eller i hvilken grad dataetterforsker kan betraktes som uavhengige med hensyn til sine undersøkelser og dokumentasjon?

4.18. Er det søkt aktivt etter informasjon som kan underbygge uskyld eller formildende omstendigheter for mistenkte, og er resultatet beskrevet i rapporten?

Enkelte framhevet at de ikke forsto hvorfor punktene omkring etikk og jus var med. En påpekte at det var mye uavklart omkring reglene rundt beslag, og at de ventet på retningslinjer som trolig ville gjøre dette klarere.

3.2.8 Endring av egen og lokal praksis

En effekt av pilotfasen, som ble tydelig gjennom intervjuene, var at fagfellene også hadde endret egen etterforsknings- og rapportskrivingspraksis. En DPA-medarbeider fortalte at hen nå brukte sjekklisten som utgangspunkt når hen skrev rapport, og sørget for at punktene ble dekket inn. Dette ville også gjøre den etterfølgende kvalitetssikringene enklere. En DPA-medarbeider beskrev det slik:

«Jeg tenkte faktisk på, det var en ting i forhold til, det er veldig bra å bruke et sånt skjema for å gjøre hverandre gode. Men jeg tenker også for min egen del så - når jeg hadde brukt skjemaet noen ganger. Jeg har sett skjema før så jeg har liksom hatt de punktene i bakhodet da, og det har kanskje fått meg til å tenke på det også når jeg skriver rapport, at det og det bør være med da. Så det var ganske sånn umiddelbart, jeg tror når jeg hadde brukt det skjemaet på kollegaen min sin rapport, så gjorde jeg meg opp noen tanker når jeg skulle skrive min egen rapport da». (M12)

En annen fortalte at hen var blitt mer nøye med å avklare mandatet for undersøkelsen. Dette ledet til mer målrettede undersøkelser og lettet den etterfølgende rapportskrivningen. En DPA-medarbeider fortalte at hen nå deler konklusjonen i to, og presiserer hva som er funn og hva som er egen tolkning av funnene basert på undersøkelsene som er gjort.

En DPA-medarbeider fortalte at å bli oppmerksom på egen praksis gjennom å motta gule trafikklys var god læring, og gjorde hen inspirert til å ta tak i dette i neste rapport. Sjekklistene ble ikke kun sett på som et verktøy til nytte for en selv og sin egen praksis, men også noe som kunne hjelpe dem å som gruppe til å bli bedre.

3.3. Eksempler på feil og håndteringen av disse

3.3.1 Feil

Resultatene viser at samtlige fagfeller fant forbedringspunkter i rapportene de vurderte. Flere av disse forbedringspunktene ble karakterisert som kritiske feil. Deltakerne ble også spurt om de hadde andre eksempler på feil fra eget politidistrikt relatert til digitale spor som hadde skjedd eller hadde vært nær ved å skje. Det ble understreket at de ikke måtte gå ut over sin taushetsplikt om straffesaker eller involverte personer. De fleste av deltakere beskrev et eller flere konkrete eksempler på feil eller nesten-feil, og noen beskrev også typiske feil i mer generelle termer. Noen av eksemplene relaterte seg til arbeid utført av dataetterforskere, mens andre handlet om saker håndtert av etterforskere eller fagkontakter. I den videre framstillingen fokuseres det i hovedsak på mekanismene som ledet til feilen eller nesten-feilen framfor hvem (hvilken funksjon) som var involvert. For ordens skyld gjøres det oppmerksom på at feilene som beskrives under ikke er avgrenset til Kvalitetskontrollprosjektets pilotfase.

Mange av eksemplene var relatert til *feil eller villedende informasjon i dataprogrammene* som brukes av DPA-medarbeiderne til tekniske analyser eller av fagkontakter/etterforskere til innholdsanalyser (gjennomgang).

- En hendelse hvor ungdommer hadde slåss var sentral i etterforskningen av en voldssak. Noen hadde filmet slåssingen med mobilen sin og distribuerte denne i sosiale medier. Etterforskningen handlet blant annet om å identifisere åstedet og tidspunkt for hendelsen. Det ble tatt beslag i - og sikret innhold fra - flere mobiltelefoner. Når videomateriale fra de ulike beslagene ble sammenholdt viste verktøyet samme geolokasjon, og det ble dermed konkludert med at dette var stedet hvor slåssingen hadde foregått. En manuell undersøkelse av lokasjonsdata fra videomaterialet viste imidlertid at det

man trodde var en eksakt lokasjon dekket et område på en kvadratkilometer, og at hendelsen hadde skjedd på et annet sted og til en annen tid enn det analyseverktøyet hadde indikert.

- I 2019 ble det oppdaget at en versjon av et verktøy feiltolket tidsstempler. DPA-medarbeiderne på enheten ble bedt om å finne ut hvilke straffesaker denne verktøyversjonen var blitt bruk i, slik at man kunne finne ut om feilen hadde fått noen uheldige følger i etterforskningen. Siden mange ikke hadde god nok notoritet på hvilke verktøy og versjoner de hadde brukt til sine undersøkelser, hadde flere DPA-medarbeidere ved enheten utfordringer med å identifisere sakene. Selv hadde DPA-medarbeideren vært nøy med å notere versjonsnummer som følge av en lignende erfaring fra tidligere som hadde skapt mye støy, og det var derfor lett å finne tilbake til de aktuelle sakene.
- En straffesak som omhandlet et drapsforsøk hadde vært i tingretten, og den datatekniske rapporten ble tatt for 'god fisk'. Når saken skulle opp i lagmannsretten ble det imidlertid klart for DPA-medarbeideren at verktøyet hadde presentert data på en unøyaktig måte, og at siktede hadde tilpasset sin forklaring til funnene i rapporten. Dette ble rettet opp gjennom nye undersøkelser, og ny rapport ble skrevet. Tiltaltes forsvarer brukte dette til å så generell tvil om rapporter fra dette fagfeltet. Dommeren la ikke vekt på dette, og tiltalte fikk strengere straff i lagmannsretten sammenlignet med dommen fra tingretten. Om feilen ikke hadde blitt avdekket, kunne den altså ført til at tiltalte fikk for lav straff for forholdet hen var tiltalt for.
- I en sak framstilte verktøyet en WIFI-tilkobling som tilkobling til lader. Dypere undersøkelse av databaser og datastrukturer avslørte feilen.

- I et tilfelle hadde bilder med amerikansk dato med måned først blitt feiltolket av verktøyet, noe som medførte at bildene ble presentert i feil rekkefølge i tidslinjen i programmet.

Mange av eksemplene handlet om *feiltolkning av funnene* i analysen. Felles for flere av feiltolkningene er at det trekkes feilaktige slutninger om aktiviteter relatert til funnene. Slike feiltolkninger så ofte ut til å ha sammenheng med utilstrekkelig datateknisk kompetanse.

- I trafikksak hvor en bilfører hadde kjørt på en fotgjenger, ble sjåføren siktet. Siktedes mobil ble sikret og gjennomgått. På mobilen ble det funnet en video som var filmet gjennom frontruten. Man kunne se dashbordet, høre at det ble spilt høy musikk og se på speedometeret at bilen kjørte i 180 km/t. Tidsstempelen på videoen stemte godt overens med tidspunktet for påkjørselen. Dermed var det kort vei til konklusjonen om at føreren råkjørte rett forut for påkjørselen. Det ble skrevet rapport som dokumenterte funnet av denne videoen. Dataetterforskere i distriktet var imidlertid kritiske til dette funnet og undersøkte om videoen og råkjøringen kunne settes i sammenheng. Nye undersøkelser viste at videoen ikke var filmet med telefonen, men at den var blitt sendt til siktede rett forut for hendelsen. Tidsstempelen i videoen var altså fra når videoen ble mottatt på siktedes telefon. Videoen var filmet i en annen bil, som riktignok var veldig lik siktedes, men hadde ingen sammenheng med påkjørselen. Dette ble heldigvis avdekket før saken ble behandlet i retten.
- I en smuglings sak ble GPS-lokasjoner funnet i innholdsanalysen feiltolket som en faktisk gjennomført kjørerute. Det ble besluttet pågripelse og ransaking. Teknisk undersøkelse viste at GPS-lokasjonene stammet fra en kjørerute siktede hadde søkt på, men som ikke var fysisk gjennomført. Pågripelse og ransaking var dermed gjennomført på feilaktig grunnlag.

- I en seksualforbrytelsessak ble det funnet videomateriale som viste overgrep, og det ble skrevet rapport om dette. Saken ble avsluttet, oversendt til påtale, og var på vei til forhåndsberamming. Påtalejuristen ba om en teknisk rapport fra dataetterforsker for å være på den sikre siden. Undersøkelsene viste at siktede aldri hadde åpnet videoene det var snakk om, men kun surfet på nettsiden. Som følge av den tekniske rapporten ble saken henlagt.
- I en sak ble funn av miniatyrbilder (thumbnails) tolket som nedlastet materiale. Det viste seg at siktede hadde besøkt nettsiden, men ikke lastet ned bilder. Miniatyrbildene hadde blitt automatisk lagret i datamaskinens minne, og var feilaktig tolket som nedlastet ulovlig materiale.
- I en straffesak som omhandlet en seksualforbrytelse ble et bilde med relevans for straffesaken funnet på en mobiltelefon. I rapporten ble det skrevet at bildet var lagret på telefonen til et gitt tidspunkt. Tekniske undersøkelser viste at bildet hadde dukket opp på telefonen som resultat av at det hadde havnet i en chat-mappe på Snap story (Snapchat), og at konklusjonen om at det var lagret av siktede dermed var feil.
- Et bilde med relevans for en sedelighetssak ble funnet, og det ble dokumentert i rapport hva bildet framstilte og hvor det var funnet. Det ble dermed ansett som besittelse av ulovlig materiale. Forsvarer tok kontakt og fortalte at ifølge siktede, så hadde hen fått bildet tilsendt i WhatsApp-chat, og at hen så hadde slettet det. Tekniske undersøkelser viste at siktetes versjon var riktig.

Feil knyttet til *feil eller unøyaktigheter i håndtering eller sikring*

- Sikret materiale fra en mobiltelefon ble lagt inn i feil prosjekt. I dette prosjektet lå det overgrepsmateriale, mens i den nye telefonsikringen fantes det ikke slikt innhold. Dermed så det ut

som det var overgrepsmateriale på mobiltelefonen som nå var sikret. Feilen ble oppdaget etter at saken var oversendt til annet distrikt, men før det ble gjort noen videre etterforskningsskritt mot siktede. Om ikke feilen var blitt avdekket i tide ville trolig vedkommende blitt pågrepet og siktet for besittelse av overgrepsmateriale.

- I en sak om seksualforbrytelse ble det konkludert uriktig på bakgrunn av et tidsstempel. Aktiviteten man tilskrev siktede viste seg å være politiets egen aktivitet på datamaskinen under ransakingen og beslaget av denne. Dersom dette ikke hadde blitt oppdaget, ville dette feilaktig kunne blitt brukt som et bevis mot at siktede hadde utført noe han faktisk ikke hadde utført.

Feil relatert til identifisering av mistenkte

- Ved et tilfelle ble en uskyldig person pågrepet og fengslet for besittelse av overgrepsmateriale. Saken stammet fra NCMEC⁴ og var oversendt til distriktet fra Kripos. Kripos hadde søkt opp abonnenten for den aktuelle IP-adressen, som var en av de store teleleverandørene i Norge, og den oppgitte gateadressen tilhørte teleleverandørens hovedkvarter. Datamaskinen som var benyttet var identifisert og det var også oppgitt et navn på en person med tilknytning til datamaskinen. Kripos hadde gjort registersøk på personen, og funnet ut at hen bodde i det aktuelle politidistriktet, og saken ble så sendt over med denne informasjonen. Personen ble pågrepet, satt i arrest og samboeren ble tatt inn til avhør. Det viste seg at IP- adressen stammet fra en ruter i teleleverandørens hovedkvarter som hvem som helst kunne koble seg opp til. Den pågrepne hadde kun ansvar for PC-en som sto koblet opp til ruterens og hadde ingenting med det mistenkte forholdet å gjøre. Den pågrepne ble løslatt, og politidistriktet har som følge av feilen innført nye rutiner for hvordan de vurderer slike saker.

4 National Center for Missing & Exploited Children

- I en sak oversendt fra et annet nordisk land feil ble et tegn feilaktig inkludert i en epost adresse, og den påfølgende IP-springen ledet dermed til feil person. Dette ledet til at en uskyldig person ble pågrepet, ransaket og satt i arrest.
- I en straffesak som omhandlet nedlasting av overgrepsmateriale fra internett ble huseier pågrepet, siden hen sto som abonnent på internettabonnementet, mens leieboer var den som faktisk hadde lastet ned det ulovlige innhold fra internett.

Villedende framstilling

- I en rapport som dokumenterte funn etter gjennomgang av databaseslag, var det presentert en samtalelogg mellom to personer i en tabell. Ut ifra tabellen framsto samtalen som 'rådata' fra en sammenhengende samtale mellom siktede og fornærmede. Siktede gjennomgikk rapporten sammen med sin forsvarer, og kjente seg ikke igjen i samtalen slik den var framstilt i rapporten. Da forsvarer ba politiet undersøke dette, viste det seg at etterforsker hadde klippet og limt ulike utdrag fra samtalen sammen fra forskjellige tidspunkt, og plassert dette i en tabell uten å opplyse at samtalen var redigert. Etterforsker fikk beskjed om å endre rapporten.

Generelle beskrivelser av *vanlige feil*

- Mange av de mer generelle beskrivelsene av feil var ifra straffesaker som omhandlet **internettrelaterte seksualforbrytelser**.
 - Flere deltakere trakk fram **manglende kvalitetssikring** av saker fra IcacCops⁵ og NCMEC som et gjentakende problem. I slike saker var det gjerne mistanke om nedlasting av bilder eller video som framstiller overgrep mot barn, og typiske feil i slike saker var mangelfulle undersøkelser av om mistankegrunnlaget, feil ved IP-sporing og feil ved identifisering av mistenkt som følge av IP-sporing. Feilene kunne føre til feilaktige pågripelser og ransakinger.
 - Ifølge deltakerne gjøres det stadig **feiltolkninger av miniatyrbilder** (thumbnails) i etterforskning av internettrelaterte seksualforbrytelser. I analyseverktøyet kan man se at bildet befinner seg på det sikrede beslaget, men verktøyet sier imidlertid ingenting om hvordan det havnet der. Slike bilder blir automatisk lastet ned og lagret midlertidig på maskinen når man besøker ulike sider på internett, selv om man ikke har trykket på bildene eller gjort andre grep for å se på dem eller laste dem ned til datamaskinen. Feiltolkning av slike bilder har gjentatte ganger dannet grunnlag for siktelse for besittelse av bilder av overgrep mot barn.
 - Flere deltakere trakk fram at **funn av ulovlig innhold ofte blir feiltolket som resultat av aktive handlinger** av den som disponerte enheten. For eksempel blir filer i datamaskinens mappe for mellomlagring (eng. cache memory) som er resultat av 'surfing' på websider ofte tolket som aktivt nedlastede bilder.

⁵ Program for overvåking av ulovlig nedlasting av merket overgrepsmateriale fra ulike fildelingsplattformer

- Manglende kunnskap om kriminalitetsfenomener og typiske spor kunne føre til at **viktige bevis ble oversett**, og hvor følgen ble at saken ikke ble så stor eller alvorlig som den burde være. Når det eksempelvis ble funnet bilder eller videomateriale fra overgrep mot barn, ble det gjerne raskt konkludert med besittelse av ulovlig materiale, og etterforskningen ble avsluttet. I slike saker kan imidlertid materialet stamme fra direkteoverførte bestillingsovergrep, som er et langt mer alvorlig forhold, og sakens egentlige omfang og alvorlighet blir dermed ikke avdekket. Å finne sporene av slik aktivitet krever kunnskap om kriminalitetsfenomenet og om hvilke spor som skapes når slike straffbare handlinger gjennomføres.
- Det trekkes feilaktige slutninger omkring **geolokasjoner**. Selv om funn av lokasjonsdata kan bety at datamaskinen har vært på eller i nærheten av et fysisk sted, er ikke dette alltid riktig. Det kan finnes lokasjonsdata som ikke har sammenheng med hvor datamaskinen har vært. Lokasjonsdataene kan for eksempel være hentet inn fra WIFI-nettverk. En deltaker fortalte at man eksempelvis kunne finne datapunkter for hele Oslo med samme tidspunkt i ett og samme beslag. Analyseverktøyet kan også fremstille lokasjoner feil eller unøyaktig, slik at posisjonen som funnet relateres til avviker fra det som var den reelle posisjonen.
- **Bakgrunnsprosesser blir feiltolket som aktiv bruk** av datamaskinen. Deltakerne fortalte at manglende teknisk kompetanse kunne føre til feiltolkning av systemgenererte spor og feilslutninger om at sporene skyldtes brukerens aktive handlinger.
- Funn blir feilaktig satt i sammenheng med den beslaglagte databæreren, mens det egentlig er resultat av **synkronisering på tvers av enheter**. Dette leder for eksempel til at

telefonsamtaler blir knyttet til en iPhone, mens de i virkeligheten skjedde fra en iPad knyttet til samme bruker.

- **Verktøy blir ansett som feilfrie.** En overdreven tro på feilfrie verktøy kan føre til at funn beskrives uten å kontrollere om verktøyet har framstilt funnet korrekt. Det varierer også hvor oppdaterte verktøyene er til å takle ulike dataformater, noe som medfører at data kan fremstilles ulikt i ett verktøy versus et annet. Verktøyversjoner har fra tid til annen 'bugs' som leder til feilaktig eller villedende framstilling av informasjonen.
- **Det skjer feil ved konvertering av tidsstempler.** Tidsstempler må konverteres til en tidssone for å få en riktig kronologisk framstilling av innholdet på den digitale enheten. En deltaker fortalte om et tilfelle hvor det ble lagt til to timer framfor å trekke fra to timer ved konvertering av tidsstemplene i en straffesak, hvor disse tidspunktene var sentrale.
- **Rapportskriver sitter med mer kunnskap enn det man får fram i rapporten,** derfor blir det ikke alltid harmoni mellom funn og konklusjon. Misforhold mellom funn og konklusjon betyr ikke nødvendigvis at det er feil, men at konklusjonen ikke er tilstrekkelig underbygget av funnene den svarer til.

3.3.2 Håndtering og varsling

Ifølge DPA-medarbeiderne var det ingen fastlagte rutiner for hvordan man skulle opptre når feil i verktøy eller metoder ble oppdaget. Det var heller ingen rutiner for å kartlegge hvilke øvrige straffesaker feilen kunne ha påvirket, og hvilke følger det eventuelt hadde ledet til.

Noen DPA-medarbeidere fortalte at de varslet andre DPA-medarbeidere via interne uformelle kommunikasjonskanaler, som f.eks. Signal og Teams, mens andre oppga at de rapporterte feilen i Synergi. En DPA-medarbeider fortalte at hen på bakgrunn av signaler fra sin leder oppfattet det å rapportere i Synergi som en eskalering

sammenlignet med de uformelle kommunikasjonskanalene, og at det ville blitt veldig mange Synergi-meldinger om alle feil skulle blitt rapportert der.

DPA-lederne henviste også til Synergi, men flere beskrev systemet som tungvint. Flere DPA-ledere uttalte at de var usikre på om informasjon om feil når ut i tide til de som trenger det om den legges inn i Synergi, og en DPA-leder uttalte:

«At man sørger for å ha en robust rigg for å ivareta den nye kunnskapen da. Om det er på Synergi – ingen ligger jo der og surfer, ingen surfer på Synergi. Så å få dette inn da, om det skal være et månedsbrev, at det blir en sånn, KO:DE, siste oppdateringer, nb, obs, obs, dette må folk lese. At hvermannsen som jobber med fagfeltet skal kunne oppsøke denne nye kunnskapen og at det ikke er kun den som har gjennomført selve evalueringen og gjennomgangen og den som har skrevet rapporten. Det er de to som sitter på kunnskapen inntil den blir offentliggjort». (L2)

Ifølge DPA-medarbeiderne fantes det ingen etablerte rutiner for å hvordan programvareleverandøren skulle varsles om feil i verktøy. DPA-lederne pekte mot både fagforvaltningen og NC3 som viktige roller ved håndtering av feil. Flere DPA-ledere uttalte at feil i verktøy bør håndteres sentralt. De fremholdt at dette var viktig slik at alle som har brukt metoden/verktøyet blir varslet, og at konsekvensene av feilen blir kartlagt og minimert raskt og effektivt.

3.4. Veien videre

3.4.1 Holdninger til kvalitetskontroll

Samtlige av de intervjuede DPA-lederne understreket betydningen av god kvalitet, og at kvalitetskontroll var et viktig tiltak for å oppnå dette. De fortalte at de stadig oftere opplever at resultater fra datatekniske

undersøkelser ikke lenger blir lagt til grunn som fakta, og at de utfordres i stadig større grad i retten.

«Ja, vi har en del straffesaker hvor vi ser at forsvarerne stiller mere spørsmål. Sågar så har de egne sakkyndige som stiller spørsmål til politiets metode, blant annet måten vi har valgt ut bevis på og så videre og så videre. Og det merker vi at det skaper en del støy da. Og ikke minst utfordringer på at, oj, det var jo..., og vi er ikke vant til at folk spør hvorfor, hvordan og på hvilken måte.» (L1)

De fleste DPA-medarbeiderne ønsket en *systematisk* kvalitetskontroll, og begrunnet dette med at det ville øke kvaliteten, hindre feil og at man ville oppnå bedre rettsikkerhet. En DPA-medarbeider beskrev det slik:

«Jeg har erfart selv å stått i retten og hatt en rapport som skapte unødvendig støy. Ikke at det var noe feil, men at jeg hadde beskrevet det for dårlig. Og det er liksom ting som kunne vært unngått hvis man hadde gjort en ordentlig gjennomgang da». (M1)

Det ble også framhevet fra flere deltakere at systematisk kvalitetskontroll ville være et sikkerhetsnett for dataetterforskerne, slik at ikke alt sto og falt på *en* persons tolkning.

Flere av DPA-medarbeiderne med sivil bakgrunn fortalte at da de startet sin jobb i politiet hadde de forventet langt mer sentral styring og sentrale føringer. I bedriftene de kom fra var det gjerne betydelig høyere krav til kvalitetssikring på produktene de leverte enn det som møtte dem da de ble ansatt i politiet.

De fleste DPA-medarbeiderne mente at å gjøre kvalitetskontroll av dataetterforskernes tekniske rapporter *obligatorisk* var viktig og riktig, og at valgfrihet ville føre til at det ble prioritert bort i en hektisk hverdag. Samtidig ble det understreket av flere deltakere at å gjøre kvalitetskontroll obligatorisk ikke nødvendigvis ville sikre kvalitet,

siden man måtte ha ressurser til å gjøre det ordentlig. En DPA-medarbeider sa det slik:

«Jeg vet at det er en utfordring, jeg vil anta det er en utfordring for flere som jobber i politidistrikter, når det er så høyt sakspress for det er vanskelig å sette av den tiden til å faktisk gjøre det. Og da er du jo redd for at mye av disse punktene bare kan være litt sånn check, check, check gjennom det hele». (M1)

Flere DPA-medarbeidere uttalte at de opplevde at leder var enig i at kvalitetskontroll var viktig og antok at de ville få støtte til å fortsette å gjennomføre kvalitetskontroll ved egen enhet også når pilotfasen av Kvalitetskontrollprosjektet var avsluttet. Flere uttalte at siden de skrev relativt få tekniske analyserapporter, ville det være oppnåelig å gjennomføre at kvalitetskontroll på nivå 4, innholdskontroll, som en del av arbeidsprosessen.

Også DPA-lederne pekte på at kvalitetskontroll burde settes i system, og gjøres *obligatorisk*. Uten klare rammer og forankring advarte flere av DPA-lederne mot at kvalitetskontroll ville bli prioritert ned framfor å få saker unna. Rutinebeskrivelser og maler, sjekklister og krav til gjennomføring ble fremhevet som viktig for å sikre et visst nivå av kvalitet, men også en mer lik praksis og likt kvalitetsnivå mellom distriktene.

Å endre en etablert praksis er vanskelig, og flere DPA-ledere understreket behovet for riktig forankring i POD og fagforvaltningsapparat. Samtlige pekte flere på at det ville kreve ressurser for å få på plass en systematisk kvalitetskontroll. Det vil innebære at saksarbeidet tar noe lenger tid i den enkelte sak. Flere DPA antok at ressursbruken ville tjenes inn på lang sikt. En sa det slik:

«Så det er utgift til inntekts ervervelse egentlig, sånn tenker i alle fall jeg at tilnærmingen er. Det har ikke nødvendigvis en kortsiktig gevinst - men en langsiktig gevinst med at vi gjør det sånn». (L5)

I tillegg til økt kvalitet ville obligatorisk kvalitetskontroll føre til færre hasteoppdrag dagen før hovedforhandling på grunn av usikkerhet omkring de digitale bevisene. Den viktigste gevinsten ville ifølge DPA-lederne være økt rettssikkerhet for de impliserte i straffesakene. En DPA-leder uttrykte det slik:

«Det er derfor jeg sier at kvalitetskontroll av rapporter – behovet er der. Og det man snakker om her er, skal vi prioritere rettssikkerheten eller skal vi prioritere mengden av arbeid?». (L3)

I likhet med DPA-medarbeiderne, påpekte flere DPA-ledere at en innføring av obligatorisk kvalitetskontroll uten styrking av ressurser kunne føre til slett kvalitetskontroll. En DPA-leder uttalte at et slikt krav kunne føre til at det ble et jag for å komme seg gjennom sjekklisten uten noen særlig refleksjon. Dette kunne dermed gjøre vondt verre, ved at tilliten til funnene blir styrket uten at det egentlig er grunnlag for det.

3.4.2 Om å bygge kvalitets- og tilbakemeldingskultur

DPA-lederne understreket betydningen av å etablere en god kvalitetskultur lokalt, med psykologisk trygghet og tillit, slik at det føltes ok både å gi og få tilbakemelding på arbeidet sitt. Med dette på plass var et naturlig neste steg være å utvide til å gjennomføre fagfellelvurdering på tvers av distrikter. Da ville man også kunne organisere blind fagfellelvurdering, men også bidra til innsikt i hverandres arbeidsmetoder og muliggjøre en større grad av harmonisering.

«Jeg ønsker å bidra inn til at det skal bli et bedre system i hele landet. Både at vi gjør ting litt likt, det er noe vi jobber mye med på de DPA-ledermøtene også gjennom hele året. Vi har månedlig møter og et par ganger har vi fysiske møter. Og det går på det at vi prøver å få til å gjøre ting på samme måte. Og få en lik kvalitet over hele landet da, som er

vanskelig når ressursene er forskjellig. Men det liksom noe med at man kan lage en minimumsløsning som alle kan være med på». (L4)

Flere DPA-ledere pekte på NC3 ved Kripos som en viktig aktør i et system for kvalitetskontroll siden de besitter spisskompetanse innenfor flere områder, og at de kanskje også har den beste oversikten over hvilke dataetterforskere i politidistriktene som har spisskompetanse innenfor ulike områder. Enkelte DPA-ledere anså også kvalitetskontroll som en oppgave de burde kunne anmode om bistand til fra NC3 når ressursene ikke strakk til i distriktet.

3.4.3 Motstand i retten

Flere deltakere trakk fram at mer detaljerte beskrivelser i rapportene om for eksempel egen kompetanse, anvendte undersøkelsesmetoder og vurderinger av funn ville gi mer motstand i retten og bedre anledning for forsvarer til å så tvil om undersøkelsene.

«Vi må nok regne med at vi kommer til å bli sett nærmere i kortene, det tror jeg. Vi har i mange mange år egentlig bare flytt på en sky at, det vi produserer og det vi viser i retten, det er opplest og vedtatt. Vi får ingen kritiske røster imot oss. Men det tror jeg er noe som kommer til å endre seg. Og vi har vel egentlig sett litt konturene av det allerede». (L1)

En DPA-medarbeider uttalte at mer motstand i retten kunne føre til at dette kunne ha en positiv effekt på kvaliteten. Selv om de fleste opplever lite motstand i retten per i dag, uttalte flere deltakere at de så tegn til ending, og at de forventet mer bruk av 'motvitner', altså partssakkyndige vitner engasjert av forsvarer, i framtiden.

3.4.4 Tid og ressurs

Under workshopene høsten 2022 (før Kvalitetskontrollprosjektets pilotfase) ble det uttrykt stor skepsis til hvorvidt det ville være mulig å finne tid til å gjennomføre kvalitetskontroller av datatekniske rapporter. I intervjuene etter at pilotfasen var gjennomført var deltakernes

innstillinger til dette nyansert noe. De fleste anså kvalitetskontroll av datatekniske rapporter som et viktig tiltak som man burde sette av tid til. Noen anså det som mulig innenfor rammene av arbeidet de utførte i dag, mens andre framhevet at det ville kreve at de fikk mer ressurser hvis de skulle opprettholde samme nivå på produksjonen som de hadde i dag. En DPA-leder uttalte:

«Jeg mener det er gjennomførbart. Men det føyer seg inn i listen over gjøremål som må utføres da. Men jeg tenker at vi av alle bør ha tid og ta oss tid til å gjøre den type ting, at det bør inngå som en del av - la oss kalle det - daglige virke, tenker jeg». (L2)

Noen DPA-medarbeidere framhevet imidlertid at de var usikre på om kvalitetskontroll totalt sett ville kreve mer ressurser. De begrunnet dette i at når det ble oppdaget feil og mangler rett før saker skulle opp i retten, så utløste dette gjerne store og omfattende undersøkelser. Dette forsinket i tillegg det ordinære saksarbeidet, som måtte settes på vent inntil det ekstraordinære var unnagjort. En DPA-medarbeider fortalte:

«ja, og så i tillegg så har du de sakene som du da jobber med. Dem må jo legges bort for den perioden. Og det er alltid, jeg vet ikke hva vi kan regne med her, men oppstartskostnaden med å legge bort en sak og ta den opp på nytt igjen. Det forsvinner ofte en halv dag før du er tilbake der du var i realiteten. Ikke bare at du mister en dag på å se på den andre saken, men du taper en halv dag til for at du skal komme deg tilbake der du var. Det har vi ikke gjort noe mye målinger på, men det føles ihvertfall ut som det tar veldig lang tid mange ganger å komme inn i det». (M4)

Drift av systemer og forvaltning av verktøy og lisenser ble trukket fram som tids- og ressurskrevende oppgaver som heller burde vært sentralisert, slik at det ble frigjort tid til andre oppgaver – som for eksempel kvalitetskontroll. En DPA-leder uttalte:

«Og det er jo synd at kvalitet og rettssikkerhet må gå på bekostning av at vi ikke har nok ressurser til å få gjort dette på en god nok måte. Men vi gjør det der vi kan». (L7)

3.4.5 Standardisering

Datateknisk etterforskning er uregulert ut over de generelle reglene for straffesaksbehandling. Behovet for en større grad av standardisering, sentral styring og koordinering ble framhevet både av DPA-ledere og DPA-medarbeidere. Det finnes per i dag ingen rutine for hvordan en datateknisk undersøkelse skal gjennomføres i straffesaksetterforskning, og det ble pekt på behov for en *nasjonal standard for dataetterforskning*, slik at det ikke ble ulike praksiser og nivå på undersøkelser og kontroll i hvert politidistrikt.

Mange etterlyste en mal for teknisk analyse, og understreket at den ville gjort både rapportskrivningen og kvalitetskontrollen mer effektiv. I tillegg pekte flere på behovet for maler for sikringsrapport og rapport for tilrettelegging for innholdsanalyse/gjennomgang. Flere fortalte at de var i gang med å lage egne rapportmaler. Dette er en form for lokal standardisering, men innebærer samtidig at det kan bli ulike standarder mellom distriktene i fravær av en sentral mal.

Manglende rutiner for beslagshåndtering, analyse, rapporter og dokumentasjon ble fremhevet som risikofaktorer for ulikhet mellom distrikter og utilstrekkelig kvalitet i arbeidet. Det er per i dag ulike forståelser og praksiser omkring hva som anses som beslag. Flere deltakere fortalte om usikkerhet omkring den digitale enheten (mobilen, PC-en eller lignende), sikringsfilen, databaser som eksporteres ut av sikringsfilen eller andre uttrekk fra sikringsfilen som kopieres inn i rapporter er å anse som enkeltstående beslag. Det var også noe uklarerheter omkring hva som skal dokumenteres i saksnotater utenfor straffesakssystemet og hva som skal inn i den datatekniske rapporten.

3.4.6 Kompetansebehov

Kompetanse er ikke bare en essensiell faktor for å oppnå kvalitet i datatekniske undersøkelser, men også for å kunne fagfellevurdere slike undersøkelser med kritisk blick og gi relevante tilbakemeldinger.

Det kan være ubehagelig å få kritikk på arbeid man har utført, og flere DPA-ledere framhevet at teknisk kompetanse ikke nødvendigvis innebærer en kompetanse i å gi tilbakemelding med fokus på læring. Å etablere et system for kvalitetskontroll vil derfor innebære nye kunnskapsbehov slik at en god tilbakemeldingskultur kan bygges og vedlikeholdes. Flere DPA-ledere trakk fram eksempler i tidligere forsøk på kvalitetskontroll hvor tilbakemeldingene ikke har ført fram, men heller ledet til frustrasjon og tap av motivasjon. En DPA-leder pekte på at det er viktig at rapportskriveren forstår hva som må til for å oppnå riktig kvalitet, og at kriteriene er forankret i et omforent rammeverk framfor en fagfelles egne personlige preferanser. Standardisering og sjekklister spiller derfor en stor rolle for forutsigbarhet og tydelighet i vurdering og formidling av hva som er godt nok og hva som kan bli bedre i en rapport. På spørsmål svarte flere at kvalitetskontroll gjerne kunne være en del av et spesialisert OÅO-opplegg for dataetterforskere.

Flere av DPA-lederne pekte på behovet for økt digital kompetanse blant påtalefaglige og etterforskningsfaglige ledere, fagkontakter og etterforskere.

Behovet for en tydeliggjøring og nivådeling av hvilke oppgaver den enkelte kan utføre, og hvilken kompetanse som kreves ble understreket fra flere. En DPA-leder viste til at man fortsatt kunne oppleve at brukerferdigheter, f.eks. i Snapchat, feilaktig ble likestilt med digital kompetanse.

Per i dag finnes ingen vedtatt standard for hvordan digital kriminalteknikk skal gjennomføres i norsk politi. I prinsippet kan derfor 'alle gjøre alt', og det er opp til den enkelte å vurdere om hen har tilstrekkelig kompetanse til oppdraget. Kompetanse og erfaring varierer også blant DPA-medarbeiderne. En DPA-medarbeider pekte på at det burde stilles sterkere krav også til denne rollen:

«Vi har jo på en måte også ønsket å ha kanskje mer krav til den rollen som vi har da. At man har kanskje utdanningskrav eller sertifiseringskrav. For nå, de fleste som jobber her har tatt ganske mye NCFI utdanninger blant annet og sånn, men vi hadde egentlig ikke trengt det, vi kunne jo latt være å ta noe som helst form for utdanning og likevel jobbet her». (M8)

En DPA-leder uttalte at økt digital forståelse i alle ledd forhåpentligvis ville føre til at flere innså sin egen begrensning, og etterspurte kvalitetssikring av resultater i større grad enn i dag. DPA-lederne fortalte at deres medarbeidere (inkludert dem selv) brukte mye tid på å bidra med kompetansehevede tiltak internt i politidistriktet.

Flere deltakere understreket at formalkompetanse alene ikke er nok, og at den må vedlikeholdes gjennom praktisk erfaring med datatekniske undersøkelser. Dersom kompetansen ikke blir brukt blir den fort utdatert. Flere DPA-ledere uttrykte stor bekymringen for vedlikehold av kompetanse hos fagkontaktene, som gjerne fikk noe opplæring, men kun sporadisk fikk benyttet kompetansen til sikring av databeslag. En DPA-leder fortalte om fagkontakter i eget distrikt som ikke hadde sikret mobiler på et halvt år, og at man i slike tilfeller ikke kunne regne med at det satt i fingrene når noe skulle sikres. Digital kompetanse handler altså ikke kun om formalkompetanse, men også kompetanse som vedlikeholdes og utvikles gjennom praksis og erfaringslæring.

Videre ble behov for *rapportskrivingskompetanse* framhevet av flere deltakere. DPA-medarbeidere med politibakgrunn har rapportskrivning som en av sine kjernekompetanser fra politiutdanningen, men de fleste politidistrikt har også ansatt DPA-medarbeidere med sivil bakgrunn. Selv om disse gjerne kommer med teknisk kompetanse i bagasjen, har de som regel aldri dokumentert undersøkelser og funn som skal brukes som bevis i straffesak. DPA-medarbeidere med sivil bakgrunn har derfor et behov for en innføring i hvordan datatekniske rapporter i straffesaks bør skrives for å bidra til en formålsstyrt etterforskning som holder tilstrekkelig kvalitet.

4. Diskusjon

Følgende spørsmål har vært førende for diskusjonen:

Kan fagfellevurdering og kvalitetskontroll i henhold til kvalitetskontrollhierarkiets nivå 4 – Innholdskontroll

- bidra til å heve kvaliteten på datatekniske rapporter gjennom å avdekke kritiske feil og/eller villedende informasjon?
- skape læring hos fagfelle og/eller rapportskriver (mottaker av fagfellevurdering)?
- lede til systemforbedring/organisasjonslæring?
- gi uønskede sideeffekter?

I tillegg vil vi diskutere noen tema som ble trukket fram av deltakerne i studien som ikke dekkes direkte av disse spørsmålene.

4.1 Kan fagfellevurdering og kvalitetskontroll bidra til å heve kvaliteten og avdekke kritiske feil og/eller villedende informasjon?

Deltakernes fortellinger om feil og nesten-feil har gitt innsikt i hvilke feil som kan oppstå i etterforskninger som bygger på eller innebefatter digitale spor og bevis. Dette viser at feil kan oppstå i alle deler av prosessen, som under ransaking og beslag av databærere på åstedet, ved sikring av innhold fra databærerne, ved analyse av innholdet og under rapportskrivningen. Feilene har noen ganger fått alvorlige følger - som pågrepelse og varetekt av uskyldige personer eller uriktige domfellelser. Beskrivelsene av feil og nesten-feil underbygger behovet for tiltak som kan avdekke slike feil før de får følger for etterforskningen og partene som denne retter seg mot.

Før Kvalitetskontrollprosjektet har man nokså ensidig satset på forebygging av feil gjennom formalkompetanse. Dette er et essensielt

virkemiddel for å oppnå kvalitet og unngå feil, men er alene ikke nok til å hindre at feil skjer i dataetterforskningsprosessen, eller at feilaktig eller villedende informasjon blir formidlet i datatekniske rapporter. Fagfellevurdering vil derfor være et viktig tiltak for å avdekke feil som selv den mest kompetente medarbeider kan komme til å gjøre.

Fire øyne prinsippet, som fagfellevurdering kan sies å være, er implementert i mange sammenhenger også utenfor politiet, for eksempel når banker utsteder kreditt, ved korrupsjonsforebygging og i pasientsikkerhetsarbeid (se for eksempel DNB-konsernet, 2020 s. 49; Schickora 2011; Kobler & Schwappach 2017 s. 7). Felles for disse er at fire øyne blir implementert når kritiske vurderinger og beslutninger skal gjennomføres. Page et al. (2019, s. 8) framhever også *dual investigator* som et mulig kvalitetssikringstiltak for å øke sjansen for å oppdage og unngå feil. Å alltid konsultere en kollega for kritiske innspill til ens egen vurdering er framhevet av Fahsing (2016, s. 20) og bygget inn som et av seks kritiske steg i etterforskningens sykliske problemløsningsprosess.

Hvor effektiv fagfellevurdering og kvalitetskontroll vil være for å avdekke kritiske feil er vanskelig å si på bakgrunn av erfaringer fra Kvalitetskontrollprosjektet alene. Vi kan konstatere at både kritiske feil (røde trafikklys) og behov for endring og justering ble avdekket ved fagfellevurdering av rapportene. Vi vet imidlertid ikke om sakene inneholdt flere feil som ville blitt avdekket gjennom kvalitetskontroll på høyere nivå, for eksempel nivå 5 - *Verifisering av utvalgte spor*, eller nivå 6 - *Verifisering av alle spor*.

Samtidig er det viktig å anerkjenne at ingen metoder kan garantere at feil ikke vil skje, eller at alle feil blir avdekket. Dette henger sammen med at årsakene til at feil oppstår kan være mange og sammensatte. Som påpekt av professor i Safety Science, Sidney Dekker (2014), bør man framfor å konkludere med menneskelige feil heller betrakte dem som *konsekvenser* av systemiske årsaker. I dette tilfellet kan det se ut som at manglende strategisk styring av DPA-arbeidet over lang tid,

fravær av kvalitetssystemer og systematisk kvalitetsarbeid, utydelige rolledefinisjoner og manglende kompetansekrav til personale som håndterer digitale spor og bevis, kan ha ledet til at slike feil som beskrevet i kapittel 3.3 har fått anledning til å skje. Dette styrkes av fortellingen om typiske feil og de manglende rutineene for varsling og håndtering av feil, som viser at feil får anledning til å skje gjentatte ganger uten at organisasjonen evner å ta lærdom og gjøre tiltak for å forhindre at feilen skal skje igjen.

Som beskrevet over viser resultatene at samtlige fagfeller fant forbedringspunkter i rapportene, og en del fant også det de mener er kritiske feil. Samtidig er det klart at kvalitetskontroll på nivå 4, Innholdskontroll, hvor kun rapporten er gjenstand for kritisk vurdering, har sine klare begrensninger med hensyn til å oppdage feil. Kvalitetskontroll opp til dette nivået vil kunne bedre klarheten og tilgjengeligheten av rapportens budskap ved fokus på språklige og presentasjonsmessige aspekter. Den vil styrke transparens og etterprøvbarehet gjennom å rette søkelyset mot beskrivelse av mandat, anvendte metoder og verktøy, og presentasjon av funn. Samtidig er fagfelle prisgitt rapportskrivernes beskrivelser, siden kontroll på dette nivået ikke innebærer å ettergå de konkrete funnene ut over det som står beskrevet i rapporten. Det betyr at kontrollen ikke vil bidra til en verifisering av funnene, og heller ingen 'second opinion' av at de er tolket og beskrevet korrekt ut ifra konteksten hvor de er funnet.

PhD-avhandlingen til N. Sunde (2022a) undersøkte blant annet om dataetterforskeres vurderinger og beslutninger ble påvirket av informasjon uten relevans for analysen de skulle gjennomføre, for eksempel informasjon om at mistenkte hadde tilstått forholdet. Forskningen viste at slik informasjon kan påvirke hvilke spor dataetterforsker finner og anser som relevant. Forskningen viste også at selv om dataetterforskere fant de samme sporene, så ble disse tolket og vurdert ulikt, og dermed beskrevet på svært forskjellige måter i dataetterforskernes rapporter. Dette underbygger ikke bare behovet

for kvalitetskontroll av resultatene, men også at kvalitetssikring underveis, bl.a. ved å implementere «fire øyne» prinsippet, er et viktig tiltak for å sørge for at saken blir tilstrekkelig opplyst, og at funnene blir gjenstand for en kritisk vurdering fra flere før de dokumenteres i analyserapporten.

Blind fagfellevurdering anses som et viktig tiltak for å hindre at fagfellens kritiske vurdering påvirkes av kunnskap om rapportskrivers kompetanse og erfaring. Samtidig er det viktig å lytte til deltakernes erfaringer med en ikke-blind fagfellevurdering. Det å kjenne hverandre, ha tillit og kunne diskutere uformelt ble framhevet som viktige momenter for læring og forbedring av framtidige rapporter. En ikke-blind fagfellevurdering internt på DPA-enheten kan også være med å styrke teamets motivasjon for å i større grad ta et felles ansvar for kvaliteten på det som DPA-enheten leverer. Samtidig pekte flere på at blind fagfellevurdering kunne være enklere siden det fjernet de sosiale barrierene man kunne kjenne på når man visste hvem man vurderte.

4.2 Kan fagfellevurdering og kvalitetskontroll skape læring hos fagfelle og/eller mottaker av fagfellevurdering?

Resultatene fra prosjektet viser at rapportskriver kan lære av at andre leser og påpeker forbedringspunkter. Ved hjelp av andres involvering blir man gjort oppmerksom på egen praksis. Rapportskriver blir klar over hva som ikke holder tilstrekkelig kvalitet og bør justeres, men også hva som er bra og bør videreføres. Ved å få andres vurdering vil man dermed kunne vedlikeholde god praksis, og kontinuerlig forbedre det som ikke holder mål.

Resultatene viser videre at fagfellen også opplever læring i dette arbeidet. Flere fagfeller fortalte at de hadde endret hvordan de skrev sine egne rapporter ved at de brukte sjekklisten og støttedokumentet

som referanse, eller på bakgrunn av det de hadde erfart ved fagfelleevaluering av andres rapporter.

Samlet sett viser dette at fagfelleevaluering ikke kun 'koster' ressurser, men at det heller bør ses på som en investering i kompetanse og kvalitet.

4.3 Kan fagfelleevaluering og kvalitetskontroll lede til systemforbedring / organisasjonslæring?

Som beskrevet i punkt 4.2, forteller både fagfelle og rapportskriver om læring. Siden dette utføres innenfor DPA-enheten vil man over tid kunne få innblikk i feil eller mangler som gjentar seg, og derigjennom avdekke kompetansebehov. Oppdagede feil og mangler i metoder og verktøy vil kunne bidra til å avverge feil blant andre som bruker samme verktøy og metoder.

Et sentralt punkt for organisasjonslæring vil være at kunnskap om feil og mangler samles og løftes opp, slik at det ikke kun blir den enkelte dataetterforsker, eller den enkelte DPA-enhet som drar nytte av denne kunnskapen. Å få på plass et system for kvalitetsledelse vil derfor være essensielt systematisk oppfølging og deling av kunnskap.

Mange av deltakerne fortalte om ulike forsøk på å få til en praksis med kvalitetskontroll i enheten hvor de jobbet. Disse tiltakene hadde stort sett vært ildsjel-drevet og basert på frivillig deltakelse. Slike tiltak så ut til å ha begrenset levetid, som følge av at ildsjelen skiftet arbeidssted eller at andre mer presserende oppgaver måtte prioriteres. Ved NC3 i Danmark fagfellevalueres alle tekniske rapporter før utsendelse til de lokale politienhetene (Wenstrup, 2022). I Forensic Science Regulators retningslinjer for England og Wales er det satt krav til at alle kriminaltekniske enheter skal ha en prosedyre for kontroll og fagfelleevaluering (eng. checking and review) av alle rapporter som inneholder kritiske funn (eng. critical findings) (Forensic Science Regulator, 2021, s. 38-39). *Kritiske funn* beskrives slik:

«Typically observations or results that meet one or more of the following criteria:

- a. They have a significant impact on the conclusion reached and the interpretation and opinion provided;
- b. They cannot be repeated or checked in the absence of the exhibit or sample;
- c. They could be interpreted differently.» (Forensic Science Regulator, 2021, s. 102)

Deltakerne i pilotprosjektet er nokså samstemte om at systematisk kvalitetskontroll er nødvendig og ønskelig, og at den bør gjøres obligatorisk for tekniske rapporter.

Hvorvidt organisasjonen evner å lære av systematiske fagfelle-vurderinger og at det faktisk fører til en forbedring på systemnivå vil trolig henge sammen med flere faktorer enn hva dette prosjektet har undersøkt. Resultatene viser at mange kjenner til politiets forbedringssystem Synergi. Likevel gir mange uttrykk for at de er usikre på hva som skal rapporteres der, og om det som rapporteres faktisk blir tatt tak i. Det er også blitt etablert andre uoffisielle kanaler hvor man bl.a. kan varsle andre i DPA-miljøet om feil med metoder og verktøy. Dette har ikke de samme funksjonalitetene som Synergi, og kan ikke sidestilles med et system som ivaretar systematisk erfaringslæring.

Flere av feilene som deltakerne fortalte om handlet om feil i verktøy som ble benyttet i dataetterforskningsprosessen. Synergi ser ikke ut til å ivareta behovet for *varsling* av andre som bruker samme verktøy slik at nye feil kan avverges. Varslingen handler imidlertid ikke bare om å avverge framtidige feil, men også om å sette i gang prosessen med å avdekke og rette opp eventuelle feil som allerede har skjedd. Når feil i verktøy og metoder oppdages, kan verktøyet ha vært i drift i noe tid, og dermed kan feilen ha påvirket straffesaker som er avsluttet. Det finnes per i dag ingen retningslinje eller rutine for hvordan pågående eller avsluttede straffesaker skal identifiseres og undersøkes ved oppdagelse

av verktøy- eller metodefeil. Til sammenligning har NC3 i Danmark etablert en gruppe som systematisk identifiserer og undersøker straffesaker når slike feil oppdages (Wenstrup, 2022).

4.4 Kan fagfellevurdering og kvalitetskontroll gi uønskede sideeffekter?

På grunn av Kvalitetskontrollprosjektets begrensede omfang, er det ikke mulig å si noe om effekter eller sideeffekter ut over læring og avdekking av feil i enkeltsakene som ble gjenstand for fagfellevurdering i prosjektets pilotfase. For å måle effekten av et slikt tiltak må det implementeres og følges opp over tid. Det er imidlertid grunn til å tro, på bakgrunn av de tidlige sporadiske lokale forsøkene på implementering av et system for kvalitetskontroll, at også et slikt tiltak må være forankret i politiledelsen, ressursatt, og rustet med tilstrekkelig kompetanse, for å ha en varig effekt på kvaliteten på datatekniske undersøkelser og rapporter fra slike undersøkelser.

Uten en tydelig forankring i politiledelse, fagforvaltning og et normativt rammeverk er det en risiko for lite kontinuitet i kvalitetskontrollen, og at det oppstår lokale variasjoner i hvordan og hvor grundig den gjennomføres.

Uten tilstrekkelig ressurssetting kan kvalitetskontrollen bli nok en oppgave i en allerede enorm oppgaveportefølje, og utført med fokus på effektivitet framfor kvalitet. Da risikerer man at kvalitetskontrollen blir en skinnprosess, hvor tilliten til sluttproduktet øker, uten at den forventede kontrollen faktisk er gjennomført.

Uten tilstrekkelig kompetanse risikerer man at det bygges en fryktkultur framfor en tilbakemeldingskultur med fokus på læring og kontinuerlig forbedring. Teknisk kompetanse er nødvendig for å oppdage feil og villedende informasjon omkring digitale bevis, men kompetanse i å gi tilbakemelding og veilede andre er essensielt for at tilbakemeldingen skal føre til refleksjon og læring hos rapportskriver.

5. Oppsummering og anbefaling

5.1 Oppsummering

Kvalitetskontrollprosjektet har bidratt med et oppdatert, utvidet og nyansert bilde av hvordan digitale bevis håndteres i straffesakskjeden. Stor etterspørsel etter digitale spor, et høyt effektivitetsfokus kombinert med manglende helhetlig strategisk styring over tid, kan ha ledet til at kvalitet i for stor grad har blitt overskygget av kravet til effektivitet. Kvaliteten til digitale spor og bevis avhenger av at både dataetterforskningsprosessen og rapportskrivningen utføres på en forsvarlig og rettssikker måte. Kvalitetskontrollprosjektet har forsøkt å operasjonalisere kvalitetskravene som bør stilles til datatekniske rapporter i sjekklister, slik at disse kan vurderes og kvalitetskontrolleres av en kompetent fagfelle. Formålet har vært å øke evnen til å avdekke feil og bidra til læring.

Kvalitetskontrollprosjektet har vist at dataetterforskere har både evne og motivasjon til å levere god kvalitet, og at de er opptatt av kontinuerlig læring og forbedring. Prosjektet har videre vist at fagfellevurdering og kvalitetskontroll er et av flere tiltak som kan forebygge og avverge feil relatert til digitale spor og bevis, og dermed bidra til å styrke rettssikkerheten for personer som etterforskes av politiet. Deltakerne i Kvalitetskontrollprosjektet er samstemte om at systematisk fagfellevurdering og kvalitetskontroll av datatekniske rapporter bør vurderes innført som et obligatorisk trinn i dataetterforskningsprosessen, og at dette bør forankres tydelig i politiledelse og fagforvaltning med klare føringer for roller og ansvar. Dette er også i tråd med I. M. Sundes forslag til tiltak i utredningen: *Effektiv, rettssikker og tillitvekkende behandling av databevis* (I. M. Sunde, 2021, s. 97-98).

I faglitteraturen finnes rikelig med beskrivelser av hva som er god praksis innenfor dataetterforskning. Dette gir imidlertid ingen tydelige føringer for hva som er en god undersøkelses- og rapporteringspraksis i norsk etterforskningskontekst, noe som har ført til en fragmentert dataetterforskningsprosess med mange involverte og liten grad av styring, koordinering og kvalitetssikring. Det er derfor behov for en omforent kvalitetsstandard som omfatter behandlingen av digitale spor gjennom hele dataetterforskningsprosessen. En kvalitetsstandard vil tjene som et felles referansepunkt, og vil kunne bidra til harmonisering av praksiser på tvers av politidistriktene og mulighet til å avdekke, følge opp og korrigere avvikende praksiser.

På bakgrunn av kunnskapen fra Kvalitetskontrollprosjektet anbefaler prosjektgruppen følgende kortsiktige og langsiktige tiltak:

5.2 Kortsiktige anbefalinger

- Fagforvaltningsapparatet overtar ansvar for forvaltning av sjekklister og støttedokumentet. De gjøres tilgjengelig på KO:DE, og oppdateres jevnlig i henhold til enhver tids beste praksis.
- Innføre obligatorisk kvalitetskontroll på minimum nivå 4, innholdskontroll, av spesialistenes tekniske rapporter.
- Politidistriktene etablerer lokale rutiner for en systematisk kvalitetskontroll av rapporter om digitale bevis. Rutinen bør som et minimum omfatte generalist (etterforsker/patrulje), fagkontakt og spesialist (dataetterforsker/digital kriminalteknikk), samt hvem som skal ha et ansvar for koordinering og ledelse av kvalitetsarbeidet.
- Politidistriktene jobber kontinuerlig mot å bygge en god tilbakemeldingskultur lokalt.
- Fagforvaltningsapparatet utarbeider en retningslinje og system for håndtering av kritiske feil i metoder og verktøy, inkludert

varslingsrutine, roller og ansvar for oppfølging og undersøkelse av feilens konsekvenser for pågående og avsluttede straffesaker.

- Kvalitetskontroll med tilbakemelding blir en del av DPA-medarbeideres OÅO.

5.3 Langsiktige anbefalinger

- Det utarbeides en nasjonal kvalitetsstandard for hele håndteringskjeden av digitale spor, som inneholder
 - Rolledefinisjoner (hvem kan gjøre hva, og hvilken kompetanse skal de ha)
 - Begrepsavklaringer
 - Prinsipper for behandling av digitale spor med fokus på etikk og rettssikkerhet
 - Beste praksis for innsamling/sikring, tilrettelegging, analyse (innholdsanalyse, teknisk analyse), rapportering - samt kvalitetssikring underveis og kvalitetskontroll av resultatet.
- Det utarbeides nasjonale retningslinjer for kvalitetssikring og kvalitetskontroll av digitale spor og bevis, herunder
 - Rutiner for validering av verktøy og metoder
 - Rutiner for kvalitetssikring underveis i dataetterforskningsprosessen
 - Rutiner for kvalitetskontroll av rapporter som omhandler digitale bevis.
- Det etableres et helhetlig system for kvalitetskontroll på tvers av distrikter, slik at blind fagfellevurdering kan gjennomføres jevnlig. Formålet er harmonisering av undersøkelses og rapporteringspraksis, kalibrering av nivå på fagfellevurderinger og læring.
- Det settes i gang kompetansehevingstiltak

- Utdanning – kvalitetssikring og kvalitetskontroll som en del av 'grunnutdanningen' innenfor digital kriminalteknikk (NCFI Core)
- OÅO rettet mot spesialistnivået
- DPA-medarbeidere tilbys kurs/opplæring i tilbakemelding.

Referanser

Andreassen, L. E., & Andresen, G. (2020). *Live Data Forensics: A quantitative study of the Norwegian Police University College students LDF examinations during their year of practice* [Masteravhandling, University College Dublin]. PIA. <https://hdl.handle.net/11250/2734964>

Bauge, R. K. (2023). *Risks and quality control in Norwegian Police's digital forensics process: The digital forensics examiner's tightrope walk* [Masteravhandling, NTNU]. NTNU Open. <https://hdl.handle.net/11250/3079082>

Bjerkan, L. (2022, 5. juli). Bedriftsleder i Steinkjer ble siktet i alvorlig overgrepssak – nå henlegges saken. *Steinkjer24*. <https://www.steinkjer24.no/nyheter/i/10Em2X/bedriftsleder-i-steinkjer-ble-siktet-i-alvorlig-overgrepssak-naa-henlegges-saken>

Borhaug, T. S. (2019). *The Paradox of automation in digital forensics* [Masteravhandling, NTNU]. NTNU Open. <http://hdl.handle.net/11250/2617753>

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Cervantes Mori, M. D., Kävrestad, J. & Nohlberg, M. (2021). *Success factors and challenges in digital forensics for law enforcement in Sweden*. Presentert ved 7th International Workshop on Socio-Technical Perspective in IS development, Trento, Italy. <https://www.diva-portal.org/smash/get/diva2:1622611/FULLTEXT01.pdf>

Dekker, S. (2014). *The field guide to understanding 'human error'*. 3rd edition. CRC Press.

DNB-konsernet. (2020). *Risiko- og kapitalstyring Redegjørelse i henhold til pilar 3, 2020*. DNB. https://www.ir.dnb.no/sites/default/files/DNB_ASA_Risiko_-_og_kapitalstyring_-_pilar_3.pdf

Erlandsen, T. E. (2019). *Fallacies when evaluating digital evidence among prosecutors in the Norwegian Police Service* [Masteravhandling, NTNU]. NTNU Open. <http://hdl.handle.net/11250/2617771>

Fahsing, I. A. (2016). *The making of an expert detective. Thinking and deciding in criminal investigations* [PhD-avhandling, Göteborgs Universitet]. PIA. <http://hdl.handle.net/11250/2428006>

Forensic Science Regulator. (2021). *Codes of practice and conduct for forensic science providers and practitioners in the criminal justice system. FSR-C-100. Issue 6*. https://assets.publishing.service.gov.uk/media/602a9a408fa8f503859508ea/FSR-C-100_Codes_of_Practice_and_Conduct_-_Issue_6.pdf

Friheim, I. (2016). *Practical use of dual tool verification in computer forensics* [Masteravhandling, University College Dublin]. Hentet fra: <https://www.researchgate.net/profile/Ivar-Friheim/publication/324784023-Practical-use-of-dual-tool-verification-in-computer-forensics/links/5ae23950aca272fdaf8fae54/Practical-use-of-dual-tool-verification-in-computer-forensics.pdf>

Gjerde, M. (2007). *Victims of success? Knowledge discovery amongst digital forensic investigators in the Norwegian police districts* [Masteravhandling, University College Dublin].

Grut, S. (2020, 3. november). Feil i analyseverktøy gjør at politiet må gjennomgå minst 57 straffesaker. *NRK Beta*. <https://nrkbeta.no/2020/11/03/feil-i-analyseverktoy-gjor-atpolitiet-ma-undersoke-flere-titalls-straffesaker-pa-nytt/>

Haraldseid, S. (2021). *Kan du stikke opp og gå gjennom databeslaget? Framgangsmåter for innholdsanalyse av databeslag og behovet for*

metodisk støtte [Masteravhandling, Politihøgskolen]. PIA.

<https://hdl.handle.net/11250/2760497>

Heitmann, O. (2019). *Digital investigation: The malnourished child in the Norwegian police family?* [Masteravhandling, NTNU]. NTNU Open.

<http://hdl.handle.net/11250/2617759>

Helm, R. (2022). Wrongful conviction in England and Wales: an assessment of successful appeals and key contributors. *Wrongful Conviction Law Review*, 3, 196–217. <https://doi.org/10.29173/wclawr79>

Horsman, G., Sunde, N. (2020). Part 1: The need for peer review in digital forensics. *Forensic Science International: Digital Investigation*, 35, 301062. <https://doi.org/10.1016/j.fsidi.2020.301062>.

House of Commons. (2022). *Post Office and Horizon - compensation: Interim report, eight report of session 2021-22*. UK Parliament, House of Commons. HC 1129, 17. februar, 2022. <https://publications.parliament.uk/pa/cm5802/cmselect/cmbeis/1129/report.html#heading-0>

Hultgren, G. (2018, 5. juli) Sove-voldtekt. Frikjent etter SMS-bevis. *Dagbladet*. <https://www.dagbladet.no/nyheter/frikjent-etter-sms-bevis/69987786>

Jahren, J. H. (2020). *Is the quality assurance in digital forensic work in the Norwegian Police adequate?* [Masteravhandling, NTNU]. NTNU Open. <https://hdl.handle.net/11250/2781174>

Justitsministeriet (2021) Besvarelse av spørsmål nr. 34 (Alm. Del) fra Folketingets Retsutvalg. 10. november 2021. <https://www.ft.dk/samling/20211/almdel/reu/spm/34/svar/1827310/2477481.pdf>

Kjær, J. S. (2019, 18. oktober). Telemaster med skæve data kan have vildledt politiet i Emilie Meng-sag. *Politiken, Inland*. <https://politiken.dk/indland/art7448590/Telemaster-med-sk%C3%A6ve-data-kan-have-vildledt-politiet-i-Emilie-Meng-sag>

- Kobler, I. & Schwappach, D. (2017). *More patient safety by design: Systemic solutions for hospitals. Foundation for patient safety*. https://www.hkb.bfh.ch/dam/jcr:bf0b80c0-df96-499c-8847-dcc19ee7e956/KD_Brosch%C3%BCre_Patientensicherheit_ENG.pdf
- Lentz, L. W. & Sunde, N. (2021). The use of historical call data records as evidence in the criminal justice system - lessons learned from the Danish telecom scandal. *Digital Evidence and Electronic Signature Law Review*, 18. <https://doi.org/10.14296/deeslr.v18i0.5235>
- Nokes, K. & Moorhead, R. (2023, 9. august). The Post Office scandal is possibly the largest miscarriage of justice in UK history – and it’s not over yet. *The Conversation*. <https://theconversation.com/the-post-office-scandal-is-possibly-the-largest-miscarriage-of-justice-in-uk-history-and-its-not-over-yet-211217>
- Page, H., Horsman, G., Sarna, A., & Foster, J. (2018). A review of quality procedures in the UK forensic sciences: What can the field of digital forensics learn? *Science & Justice*, 59(1), 83–92. <https://doi.org/10.1016/j.scijus.2018.09.006>
- Peachy, K. (2022, 22. mars). Post Office scandal: What the Horizon saga is all about. *BBC, News*. <https://www.bbc.com/news/business-56718036>
- Post Office Corporate. (2023). *Overtured convictions and compensation: Information on progress*. <https://corporate.postoffice.co.uk/en/horizon-scandal-pages/overtured-convictions-and-compensation-information-on-progress>
- Reedy, P. (2020). Interpol review of digital evidence 2016-2019. *Forensic Science International: Synergy*, 2, 489–520. <https://doi.org/10.1016/j.fsisyn.2020.01.015>

Reedy, P. (2023). Interpol review of digital evidence for 2019–2022. *Forensic Science International: Synergy*, 6, 100313. <https://doi.org/10.1016/j.fsisyn.2022.100313>

Riksadvokaten. (2021) *Feil i sikringsverktøy til bruk ved uthenting av informasjon fra IOSenheter - resultater fra gjennomgang av aktuelle saker*. Brev av 18.02.2021. <https://www.riksadvokaten.no/wp-content/uploads/2021/03/Cellebrite-gjennomgang-180221.pdf>

Riksrevisjonen. (2021). *Riksrevisjonens undersøkelse av politiets innsats mot IKT-kriminalitet*, rapport av 2. februar 2021, Dokument 3:5 (2020-2021).

Schickora, J. T. (2011). Bringing the four-eyes-principle to the lab. *Discussion Papers in Economics 2011-3*. <https://doi.org/10.5282/ubm/epub.12160>

Stoykova, R., Andersen, S., Franke, K. & Axelsson, S. (2022). Reliability assessment of digital forensic investigations in the Norwegian police. *Forensic Science International: Digital Investigation*, 40, 301351. <https://doi.org/10.1016/j.fsidi.2022.301351>

Stølan, I. B. (2022, 12. juli) I fjor ble Erik bedt om å brenne sakte i helvete. Den siste uka har han fått helt andre meldinger. *Steinkjer24*. <https://www.steinkjer24.no/nyheter/i/Or41lq/i-fjor-ble-erik-bedt-om-aa-brenne-sakte-i-helvete-den-siste-uka-har-han-faatt-helt-andre-meldinger>

Sunde, I. M. (2021). *Effektiv, tillitvekkende og rettssikker behandling av databevis: En straffeprosessuell utredning om ransaking, sikring og beslag i data*. Avgitt til Justis- og beredskapsdepartementet 18. Juni 2021. PIA. <https://hdl.handle.net/11250/2762721>

Sunde, N. (2017). *Non-technical sources of errors when handling digital evidence within a criminal investigation* [Masteravhandling, NTNU]. NTNU Open. <http://hdl.handle.net/11250/2450280>

Sunde, N. (2021). What does a digital forensics opinion look like? A comparative study of digital forensics and forensic science reporting practices. *Science & Justice*, 61(5) 586-596. <https://doi.org/10.1016/j.scijus.2021.06.010>

Sunde, N. (2022a). *Constructing digital evidence. A study on how cognitive and human factors affect digital evidence* [PhD avhandling, Universitetet i Oslo]. DUO vitenarkiv. <https://www.duo.uio.no/bitstream/handle/10852/97851/PhD-Sunde-2022.pdf>

Sunde, N. (2022b). Strategies for safeguarding examiner objectivity and evidence reliability during digital forensic investigations. *Forensic Science International: Digital Investigation*, 40, 301317. <https://doi.org/10.1016/j.fsidi.2021.301317>.

Sunde, N. (2022c). Unpacking the evidence elasticity of digital traces. *Cogent Social Sciences* 8(1). <http://dx.doi.org/10.1080/23311886.2022.2103946>.

Sunde, N. & Dror, I. E. (2021). A Hierarchy of Expert Performance (HEP) applied to digital forensics: Reliability and biasability in digital forensics decision making. *Forensic Science International: Digital Investigation*, 37, 301175. <https://doi.org/10.1016/j.fsidi.2021.301175>.

Sunde, N. & Horsman, G. (2020). Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations. *Forensic Science International: Digital Investigation*, 36, 301074. <https://doi.org/10.1016/j.fsidi.2020.301074>

Søe, C-E. (2019, 4. oktober). Medarbejdere i Rigspolitiet blev bedt om at lyve i sag om teledata. *TV2, nyheder*. <https://nyheder.tv2.dk/politik/2019-10-04-medarbejdere-i-rigspolitiet-blev-bedt-om-at-lyve-i-sag-om-teledata>

Wenstrup, L. (2022, oktober). *Kvalitetssikring af erklæringer. Erfaringer fra Danmark*. Foredrag ved Kvalitetskontrollprosjektets Workshop II, Politihøgskolen, 26.10.2022.

Vincze, E. A. (2016). Challenges in digital forensics. *Police Practice and Research*, 17(2), 183–194. <https://doi.org/10.1080/15614263.2015.1128163>

Vedlegg

1. Sjekkliste for kvalitetskontrollhierarkiet, v. 1.2
2. Støttedokument til kvalitetskontrollhierarkiet, v.1.1
3. Sunde, N. (2022). Fagfellevurdering av rapporter fra datatekniske undersøkelser. (Sammendrag av artiklene: Horsman & Sunde, 2020; Sunde & Horsman, 2021).
4. Refleksjonslogg, fagfelle
5. Refleksjonslogg, leder
6. Informasjonsskriv med samtykkeerklæring

Sjekkliste for Kvalitetskontrollhierarkiet

Sjekklisten er utarbeidet i henhold til nivåene i Kvalitetskontrollhierarkiet (eng: Peer Review Hierarchy for Digital Forensics (Horsman og Sunde, 2020; Sunde og Horsman, 2021)). Dette rammeverket er utarbeidet med henblikk på fagfelleevaluering og kvalitetskontroll av rapporter som omhandler resultatet fra analysefasen i dataetterforskningsprosessen. Analysefasen forstås her som fasen hvor gjennomgang/ransaking av datamateriale skjer med formål om å finne relevant informasjon som belyser straffesakens bevisstema og øvrige informasjonsbehov, samt tekniske undersøkelser for å belyse påliteligheten (usikkerheten) av slik informasjon. Sjekklisten er justert og noe utvidet sammenlignet med opprinnelig versjon bl.a. etter innspill fra deltakere fra politidistrikt og særorgan på Workshop I og II i Kvalitetskontrollprosjektet høsten 2022.



Veiledning for fargekoding:

	Rapporten inneholder kritiske feil eller villedende informasjon som må rettes opp
	Rapporten inneholder uklarheter, mindre feil eller avvik i henhold til gjeldende maler, rutiner, prinsipper og metoder som bør rettes opp
	Rapporten er vurdert til å holde forsvarlig og tilstrekkelig kvalitet (nivå 5-6: funnene er verifisert)
	Ikke vurdert eller ikke relevant

Nivå 1 – Administrativ kontroll

Kriteria:				Kommentar:
1. Er mandatet/oppdraget og formålet (inkludert omfang og avgrensninger) nøyaktig beskrevet i rapporten?				
2. Svarer rapporten på mandatet/oppdraget og formålet med undersøkelsen i tilstrekkelig grad?				
3. Framgår det tydelig hvem som har gjennomført undersøkelsen? Dersom flere har vært involvert – beskrives det tydelig hvilke oppgaver den enkelte har utført?				
4. Er saksinformasjonen som rapportskriver hadde tilgjengelig før og under undersøkelsen beskrevet?				
5. Er egen kompetanse med relevans for det utførte oppdraget beskrevet?				
6. Er rapporten skrevet i tråd med gjeldende mal eller rapportstruktur (sentral/lokal)?				
7. Er datagrunnlaget nøyaktig beskrevet, og er det redegjort for datagrunnlagets autenticitet?				
8. Er verktøy med versjonsnummer oppgitt, og framgår det tydelig til hvilke deler av undersøkelsen de enkelte verktøy er brukt?				
9. Er rapporten komplett – dvs. inneholder den alle oppgitte sider og vedlegg?				
10. Er dokumentasjonen i tråd med gjeldende rutine eller standard?				

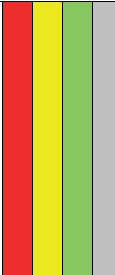
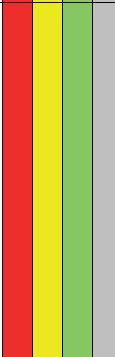
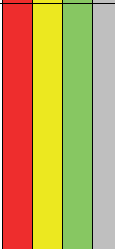
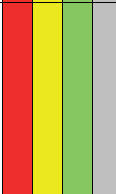
11. Er tidsrommet for undersøkelsen beskrevet?					
--	--	--	--	--	--

Nivå 2 - Språkvask

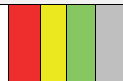
Kriteria:					Kommentar:
1. Inneholder rapporten skrivefeil som bør rettes opp?					
2. Inneholder rapporten grammatiske feil som bør rettes opp?					
3. Dersom forkortelser er brukt, blir de forklart første gang, og blir de så brukt på en lik måte gjennom rapporten?					

Nivå 3 - Kontroll av klarhet og tilgjengelighet

Kriteria:					Kommentar:
1. Er språket i rapporten forståelig?					
2. Er tekniske ord og uttrykk forklart?					
3. Gjør rapportskriver det klart om rapporten inneholder en eller flere av disse rapporteringsformene: <ul style="list-style-type: none"> - Beskrivelser av funn (deskriptiv) - Fortolkninger av hva funnene er, og hva de betyr - Vurdering (foreløpig eller endelig) av funnernes bevisverdi 					
4. Beskrives det tydelig om rapporten presenterer foreløpige eller endelige resultater?					
5. Dersom rapporten inneholder <u>tidsstempler</u> :					

<ul style="list-style-type: none"> - Oppgis tidssone for alle tidsstempler? - Er tidsstemplene konvertert til norsk tid? - Har tidsstempler konsistent format gjennom hele rapporten? 		
<p>6. Dersom rapporten inneholder <u>tabeller</u>:</p> <ul style="list-style-type: none"> - Er tabellene nummerert? - Har tabellene en oppsummerende beskrivelse? - Har kolonner/rader informative navn? - Er unødvendige kolonner/rader fjernet? 		
<p>7. Dersom rapporten inneholder <u>figurer eller andre visuelle objekter</u>:</p> <ul style="list-style-type: none"> - Er figuren/objektet nummerert? - Har objektet en oppsummerende beskrivelse? 		
<p>8. Dersom rapporten inneholder <u>objekter (f.eks. bilder, skjermtklipp) med tekst</u>:</p> <ul style="list-style-type: none"> - Er teksten gjort tilgjengelig for søk? 		

Nivå 4 – Innholdskontroll

Kriteria:	Kommentar:
Undersøkellesprosessen:	
<p>1.Dersom etterforskningen er hypotesedrevet - er de overordnede hypotesene som</p>	

danner grunnlaget for undersøkelsen beskrevet?				
2. Dersom etterforskningen er hypotesedrevet - er relevante under-hypoteser (aktivitets og kilde nivå) utarbeidet på grunnlag av de overordnede hypotesene eller mandatet/oppdraget?				
Transparens og robusthet i undersøkelsesprosessen:				
3. Er framgangsmåte, verktøy og metodebruk beskrevet så nøyaktig at undersøkelsen kan gjentas av andre?				
4. Er datagrunnlaget tilstrekkelig og fullstendig for å oppnå formålet med undersøkelsen?				
5. Gir undersøkelsene som er utført (observasjoner, testing, eksperimentering) tilstrekkelig grunnlag for resultatet som presenteres og slutningene/konklusjonene som trekkes på bakgrunn av disse?				
6. Er forbehold, usikkerhet og begrensninger med metoder og/eller verktøy beskrevet?				
Presentasjon av resultatet fra undersøkelsen:				
7. Skilles det tydelig mellom rene beskrivelser av funn og fortolkning/evaluering av funn?				
8. Beskrives funnene nøyaktig?				
9. Framgår det tydelig - hvilket beslag funnet knyttes til?				

<p>- hvor i datagrunnlaget funnet peker (logisk) til, slik at det er klart hvor i datagrunnlaget funnet er gjort?</p>					
<p>10. Relateres funnene til konteksten (den innholdsmessige sammenhengen) de er funnet i?</p>					
<p>11. Er negative funn (det som er søkt etter, men ikke funnet) beskrevet?</p>					
<p>12. Er illustrasjoner i form av tabeller, figurer eller andre objekter relevante for formålet?</p>					
<p>13. Dersom undersøkelsen er hypotesedrevet, er resultatet av undersøkelsene beskrevet i lys av minst to konkurrerende hypoteser, derav en som omfatter uskyld/ikke straffbart forhold?</p>					
<p>14. Dersom rapporten inneholder en konklusjon - er resultatene, omstendighetene og resonnementene som konklusjonen bygger på beskrevet? - er det er samsvar mellom konklusjonens styrke og funnene den er basert på? - dersom konklusjonsgrader benyttes, er disse forklart eller henvises til et etablert rammeverk med beskrivelse av de ulike konklusjonsgradene?</p>					
<p>15. Dersom det er gjort en evaluering av funnenes bevisverdi – er dette utført i henhold til en gjeldende standard/rammeverk?</p>					

Jus/etikk:				
16. Er undersøkelsen og rapportskrivningen gjennomført i tråd med gjeldende straffeprosessuelle regler og relevante etiske retningslinjer?				
17. Framgår det av rapporten hvorvidt eller i hvilken grad dataetterforsker kan betraktes som uavhengige med hensyn til sine undersøkelser og dokumentasjon?				
18. Er det søkt aktivt etter informasjon som kan underbygge uskyld eller formildende omstendigheter for mistenkte, og er resultatet beskrevet i rapporten?				

Nivå 5 – Verifisering av utvalgte spor

Kriteria:					Kommentar:
1. Har dataetterforsker gjort alle funn, samt de data funnene baserer seg på, tilgjengelig for verifisering?					
2. Har fagfellen gjort rede for hvordan utvalget av spor er foretatt og begrunnet dette? (tilfeldighetsutvalg, representativt utvalg, andre kriterier)					
3. Har fagfellen gjort rede for hvordan verifiseringen er gjennomført ('dual tool' verifisering, manuell undersøkelse, basert på tidligere erfaring, basert på forskning eller lignende)					
4. Har fagfellen dokumentert verifiseringsprosessen og					

resultatet av denne på en nøyaktig måte?				
--	--	--	--	--

Nivå 6 – Verifisering av alle spor

Kriteria:					Kommentar:
1. Har dataetterforsker gjort alle funn, samt de data funnene baserer seg på, tilgjengelig for verifisering?					
2. Har fagfellen gjort rede for hvordan verifiseringen er gjennomført ('dual tool' verifisering, manuell undersøkelse, basert på tidligere erfaring, basert på forskning eller lignende)					
3. Har fagfellen dokumentert verifiseringsprosessen og resultatet av denne på en nøyaktig måte?					

Oppsummering

Beskrivelse av rapportens styrker:

Beskrivelse av rapportens utfordringer og forbedringspunkter:

Andre kommentarer:

Synergi

Beskrivelse av forhold som skal meldes inn, og hvem som foretar dette:

Versjonslogg	
1.0 september 2020	«PARS Peer Review Hierarchy template» Opprinnelig versjon presentert i artikkelen: Sunde, N., Horsman, G. (2020). Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations. <i>Forensic Science International: Digital Investigation</i> , 36, 301074. https://doi.org/10.1016/j.fsidi.2020.301074
1.1 september 2022	«Sjekkliste for kvalitetskontrollhierarkiet», oversatt og bearbeidet versjon til Workshop I, høsten 2022
1.2 desember 2022	«Sjekkliste for kvalitetskontrollhierarkiet», bearbeidet etter innspill fra deltakere på Workshop II, høsten 2022

Støttedokument til

Sjekkliste for Kvalitetskontrollhierarkiet

Dette dokumentet gir en utdypende forklaring på hva som skal vurderes under sjekkpunktene, hvorfor, og med illustrerende eksempler der det er relevant. Hvert nivå innledes med en kort beskrivelse av hva en fagfelle vurdering på dette nivået omfatter og hvem som kan utføre kvalitetskontrollen. Kvalitetskontrollen omfatter resultatet av analyse- og presentasjonsfasen av dataetterforskningsprosessen slik den beskrives av Flaglien (2018). Undersøkelsen som gjennomføres er en *ransaking* i juridisk forstand, men i sjekklisten og dette støttedokumentet brukes for enkelthets skyld begrepet 'undersøkelse' om dette arbeidet.

Nivå 1 – Administrativ kontroll

En administrativ kontroll omfatter hvorvidt etterforskningen er gjennomført i tråd med formelle krav og i henhold til oppdraget, altså at avtalte undersøkelser er gjennomført på de beslaglagte enhetene. En administrativ kontroll omfatter *ikke* etterforskningsfaglige vurderinger av funn som presenteres i rapporten eller de vurderinger og slutninger som trekkes på bakgrunn av disse (dette gjøres på nivå 4, Innholdskontroll).

Kan utføres av: Personell uten spesialistkompetanse på dataetterforskning.

<p>1. Er mandatet/oppdraget og formålet (inkludert omfang og avgrensninger) nøyaktig beskrevet i rapporten?</p>
--

<p>Med mandat menes oppgaven eller oppdraget som er gitt dataetterforsker, altså hva dataetterforsker er bedt om å utføre. Rammene eller avgrensningene for oppdraget, altså hvilken tidsramme som er gitt, hva oppdraget/undersøkelsen skal avgrense mot og hva som forventes levert er viktige elementer i denne beskrivelsen. Denne informasjonen er viktig for å få et inntrykk av hva man kan forvente at undersøkelsen omfatter, og hva som ikke er undersøkt.</p>
--

<p>2. Svarer rapporten på mandatet/oppdraget og formålet med undersøkelsen i tilstrekkelig grad?</p>

<p>Her sammenlignes resultatet fra undersøkelsen med det opprinnelige oppdraget eller mandatet for å se om rapporten svarer tilstrekkelig konkret, detaljert og fullstendig på oppdraget som ble gitt, og om informasjon som går ut over oppdraget eller mandatet er utelatt.</p>

<p>NB: Informasjon som ikke omfattes av oppdraget eller mandatet bør i utgangspunktet ikke inngå i rapporten, men kan likevel være av betydning for etterforskningen, for eksempel ved at det kan danne grunnlag for andre etterforskningskritt. Etterforskningsledelsen bør derfor gjøres kjent med informasjonen, og vurdere og beslutte hvordan den skal dokumenteres og følges opp.</p>

<p>3. Framgår det tydelig hvem som har gjennomført undersøkelsen? Dersom flere har vært involvert – beskrives det tydelig hvilke oppgaver den enkelte har utført?</p>
--

<p>Dersom én person har utført undersøkelsen, er det samme person som skriver rapport og signerer, og det er dermed ingen behov for presisering av hvilke deloppdrag vedkommende har utført. Dersom flere har vært involvert, f.eks. at en person har utført prosesseringen, og den tekniske</p>
--

analysen, mens en annen har gjort ransaking/gjennomgang av innhold – og det utarbeides en felles rapport, må det framgå hvem som har gjort hva. Dette har bl.a. betydning for hvem som skal vitne om undersøkelsene i retten, og for vurderingen av om den enkelte har hatt riktig kompetanse til å utføre delopdraget.

4. Er saksinformasjonen som rapportskriver hadde tilgjengelig før og under undersøkelsen beskrevet?

Kunnskap om saken kan være viktig for å kunne gjøre effektive og målrettede undersøkelser av databeslaget. Slik kunnskap vil være en pekepinn på om informasjonen man finner er relevant eller ikke for saken som etterforskes. Samtidig kan informasjon om saken påvirke den som gjennomfører undersøkelsen på en uønsket måte, dvs. at informasjonen kan føre til bekreftelsestendens og systematiske skjevheter (bias) i hva man anser som relevant, hvordan man fortolker informasjonen man kommer over og hvordan man trekker slutninger (Sunde og Dror, 2021). For å kunne fange opp slike tendenser er det viktig at man gjør rede for hva man visste når undersøkelsene ble gjennomført.

Eksempel:

- *Jeg har vært innlemmet i etterforskningsteamet, og har hatt tilgang til sakens dokumenter.*
- *Jeg var med å gjennomføre ransakingen av mistenktes bopel, hvor mistenkte også ble pågrepet.*
- *Jeg leste avhørene av fornærmede og mistenkte i forkant av undersøkelsen, siden jeg ble anmodet om å søke etter spor som kunne bekrefte eller avkrefte deres forklaringer.*
- *Jeg har ingen kjennskap til etterforskningen ut over informasjonen som er beskrevet i mandatet for undersøkelsen.*

5. Er egen kompetanse med relevans for det utførte oppdraget beskrevet?

Kompetanse omfatter både formalkompetanse (dokumentert kompetanse gjennom kurs, utdanning, sertifisering) og reell kompetanse (erfaring, ferdigheter). Kompetansen er viktig i alle aspekter av dataetterforskningsprosessen, også analysefasen – og har betydning blant annet for evnen til å

- kunne forutsi hvor relevante spor kan befinne seg
- kunne gjøre en kvalifisert vurdering av hvilke verktøy og metoder som er best egnet for å søke etter de aktuelle sporene, samt forstå begrensningen i verktøyene og metodene
- forstå hva de enkelte dataspor er og betyr, og kunne verifisere disse på en tilstrekkelig måte
- kunne undersøke og forstå hvilken usikkerhet som er knyttet til de ulike sporene og til å kommunisere dette

Dokumentasjon av kompetanse er et krav i flere standarder og beste praksis dokumenter som retter seg mot datatekniske undersøkelser, bl.a. ENFSI (2015b) Appendix G., Council of Europe (2017) Appendix J.

Eksempel:

- *Jeg har grunnleggende kompetanse innenfor datatekniske undersøkelser fra NCFI Core (15 SP), og Advanced Computer Forensics (7,5 SP), og 4 års erfaring som spesialletterforsker innenfor digitalt politiarbeid. Jeg anser denne kompetansen som relevant og tilstrekkelig for å utføre sikring og undersøkelse av den beslaglagte bærbare datamaskinen og minnepinnene i straffesaken.*

6. Er rapporten skrevet i tråd med gjeldende mal eller rapportstruktur (sentral/lokal)?

Dersom det finnes en nasjonalt vedtatt rapportstruktur, er det denne som primært skal brukes i vurderingen. I fravær av dette vil lokal mal eller rapportstruktur være relevant å benytte.

7. Er datagrunnlaget nøyaktig beskrevet, og er det redegjort for datagrunnlagets autentisitet?

Det må ikke være tvil om hvilke beslaglagte enheter undersøkelsene omfatter. Her angis både type enhet og en unik identifikator (løpenummer/lokalt beslagsnummer). Datagrunnlagets autentisitet har stor betydning for resultatenes bevisverdi, og dokumentasjon på dette bør inntas i rapporten f.eks. datagrunnlagets sjekksum (MD5, SHA-2, SHA-256).

Eksempel:

- *Undersøkelsen omfatter data fra beslag nr. XXXXX, Minnepenn Data Traveller (3GB). Ifølge sikringsrapporten ble enheten sikret slik at autentisiteten til datagrunnlaget ble ivaretatt. Undersøkelsen ble gjennomført på en identisk kopi av denne sikringsfilen, med sjekksum, SHA-2: xxxxxxxxxxxxxxxx. Etter gjennomført undersøkelse var sjekksummen den samme, som betyr at integriteten til sikringsfilen ble ivaretatt under analysen.*

8. Er verktøy med versjonsnummer oppgitt, og framgår det tydelig hvilke deler av undersøkelsen de enkelte verktøy er brukt?

Det er svært viktig at verktøy og versjonsnummer blir oppgitt. Dette er først og fremst for å kunne etterprøve om verktøyet var egnet til formålet. Dersom det på et senere tidspunkt viser seg at versjonen hadde implementeringsfeil eller 'bugs' som fortolket data uriktig er informasjonen essensiell med hensyn til å kunne gå tilbake og vurdere om feilen har påvirket etterforskningen eller utfallet av saken på en måte som krever ny behandling av saken.

Eksempel:

- *Verktøyet X-Ways 20.6 ble brukt for å undersøke databeslaget. Funnene ble verifisert med Autopsy 4.19.3.*

9. Er rapporten komplett – dvs. inneholder den alle oppgitte sider og vedlegg?

En rapport er ofte flere sider, og kan inneholde mange ulike vedlegg. Her sjekkes det om sidetallet som er oppgitt stemmer, om det er henvist til vedlegg i rapporten, og om alle vedlegg faktisk er lagt ved.

10. Er dokumentasjonen i tråd med gjeldende rutine eller standard?

Per i dag finnes ingen vedtatt kvalitetsstandard for dataetterforskning på nasjonalt nivå. Fagforvaltningsapparatet utvikler rutiner, og når disse finnes vil de være gjeldende på nasjonalt nivå. I fravær av nasjonale rutiner vises det til lokale rutiner. Når dette finnes, kan det henvises til denne, og eventuelle avvik beskrives og begrunnes. Dersom det ikke finnes en vedtatt/gjeldende rutine, må framgangsmåten beskrives i detalj i rapporten.

11. Er tidsrommet for undersøkelsen beskrevet?

I tillegg til å opplyse hvor lang tid undersøkelsene har tatt (start og slutt) kan det i noen tilfeller være aktuelt å kommentere årsaker til at undersøkelsene har trukket ut i tid. Dette kan f.eks. være at politiet har brukt lang tid på å skaffe seg tilgang til en beslaglagt enhet som mistenkte ikke har ønsket å oppgi passord til eller annen informasjon som er nødvendig for tilgang.

Eksempel:

- *Oppdraget ble mottatt xx.xx.xx, og undersøkelsen ble avsluttet xx.xx.xx. Den beslaglagte harddisken var kryptert. Siktete ønsket ikke å oppgi passordet som åpnet krypteringen, og en stor del av denne tiden gikk med til å omgå krypteringen for å kunne ransake innholdet.*

Nivå 2 – Språkvask

Språkvask omfatter en vurdering av om rapporten holder tilstrekkelig språklig kvalitet, og at den ikke/i liten grad inneholder stave- og grammatiske feil.

Kan utføres av: Personell uten spesialistkompetanse på dataetterforskning.

1. Inneholder rapporten skrivefeil som bør rettes opp?

2. Inneholder rapporten grammatiske feil som bør rettes opp?

3. Dersom forkortelser er brukt - blir de forklart første gang, og blir de så brukt på en lik måte gjennom rapporten?

Eksempel:

- *3 stk. Secure Digital (SD) minnekort ble sikret og undersøkt. Det ble ikke funnet informasjon med relevans for etterforskningen på SD minnekortene.*

Forkortelser bør imidlertid brukes med omhu. En rapport med mange forkortelser kan være tung å lese, og det kan være vanskelig for leser å huske hva alle de ulike forkortelsene betyr. Dersom en forkortelse kun brukes noen få ganger kan det være mere leservennlig å skrive ut uttrykket i sin helhet gjennom rapporten.

Nivå 3 - Kontroll av klarhet og tilgjengelighet

Kontroll av klarhet og tilgjengelighet innebærer å vurdere om rapportskriver har framstilt resultatet av undersøkelsen på en tydelig, forståelig og ryddig måte for en leser uten særskilt teknisk kompetanse.

Kan utføres av: Personell uten spesialistkompetanse på dataetterforskning.

1. Er språket i rapporten forståelig?

Det er utfordrende å skrive en forståelig rapport uten at det går på kompromiss med nødvendig faglig presisjon. Rapporten skal leses og forstås av ulike aktører som gjerne har lavere teknisk kompetanse enn rapportskriver, og rapporten må derfor skrives slik at den kan forstås av en 'ikke-teknisk' leser.

2. Er tekniske ord og uttrykk forklart?

Om tekniske ord og uttrykk brukes er det viktig at disse forklares slik at de kan forstås av en 'ikke-teknisk' leser, enten i løpende tekst eller i en ordliste.

3. Gjør rapportskriver det klart om rapporten inneholder en eller flere av disse rapporteringsformene:

- **Beskrivelser av funn (deskriptiv)**
- **Fortolkninger av hva funnene er, og hva de betyr**
- **Vurdering (foreløpig eller endelig) av funnenes bevisverdi**

En 'ikke-teknisk' leser av rapporten uten spesialistkompetanse kan ha en tendens til å oppfatte både funn og fortolkninger av funn som uomtvistelige 'fakta', og kan ha begrenset evne til å forholde seg kritisk til det som presenteres i rapporten. Det må derfor framgå klart av rapporten hva som er funn, og hva som er dataetterforskers fortolkninger/vurderinger av funn. Dette kan gjøres ved å presentere dette under ulike overskrifter, som 'Funn' og 'Fortolkning av funn'.

Å vurdere funnets bevisverdi (sannsynlighet) skal gjøres i lys av minst to konkurrerende hypoteser og de omstendigheter som er lagt til grunn. En slik vurdering gjennomføres på bakgrunn av et konkret oppdrag (f.eks. sakkyndigoppdrag) – og bør følge en etablert framgangsmåte, som ENFSI (2015). ENFSI Guideline for evaluative reporting in forensic science (https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf), slik at både funnens verdi (for eksempel i form av en sannsynlighetsrate) og begrensninger/usikkerhet blir tydelig kommunisert. Dersom oppdraget innebærer å foreta en vurdering av funnens bevisverdi, bør resultatet presenteres under en overskrift som tydeliggjør dette, som 'Vurdering av bevisverdi'.

4. Beskrives det tydelig om rapporten presenterer foreløpige eller endelige resultater?

Noen ganger gjøres avgrensede undersøkelser med et klart formål, f.eks. sjekksum-baserte søk etter overgrepsmateriale i forkant av et avhør eller en fengsling. Slike undersøkelser er ofte ikke fullstendige, men følges gjerne opp med mer omfattende undersøkelser. Dersom dette er tilfelle må det framgå klart av rapporten hva formålet med undersøkelsene var, og hvilke avgrensninger som ble gjort.

5. Dersom rapporten inneholder tidsstempler:

- **Oppgis tidssone for alle tidsstempler?**
- **Er tidsstemplene konvertert til norsk tid?**
- **Har tidsstempler konsistent format gjennom hele rapporten?**

Dersom tidsstemplene er konvertert til norsk tid, bør det også forklares kort hva som er gjort, og hva dette innebærer – slik at det ikke er noen tvil om hvordan tidspunktene relatert til funnene som beskrives skal forstås i straffesaken som etterforskes.

Eksempel:

Tidsstemplene i beslaget var angitt i Pacific Standard Time (PST). De ble konvertert til norsk tid, altså Universal Time Coordinated (UTC) +2 timer for sommertid og +1 timer for vintertid.

Det anbefales at tidsstempler oppgis slik: dd.mm.åååå, tt:mm:ss.

6. Dersom rapporten inneholder tabeller:

- **Er tabellene nummerert?**
- **Har tabellene en oppsummerende beskrivelse?**
- **Har kolonner/rader informative navn?**
- **Er unødvendige kolonner/rader fjernet?**

Tabeller kan bidra til å gjøre informasjonen som presenteres mer tilgjengelig og forståelig. For at en tabell skal være lett å henvise til, f.eks. i retten, bør den nummereres.

Eksempel:

- *Tabell 1: Tabellen viser epost til og fra adressen xxx@gmail.com i tidsrommet 01.09.2022 til 05.09.2022.*

7. Dersom rapporten inneholder figurer eller andre visuelle objekter:

- Er figuren/objektet nummerert?
- Har figuren/objektet en oppsummerende beskrivelse?

Eksempel:

- *Figur 1: Figuren viser et utklipp fra mappestrukturen, slik den framsto i Windows Filutforsker på PC, beslag X.*

8. Dersom det er inkludert objekter (bilder, skjermtklipp) med tekst:

- Er teksten gjort tilgjengelig for søk?

Bruk av skjermtklipp fra verktøy, bilder eller lignende kan inneholde tekst, tidsstempler eller annen informasjon som er relevant for straffesaken. Det tekstlige innholdet i slike objekter er ikke søkbart i andre verktøy som f.eks. Indicia prosjekt. Dersom slike objekter brukes, bør det vurderes å gjøre teksten søkbar på andre måter – f.eks. ved å skrive den inn i rapporten.

Nivå 4 – Innholdskontroll

Innholdskontroll omfatter en grundig kontroll av rapportens innhold som beskriver resultatet av etterforskningen, men avgrenses mot en verifisering av funn/resultater. Hovedfokus rettes mot det vitenskapelige og logiske fundamentet i rapporten. Vurdering av sammenhengen mellom bevisene som presenteres og konklusjonen er spesielt viktig. Prinsippene om balanse, logisk sammenheng, robusthet, transparens og lovlighet/etikk danner rammen for elementene som skal vurderes.

Kan utføres av: Personell med spesialistkompetanse innen dataetterforskning, og helst høyere kompetanse enn rapportskriver på området kontrollen omfatter.

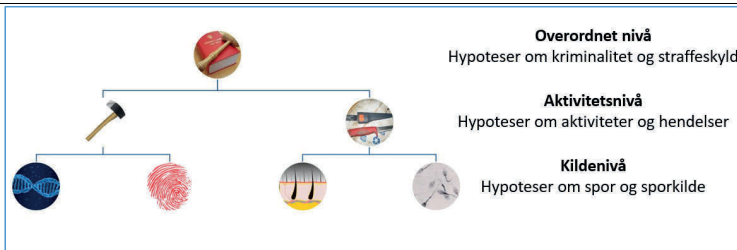
Undersøkellesprosessen:

1. Dersom etterforskningen er hypotesedrevet - er de overordnede hypotesene som danner grunnlaget for undersøkelsen beskrevet?

Dersom mandatet eller oppdraget viser til sakens overordnede hypoteser, altså hypotesene som er beskrevet i etterforskningsplanen/Indicia, bør disse beskrives. Dette er viktig fordi hypotesesettet som danner grunnlag for etterforskningen er dynamisk, og kan endres over tid.

2. Dersom etterforskningen er hypotesedrevet - er relevante under-hypoteser (aktivitets og kildnivå) utarbeidet på grunnlag av de overordnede hypotesene eller mandatet/oppdraget?

En hypotesedrevet etterforskning ved undersøkelser av databeslag vil gjerne rette seg mot hendelser/aktiviteter som er utført på eller ved hjelp av digitale enheter. Den kan også omfatte systemgenererte spor på enheten med relevans for straffesakens bevistema, f.eks. brukerens bevegelser, lokasjoner, nettverksforbindelser osv. Det er derfor hensiktsmessig å 'bryte ned' de overordnede hypotesene i underhypoteser. Innenfor forensiske undersøkelser/kriminalteknikk deles dette gjerne i to nivå.



Aktivitetsnivå handler om aktiviteter og hendelser – samt tid og sted. Med henblikk på etterforskningens grunnspørsmål er dette altså en undersøkelse av «Hva har blitt gjort/ Hva har skjedd» samt «Hvor» og «Når». Aktiviteter kan gi verdifull informasjon om både objektive og subjektive vilkår i straffebudet som etterforskes. Aktiviteter som for eksempel lagring, nedlasting, deling, installasjon av programvare og kjøring av program kan belyse objektive vilkår for straff. Søkeshistorikk, organisering av filstruktur, sletting og kryptering er eksempler på aktiviteter som kan belyse forsettet i den straffbare handlingen.

Eksempel på hypoteser på aktivitetsnivå:

- *Det er delt bilder med andre som viser seksuelle overgrep mot barn fra mistenktes mobil vs. det er ikke delt bilder med andre som viser seksuelle overgrep mot barn fra fornærmedes mobil*

Kildenivå handler om å knytte entiteter til aktivitetene eller hendelsene, f.eks. en brukerkonto på datamaskinen, epost adresse, telefonnummer. Formålet med dette er å komme nærmere et svar på «Hvem har hatt fingrene på tastaturet» ved relevante aktiviteter eller hendelser på datamaskinen. I enkelte tilfeller kan man knytte personer til entitetene på bakgrunn av undersøkelser av databeslaget, men ofte ikke.

Eksempel på hypoteser på kildenivå:

- *Mappene med overgrepstilbildene ble opprettet av brukeren XXXX vs. mappene med overgrepstilbilder ble opprettet av en annen bruker enn XXXX.*

Transparens og robusthet i undersøkelsesprosessen:

3. Er framgangsmåte, verktøy og metodebruk beskrevet så nøyaktig at undersøkelsen kan gjentas av andre?

I tillegg til verktøy og versjon (jf. Nivå 1 punkt 8) er det viktig at framgangsmåten som ledet fram til funnene (eller negative funn) beskrives detaljert.

Vurderingen her handler altså først og fremst om rapporten i tilstrekkelig grad tilfredsstillende kravene til repeterbarhet og reproducerbarhet. Her kan det henvises til etablerte rutiner/prosedyrer eller metoder dersom disse finnes. Eventuelle avvik må da beskrives. Om ikke det finnes etablerte rutiner/prosedyrer eller metoder, må framgangsmåte, verktøy og metodebruk beskrives så nøyaktig at de kan gjentas av andre med samme resultat.

En nøyaktig beskrivelse har også betydning for å kunne vurdere behovet for nye undersøkelser på et senere tidspunkt i etterforskningen. Informasjonsbehovet kan endre seg over tid, og en detaljert

beskrivelse av undersøkelsene vil være til god hjelp for å kunne vurdere hvordan nye undersøkelser/andre fremgangsmåter kan bidra til andre eller flere spor.

4. Er datagrunnlaget tilstrekkelig og fullstendig for å oppnå formålet med undersøkelsen?

En forutsetning for å kunne oppnå formålet med undersøkelsen er å ha et riktig og tilstrekkelig datagrunnlag. Dersom datagrunnlaget er ufullstendig eller ikke inneholder data som treffer formålet kan resultatet bli misvisende.

For eksempel, dersom undersøkelsene har som formål å avdekke kommunikasjon mellom mistenkte og fornærmede, og man kun har innhentet data fra mistenktes mobil – så kan man gå glipp av kommunikasjon som har foregått via applikasjoner på andre enheter som for eksempel mistenktes datamaskin. En slutning om at det ikke er funnet kommunikasjon mellom mistenkte og fornærmede vil dermed kunne villede leseren av rapporten til å tro at slik kommunikasjon ikke finnes – og at man glipper på å sikre relevante enheter/undersøke sporsteder som kunne inneholdt slik kommunikasjon.

5. Gir undersøkelsene som er utført (observasjoner, testing, eksperimentering) tilstrekkelig grunnlag for resultatet som presenteres og slutningene/konklusjonene som trekkes på bakgrunn av disse?

Å skape et robust grunnlag for å trekke slutninger om aktiviteter/hendelser kan være utfordrende. Å trekke slutninger om årsakssammenheng (for eksempel hvordan en fil havnet på en datamaskin) basert på observasjoner av data vil være forbundet med stor usikkerhet, spesielt siden teknologien utvikles og oppdateres kontinuerlig. Noen ganger kan man bygge på forskning eller erfaringer fra tidligere saker, men siden teknologien er i stadig endring kreves gjerne testing eller eksperimentering for å danne et grunnlag for å trekke pålitelige slutninger om resultatet.

6. Er forbehold, usikkerhet og begrensninger med metoder og/eller verktøy beskrevet?

Validering er en viktig mekanisme for å få en objektiv bekreftelse på at metoden eller verktøyet er egnet til formålet, og for å forstå metodens eller verktøyets begrensninger (se f.eks. beskrivelsene i ENFSI 2015b; Forensic Science Regulator, 2020). Validering er derfor et essensielt tiltak for å redusere risikoen for menneskelige eller tekniske feil. Dersom verktøy eller metoder benyttes uten at disse er validert bør dette kommenteres i rapporten.

Presentasjon av resultatet av undersøkelsen:

7. Skilles det tydelig mellom rene beskrivelser av funn og fortolkning/vurdering av funn?

En sammenblanding av beskrivelser av funn, fortolkning av funn og vurdering av funnens sannsynlighet eller betydning kan lett føre til misforståelser senere i etterforskningen, og det er viktig at det skilles tydelig mellom disse rapporteringsformene i rapporten. Fortolkning av funn kan eksempelvis være beskrivelser, basert på rapportskrivners kunnskap og erfaring, av hva et objekt er eller hva det betyr. Dette er ofte en viktig informasjon for etterforskningen, men kan samtidig lett misforstås som et objektivt og pålitelig *fakta* som legges til grunn som sannhet av den som leser rapporten. Et tydelig skille kan kommuniseres gjennom å bruke overskrifter som 'Funn' og 'Fortolkninger av funn'.

8. Beskrives funnene nøyaktig?

En nøyaktig og detaljert beskrivelse av funnet er viktig slik at den som leser rapporten er i stand til å vurdere verdien av det enkelte funnet, og til å sammenstille det med sakens øvrige spor. Dersom utklipp fra analyseprogrammet brukes for å forklare funnet, må det framgå klart at dette ikke er en gjengivelse av hvordan sporet så ut på enheten sporet er sikret fra.

9. Framgår det tydelig

- **hvilket beslag funnet knyttes til?**
- **hvor i datagrunnlaget funnet peker (logisk) til, slik at det er klart hvor i datagrunnlaget funnet er gjort?**

Denne informasjonen er viktig for mistenktes mulighet til kontradiksjon. Tydelig informasjon om hvilke beslag funnene knytter seg til, og hvor på beslaget informasjonen er funnet er også et viktig grunnlag for strafferettslige beslutninger om de ulike beslagene, som f.eks. inndragning, utlevering eller destruksjon. I tillegg er dette nødvendig informasjon med tanke på verifisering av funn.

10. Relateres funnene til konteksten (den innholdsmessige sammenhengen) de er funnet i?

Denne vurderingen handler om det framgår tydelig hvilken sammenheng funnet er gjort. Dersom funnene presenteres isolert og uten relasjon til konteksten kan de framstå som relevante for saken selv om de reelt sett ikke har noen verdi. For eksempel, hos en som er mistenkt for å ha drept sin kone med skytevåpen kan søkene 'rifle' og 'kniv' i mistenktes søkelogg underbygge et drapsforsett. Dersom søkene er gjort i relasjon til planlegging av høstjakta, og er plukket ut av en innholdsmessig sammenheng hvor også søkeordene 'elgjakt', 'tursko', 'kikkertsikte' og 'hjordestek' inngår, er trolig relevansen av denne informasjonen for straffesaken en annen.

11. Er negative funn (det som er søkt etter, men ikke funnet) beskrevet?

Fravær av informasjon kan også være relevant informasjon i straffesaken. Det er derfor viktig å dokumentere såkalte 'negative funn' – altså det som er søkt etter og ikke funnet. Siden fravær av informasjon ikke er ensbetydende med at en aktivitet/hendelse ikke har skjedd, bør vurderingen også omfatte om formuleringene i rapporten går for langt i denne retningen – som vist i eksemplet under.

F.eks.

- *Det finnes ikke overgrepbilder på mistenktes datamaskin.*
- vs
- *Det ble ikke funnet overgrepbilder på mistenktes datamaskin.*

12. Er illustrasjoner i form av tabeller, figurer eller andre objekter relevante for formålet?

Visuelle framstillinger av informasjon kan ha stor påvirkningskraft. Det er derfor viktig at det gjøres en grundig vurdering av hvilken informasjon som er relevant å framstille visuelt, og hvilke virkemidler som benyttes. Her er det også viktig å gjøre en grundig vurdering av hvorvidt den visuelle framstillingen kan skape et skjevt inntrykk av informasjonen eller at den er et resultat av 'cherry picking' – altså kun viser informasjon som støtter en side/hypotese i saken.

13. Dersom undersøkelsen er hypotesedrevet - er resultatet av undersøkelsene beskrevet i lys av minst to konkurrerende hypoteser, derav en som omfatter uskyld/ikke straffbart forhold?

Det kan være fristende å oppsummere resultatet i lys av kun den ene hypotesen. Dette kan imidlertid få resultatet til å framstå som av høyere verdi for etterforskningen enn det faktisk er. En mer balansert framstilling vil oppnås ved å presentere funnene i lys av to konkurrerende hypoteser.

- *Følgende hypoteser ble undersøkt: H1 Det er delt bilder med andre som viser seksuelle overgrep mot barn fra mistenktes mobiltelefon vs. H2 Det er ikke delt bilder med andre som viser seksuelle overgrep mot barn fra fornærmedes mobiltelefon.*
- *Når det gjelder H1, så er det ikke funnet informasjon som tyder på at det er delt bilder som viser seksuelle overgrep mot barn. I forhold til H2, så styrker undersøkelsene denne hypotesen. Det må likevel nevnes at resultatet ikke utelukker at slik deling har skjedd. Analyseprogrammet som ble brukt i undersøkelsene har ikke støtte for kommunikasjonsapplikasjonene XChat og Cryptic som var installert på mistenktes mobiltelefon, og det er derfor uvisst om det er delt informasjon som viser seksuelle overgrep mot barn der.*

14. Dersom rapporten inneholder en konklusjon

- **er resultatene, omstendighetene og resonnementene som konklusjonen bygger på beskrevet?**
- **er det samsvar mellom konklusjonens styrke og funnene den er basert på?**
- **dersom konklusjonsgrader benyttes, er disse forklart eller henvises det til et etablert rammeverk med beskrivelse av de ulike konklusjonsgradene?**

Digitale spor og bevis blir gjerne betraktet som objektive og pålitelige bevis av straffesakens aktører uten spesialistkompetanse. Det er derfor svært viktig at usikkerhet og uvisshet blir kommunisert, slik at ikke sporenes verdi feilaktig blir tillagt for stor vekt – og betraktet som absolutte sannheter. Det finnes ulike rammeverk for å kommunisere graden av sikkerhet/usikkerhet innenfor forensiske undersøkelser, men ikke alle er egnet for digitale spor. Kripos +/- 4 skala for konklusjonsgrader er eksempelvis beregnet på undersøkelser på kilde nivå som innebærer sammenligning, f.eks. av hvorvidt verktøyspor og verktøy eller fottøyavtrykk og fottøy svarer til hverandre.

Fordelen med å henvise til et etablert rammeverk er at man da ikke trenger å forklare uttrykkene man bruker for å beskrive usikkerhet. Det er per i dag ikke lagt noen sentrale føringer for bruk av konklusjonsgrader for digitale bevis – og det som følger nedenfor er derfor eksempler på ulike alternativer.

Eksempler på konklusjonsgrader i dataetterforskning er

Bevisstyrke skalaen (eng: Strenght of evidence scale) (Casey, 2020) for vurdering av funnernes verdi:

Proposed scale for expressing strength of digital evidence for evaluative opinions.

C-Value	Verbal level	Illustrative indicators
C0	Erroneous/incorrect	Evidence contradicts known facts (extreme dissonance of observations in light of the hypothesis).
C1	Extremely weak evidence	Evidence is highly questionable (very strong dissonance of observations in light of the hypothesis).
C2	Very weak evidence	Only one source of evidence that is not difficult to tamper with.
C3	Weak evidence	The source(s) of evidence are more difficult to tamper with but there is not enough evidence to support a firm conclusion or there are unexplained inconsistencies (dissonance) in the observed evidence in light of the hypothesis.
C4	Strong evidence	The source(s) of evidence are much more difficult to tamper with evidence from multiple, independent sources (strong harmonious observations in light of the hypothesis).
C5	Very strong evidence	The source(s) of evidence are very much more difficult to tamper with and evidence from multiple, independent sources (very strong harmonious observations in light of the hypothesis). However, small uncertainties exist (e.g. temporal error, data loss).
C6	Extremely strong evidence	The evidence is tamper proof (or tamper evident) and extremely strong harmonious evidence in light of the hypothesis unquestionable

DERDs (eng: Digital Evidence Certainty Descriptors) (Horsman, 2019):

The digital evidence certainty descriptors.

1 'Conclusive Fact'	Here we are looking at 100% proof in relation to a given scenario for a practitioner to describe it as a 'conclusive fact'. In this case, a set of consequences can be attributable to an act where no other viable alternative reasoning exists. It must be stressed that in line with the true meaning of fact, no potential can exist for an alternative theory. As a result, the use of 'conclusive fact' as a descriptor should be reserved for such cases and should not be utilised in looser terms to describe instances where strictly speaking, alternative theories may exist. Practitioners should utilise caution when opting for 'conclusive fact' as an evidence descriptor as it requires an examination of all alternative hypotheses which may exist regarding a given scenario, where each have been robustly tested and refuted, leaving only one given outcome.
2 'Persuasive'	'Persuasive' describes a digital evidence which cannot be described in terms of a 'Conclusive Fact', yet all available digital data supports the proposed hypothesis. For example, in regards to the interpretation of a digital event provided by the practitioner, it cannot be factually proven, but all actions depicted on a system are consistent with known, accepted and documented device functionality. Whilst the digital data cannot factually prove a given hypothesis, there is no apparent digital data which refutes it. In essence, everything on a system suggests that what the practitioner has described has occurred, with no data present which can refute it, however there are other external factors which could have been responsible.
3 'Conceivable'	A 'Conceivable' descriptor is the lowest form of certainty which a practitioner can use to describe digital evidence. A 'Conceivable' descriptor is appropriate where available digital data partially supports a proposed hypothesis but one or more core requisites are missing in order to fully validate the scenario. Therefore the proposed hypothesis is not disproved due to the given available digital content and remains possible. However, alternative or competing hypotheses remain equally viable. A 'Conceivable' descriptor may be used in conjunction with an 'Insufficient Information' descriptor.
4 'Insufficient Information'	An 'Insufficient Information' descriptor is to be used in cases where there is not enough digital content available to reliably establish and validate a given set of suspect actions. It should be noted that this descriptor does not preclude the chance of any suspected actions from having actually occurred, but there is not enough information available to the practitioner to determine they have occurred with any level of certainty. Circumstances which are described with 'Insufficient Information' may still be feasible.
5 'Implausible'	'Implausible' should be used to be used to describe an event which remains possible but there is no data present to suggest the event has occurred. 'Implausible' describes an event which would require influence from external factors which are both not present within the current case and there are no signs of such involvement. 'Implausible' also describes an event where the core requisites needed in order to validate a current hypothesis are missing. Conversely, such requisites would be present during the accepted normal functionality of the technology subject to an investigation.
6 'Impossible'	'Impossible' is to be used to describe a set of digital events which cannot possibly occur taking into account the current scenario and technology under investigation. Similar to the use of 'Conclusive Fact', this descriptor must be approached with caution due to the diversity of technological function and the potential for undocumented functionality to exist. To provide an example of an accepted use for the 'Impossible' descriptor would be to describe the event where two standard IDE hard drives are suspected of transferring digital data between the two via the physical touch of their external casing.

15. Dersom det er gjort en vurdering av funnenes bevisverdi – er dette utført i henhold til en gjeldende standard/rammeverk?

Vurdering av funnets bevisverdi (sannsynlighet) gjøres i lys av minst to konkurrerende hypoteser og de omstendigheter som er lagt til grunn. En slik vurdering gjennomføres på bakgrunn av et konkret oppdrag (f.eks. sakkyndigoppdrag) – og bør følge en etablert framgangsmåte, som beskrevet i ENFSI (2015a; 2015b), slik at både funnenes verdi og begrensninger/usikkerhet blir tydelig kommunisert.

Jus/etikk:

16. Er undersøkelsen og rapportskrivningen gjennomført i tråd med gjeldende straffeprosessuelle regler og relevante etiske retningslinjer?

Undersøkelser av databaseslag og rapportskrivning som gjennomføres i straffesak er etterforskning, og dermed en formålsstyrt virksomhet som er regulert i straffeprosessloven. Undersøkelser av databaseslagets innhold/spor er et inngrep i noens privatliv, og regnes som en ransaking i juridisk forstand. Straffeprosesslovens § 170 a setter krav om at ransakingen skal være nødvendig og forholdsmessig.

Etiske retningslinjer for politiet (POD 2018/05) skal også følges:
<https://www.politiet.no/globalassets/05-om-oss/04-om-oss/00-jobb-hos-oss/etiske-retningslinjer-for-politiet.pdf>

17. Framgår det av rapporten hvorvidt eller i hvilken grad dataetterforsker kan betraktes som uavhengig med hensyn til sine undersøkelser og dokumentasjon?

Dataetterforskning gjennomføres internt i politiet, og det er dermed ingen forventning om at dataetterforsker er uavhengig – slik som eksterne sakkyndige vil være. Unntaket er ved ransaking av datamateriale for å skille ut advokatkorrespondanse eller annet materiale som omfattes av

beslagsforbud, se HR-2022-1317-A. I slike tilfeller bør rapporten inneholde en beskrivelse av hvordan denne uavhengigheten er ivarettatt.

18. Er det søkt aktivt etter informasjon som kan underbygge uskyld eller formildende omstendigheter for mistenkte, og er resultatet beskrevet i rapporten?

Uskyldspresumsjonen er et viktig prinsipp som operasjonaliseres gjennom å aktivt søke etter informasjon som underbygger uskyld. Dersom noen er mistenkte skal man etter straffeprosesslovens § 226 3. ledd søke å kartlegge det som taler mot og det som taler til fordel for ham. Dette kan for eksempel være å søke etter informasjon som bekrefter opplysninger som mistenkte har kommet med i sin forklaring. Hvilke opplysninger det kan være aktuelt å søke etter bør avklares med etterforskningsledelsen, og beskrives i mandatet/oppdraget.

Nivå 5 – Verifisering av utvalgte spor

Dette nivået omfatter undersøkelser av databeslaget i tillegg til selve rapporten. Her tar fagfellen utgangspunkt i et utvalg av funnene, og undersøker disse med tanke på å verifisere dem. Et viktig prinsipp for å redusere usikkerheten i verifiseringen er å bruke et annet verktøy enn det første dataetterforsker benyttet, og/eller – dersom fagfellen har tilstrekkelig kompetanse – verifisere funnet gjennom manuell undersøkelse av datagrunnlaget.

Kan utføres av: Personell med spesialistkompetanse innen dataetterforskning, og helst høyere kompetanse enn rapportskriver på området kontrollen omfatter.

1. Har dataetterforsker gjort alle funn, samt de data funnene baserer seg på, tilgjengelig for verifisering?

2. Har fagfellen gjort rede for hvordan utvalget av spor er foretatt og begrunnet dette? (tilfeldighetsutvalg, representativt utvalg, andre kriterier)

3. Har fagfellen gjort rede for hvordan verifiseringen er gjennomført ('dual tool' verifisering, manuell undersøkelse, basert på tidligere erfaring, basert på forskning eller lignende)

4. Har fagfellen dokumentert verifiseringsprosessen og resultatet av denne på en nøyaktig måte?

Nivå 6 – Verifisering av alle spor

Dette nivået omfatter, i likhet med nivå 5, undersøkelser av databeslaget i tillegg til selve rapporten. Her tar fagfellen utgangspunkt i alle funn, og undersøker disse med tanke på å verifisere dem. Som i nivå 6 er det viktig å bruke et annet verktøy enn det første dataetterforsker benyttet, og/eller – dersom fagfellen har tilstrekkelig kompetanse – verifisere funnet gjennom manuell undersøkelse av datagrunnlaget.

Kan utføres av: Personell med spesialistkompetanse innen dataetterforskning, og helst høyere kompetanse enn rapportskriver på området kontrollen omfatter.

1. Har dataetterforsker gjort alle funn, samt de data funnene baserer seg på, tilgjengelig for verifisering?

2. Har fagfellen gjort rede for hvordan verifisering er gjennomført ('dual tool' verifisering, manuell undersøkelse, basert på tidligere erfaring, basert på forskning eller lignende)

3. Har fagfellen dokumentert verifiseringsprosessen og resultatet av denne på en nøyaktig måte?

Oppsummering / Synergi

Beskrivelse av rapportens styrker:

Sjekklisten vil lett kunne bli et verktøy for å identifisere feil, utfordringer og forbedringspunkter, mens god dokumentasjons- og rapporteringspraksis kommer i bakgrunnen. Dette punktet blir derfor ekstra viktig for å identifisere og gi tilbakemelding på hva som fungerer spesielt godt i rapporten, og som rapportskriver dermed bør fortsette med.

Beskrivelse av rapportens utfordringer og forbedringspunkter:

I dette punktet skal det gis en oppsummerende beskrivelse av det som er rapportens utfordringer og hvor det er behov for forbedring. Her kan det være relevant å peke på enkelte elementer som er spesielt kritiske, men det kan også være summen av mange 'gule' avkryssninger som samlet sett utgjør rapportens forbedringspunkt.

Andre kommentarer:

Det kan finnes forhold i rapporten som ikke fanges opp av sjekklistens punkter og som det er behov for å kommentere – og slike punkter kan gjerne beskrives her.

Beskrivelse av forhold som skal meldes inn i Synergi, og hvem som foretar dette:

Dersom det som følge av fagfelleevalueringen oppdages feil som relateres til systematiske feilkilder, som for eksempel feil i verktøy eller etablerte metoder/prosedyrer må disse meldes inn i Synergi så snart som mulig. Rapportskriver og fagfelle avtaler hvem som skal melde dette inn.

Referanser:

- Ask, K., Rebellius, A., & Granhag, P. A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 22(9), 1245-1259.
<https://doi.org/10.1002/acp.1432>
- Casey, E. (2002). Error, uncertainty, and loss in digital evidence. *International Journal of Digital Evidence*, 1 (2).
- Casey, E. (2020). Standardization of forming and expressing preliminary evaluative opinions on digital evidence. *Forensic Science International: Digital Investigation*, 32, 200888.
<https://doi.org/10.1016/j.fsidi.2019.200888>
- Council of Europe (2017). Digital Forensics. A basic guide for the management and procedures of a digital forensics laboratory. Version 1.1.

- ENFSI (2015a) *ENFSI Guideline for evaluative reporting in forensic science. Strengthening the evaluation of forensic results across europe (STEOFRAE)*. European Network of Forensic Science Institutes. https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf
- ENFSI (2015b). *Best practice manual for the forensic examination of digital technology, ENFSI-BPM-FOT-01, Version 01 (November 2015)*. European Network of Forensic Science Institutes. https://enfsi.eu/wp-content/uploads/2016/09/1_forensic_examination_of_digital_technology_0.pdf
- Flaglien, A. O. (2018). Digital Forensics Process. I: A. Årnes (red.) *Digital Forensics*. Wiley.
- Forensic Science Regulator (2020). *Codes of practice and conduct. Appendix: Digital Forensic Services. FSR-C-107 Issue 2*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/912389/107_FSR-C-107_Digital_forensics_2.0.pdf.
- Horsman, G. (2020). Digital evidence certainty descriptors (DECDS). *Forensic Science International: Digital Investigation*, 32, 200896. <https://doi.org/10.1016/j.fsidi.2019.200896>
- Norges Høyesterett (2022). HR-2022-1317-A, Kjennelse avsagt 30. juni 2022 av Høyesterett. <https://www.riksadvokaten.no/wp-content/uploads/2022/07/Kjennelse-avsagt-30.-juni-2022-av-H%C3%B8yesterett-1.pdf>
- Sunde, N., & Dror, I. E. (2019). Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. *Digital Investigation*, 29, 101-108. <https://doi.org/10.1016/j.diin.2019.03.011>
- Sunde, N., & Dror, I. E. (2021). A hierarchy of expert performance (HEP) applied to digital forensics: Reliability and biasability in digital forensics decision making. *Forensic Science International: Digital Investigation*, 37. <https://doi.org/10.1016/j.fsidi.2021.301175>

Versjonslogg	
1.0 september 2022	«Støttedokument til Sjekkliste for Kvalitetskontrollhierarkiet» utarbeidet til Workshop I, Kvalitetskontrollprosjektet
1.1 desember 2022	«Støttedokument til Sjekkliste for Kvalitetskontrollhierarkiet» - oppdatert etter innspill fra deltakere i Workshop II, Kvalitetskontrollprosjektet

Fagfellevurdering av rapporter fra datatekniske undersøkelser

v/Nina Sunde

Teksten er et sammendrag av artiklene:

Horsman, G., Sunde, N. (2020). Part 1: The need for peer review in digital forensics. *Forensic Science International: Digital Investigation* 35, 301062. <https://doi.org/10.1016/j.fsidi.2020.301062>

Sunde, N., Horsman, G. (2021). Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations. *Forensic Science International: Digital Investigation*, 36, 301074. <https://doi.org/10.1016/j.fsidi.2020.301074>

NB: PARS konseptet som beskrives i andre artikkel (Part 2) omhandler både *kvalitetssikring* underveis i dataetterforskningsprosessen og *fagfellevurdering/kvalitetskontroll* av resultatet. Dette sammendraget omhandler kun det siste aspektet – altså kvalitetskontroll av resultatet.

1. Innledning

Digitale spor og bevis innhentes i dag i de fleste etterforskninger, uavhengig av sakens størrelse og kompleksitet. Slike spor og bevis håndteres av personell med ulik bakgrunn og kompetanse, og behovet for kvalitetssikring er derfor presserende. Selv om det er lett å være enig i viktigheten av å levere høy kvalitet, er dette en av de største utfordringene i dataetterforskningsfeltet som fortsatt ikke er tilstrekkelig adressert.

Vi skiller mellom to dimensjoner av kvalitetsstyring (Doyle, 2019): *kvalitetssikring* (eng: quality assurance), som er en forebyggende virksomhet og som fokuserer på et system som tilrettelegger for at kvalitetskravene kan oppfylles og *kvalitetskontroll* (eng: quality control), som er reaktivt, og som fokuserer på å kontrollere at kvalitetskravene er oppfylt i etterkant av utført arbeid.

Kjernen i denne utfordringen knytter seg til påliteligheten av informasjonen som produseres gjennom dataetterforskningsprosessen. Å kontrollere påliteligheten har først og fremst betydning for å ivareta rettsikkerheten, ved å forebygge at villedende eller feilaktig informasjon brukes i etterforskningen eller retten. Det har også betydning for skadebegrensning, altså å begrense skade for organisasjonens eller den enkelte etterforskers omdømme. Et tredje aspekt handler om kostnadsbesparelse. Feil som avdekkes og rettes tidlig forhindrer kostnader ved å rette feilene senere i etterforskningen eller når rettsprosessen tar til.

Standardisering er et av flere virkemidler for å oppnå kvalitet. I England og Wales er det krav til at dataetterforskningsenheter (laboratorier) skal akkrediteres etter standarden ISO/IEC 17025. En undersøkelse av implementeringen av standarden i laboratoriene viste imidlertid at standarder ikke er noen absolutt garanti for kvalitet. Akkrediteringen setter krav om at kvalitetssystemer er på plass, men undersøkelsen viste at selv om utøveren er dyktig, så er det avgjørende med nødvendig utstyr, tilstrekkelige ressurser og kontinuerlig opplæring – ellers vil ikke systemet fungere som det skal (Tully et al, 2020).

2. Feilkilder i datatekniske undersøkelser:

Kvalitetssystemer handler ikke bare om å avdekke feil, men også å identifisere kilden til feil – slik at framtidige feil kan forebygges. Vi kan dele feilkilder inn i to hovedkategorier: 'tekniske feilkilder' og 'ikke-tekniske feilkilder'.

Tekniske feilkilder

Tekniske feilkilder kan deles inn i tre kategorier: feil i teknikker, implementeringsfeil og tolkningsfeil (SWGDE, 2018).

Feil som kan tilskrives *teknikker* (algoritmer) er de som kan oppstå tilfeldig, og beregnes gjennom feilrater. Det er imidlertid ikke enkelt å vurdere og estimere feilraten for dataetterforskningsverktøy, siden dette krever en 'populasjon' – altså et sett med kjente og dokumenterte data som viser en spesifikk teknikk oppførsel. På grunn av den raske teknologiutviklingen, samt kompleksiteten og mangfoldet av enheter og teknologi, kan det være vanskelig å oppnå gode estimater for feilrater – og i mange tilfeller ikke mulig (SWGDE, 2018)

Implementeringsfeil er feil som skyldes implementering av teknikkene/algoitmene i verktøy. Siden dette er logiske feil, vil dette være systematiske feil som gjentar seg helt til feilen er rettet. Implementeringsfeil oppstår gjerne under utvikling av verktøy, for eksempel programvare for analyse av digitale spor. Selv om verktøyet testes før distribusjon er det ofte ikke mulig å garantere feilfrihet, noe som ofte gjenspeiles i verktøyets sluttbrukerlisensavtale (EULA) (Horsman, 2018, 2019). Implementeringsfeil deles gjerne inn i tre typer:

- a) Ufullstendighet – innebærer at ikke all tilgjengelig informasjon identifiseres av verktøyet.
- b) Unøyaktighet – innebærer at verktøyet rapporterer artefakter som ikke eksisterer, grupperer ikke-relaterte elementer sammen eller endrer data på en måte som endrer meningen.
- c) Feiltolkning – innebærer at resultatene blir tolket og presentert feil av verktøyet.

Denne siste kategorien kan også relateres til interaksjonen mellom menneske og teknologi, og handler om feil ved verktøybruk og tolkning av resultater. Selv om verktøyet er testet og fungerer korrekt kan det fortsatt oppstå brukerfeil ved at verktøyet brukes på feil måte eller til oppgaver det ikke er laget for å håndtere.

Ikke-tekniske feilkilder

Ikke-tekniske feilkilder kan lede til både tilfeldige og systematiske feil. Ikke-tekniske feil kan være vanskelig å gradere seg mot, siden de kan forekomme i alle deler av etterforskningsprosessen, også hvor det er forventet at utøveren vil presentere et feilfritt resultat. Behovet for kvalitetskontroll er derfor stort for å hindre feil som skyldes ikke-tekniske feilkilder. Systematiske feil bør være en større bekymring enn tilfeldige, siden de har en tendens til å gjenta seg.

Mangelfull *kunnskap* er en viktig kilde til feil. Kunnskap er nødvendig for å finne relevante spor, forstå deres betydning og være i stand til å undersøke og fastslå hvorvidt verktøyet som blir benyttet kan ha fortolket datagrunnlaget feilaktig. Å ha forståelse for hvor langt egen kunnskap rekker, og hvor man har behov for å søke bistand hos noen med ytterligere ekspertise er avgjørende for å unngå feil.

Erfaring er viktig i forebyggingen av feil. En erfaren utøver kan være bedre i stand til å oppdage noe som avviker fra det forventede utfallet. Erfaring kan også sette utøveren i bedre stand til å gjenkjenne situasjoner hvor feil har oppstått tidligere, og dermed unngå å gå i samme fellen på nytt.

Manglende *prosedyrer* kan være en kilde til feil. Det finnes veletablerte prinsipper i dataetterforskning som tjener som et kompass i arbeidet, som å etterstrebe ivaretagelsen av bevisets integritet og en ubrutt beviskjede¹ (Casey, 2011; Årnes, 2018), eller som beskrevet i ISO/IEC 27037:2012 – å oppnå etterprøvnbarhet, forsvarlighet, repeterbarhet og reproduserbarhet². Prinsippene gir imidlertid ingen veiledning i hvordan dataetterforsker skal gå fram for å etterleve dem, og til dette trengs kvalitetsssikrede prosedyrer og metoder. Forskning har vist at klare prosedyrebeskrivelser og støtteverktøy som sjekklister er viktige for å hindre at det glipper, selv for eksperter.

¹ Omtales gjerne som 'chain of custody' eller 'audit trail' på engelsk

² Auditability, justifiability, repeatability, reproducibility

Enhver beslutningsprosess som involverer mennesker innebærer risiko for *menneskelige tanke- og beslutningsfeil*, og spesielt når prosessen innebærer høy grad av subjektivitet, fortolkning, vurderinger og beslutninger. Dette er faktorer som gjør seg gjeldende i samtlige av dataetterforskningsprosessens faser, men kanskje spesielt i analyse og rapporteringsfasen. Forskning har vist at kognitive bias (systematiske skjevheter, slagsider) er en utfordring i forensisk/kriminalteknisk arbeid generelt.³ Bias settes i sammenheng med indre og ytre påvirkningsfaktorer knyttet til den menneskelige natur og hjernens funksjoner, sosiale og kulturelle miljøfaktorer, erfaring og saksspesifikke faktorer. En utfordring her er at mange utøvere av forensiske/kriminaltekniske undersøkelser er blinde for eget bias, og tror at påvirkningsfaktorene kan elimineres ved ren viljestyrke. I tillegg til de nevnte påvirkningsfaktorene har det gjennom forskning blitt identifisert en rekke situasjonelle faktorer som kan øke risikoen for bias, som tidspress, følelser, opplevd ansvar og bekymring for framtidige konsekvenser (Ask, 2013).

3. Formål med fagfellevurdering

Formålet med fagfellevurdering er først og fremst å avdekke og rette feil i digitale spor og bevis med verdi for straffesaken. Fagfellevurdering er også innrettet mot å vurdere kvaliteten av etterforskningsprosessen som ledet fram til resultatene. Selv om informasjonen ikke er feilaktig, kan den bli framstilt med på en villedende måte. For eksempel kan funnenes relevans eller sikkerhet overdrives ved at usikkerhetsmomenter eller forbehold underkommuniseres eller at viktig informasjon er utelatt. Et tredje formål med fagfellevurdering er å forbedre kvaliteten på framtidig dataetterforskningsarbeid gjennom erfaringslæring. Fagfellevurdering kan føre til oppdagelse av kritiske systematiske feil, og muliggjøre endring i prosedyrer eller verktøy slik at den systematiske feilen ikke gjentas i framtidig undersøkelse.

4. utfordringer med kvalitetskontroll i dataetterforskning

Det finnes lite forsknings- eller erfaringsbasert kunnskap kvalitetskontroll i dataetterforskningsarbeid, og behovet for mer kunnskap er stort. Vi mangler kunnskap om:

- i hvilket omfang kvalitetskontroll blir gjennomført ved dataetterforskningsenhetene
- hvordan kvalitetskontroll blir gjennomført
- hvilke prosess-steg som omfattes av kvalitetskontroll
- hvilket nivå av kvalitetskontroll som gjennomføres
- hvem som innehar rollen som fagfelle
- hvordan dataetterforskere responderer på fagfellevurderingen
- opplæring i fagfellevurderingsarbeid og hva som er tilstrekkelig kompetanse og erfaring
- hva som kjennetegner en effektiv kvalitetskontroll
- hvordan fagfellevurderingsarbeid bør evalueres
- kvalitetskontroll gjennomføres som en selvregulert aktivitet eller om den organiseres av leder
- følelsesmessige aspekter av å gi eller få fagfellevurdering
- kvalitetskontroll oppleves som meningsfullt arbeid
- digitale spor og bevis er for komplekse for tradisjonell vurdering eller verifikasjon
- hvordan aksept for fagfellevurderingens resultater oppnås
- den bevismessige verdien av statusen 'fagfellevurdert'

³ Nylige studier har vist at problemet også omfatter dataetterforskere (Sunde og Dror, 2019; 2021)

5. Kvalitetskontroll hierarkiet for dataetterforskning.

Kvalitetskontrollhierarkiet består av 7 nivå, med tilhørende beskrivelse av hva kontrollen skal omfatte og sjekklister som støtteverktøy (Sunde og Horsman, 2021). Hierarkiet er en videreutvikling av et rammeverk som opprinnelig ble presentert av Page et al. (2019). Nivådelingen er begrunnet i en tanke om at det er viktig med transparens omkring omfanget av kvalitetskontrollen, altså hva som er vurdert/ ikke vurdert, på hvilken måte, og hvilke feil den dermed er egnet til å oppdage. I tillegg er nivådelingen hensiktsmessig for å kunne variere og tilpasse omfanget av kvalitetskontroll i en organisasjon som gjerne preges av stort sakspress og knapphet i ressurser.

Hierarkiets laveste nivåer anses å være mindre ressurskrevende enn nivåene over, men samtidig også mindre egnet til å avdekke kritiske feil enn de høyeste nivåene. Hierarkiet er videre organisert slik at nivåene opp til og med nivå 6 innebefatter aktivitetene i nivåene under. Altså, en innholdskontroll (nivå 4) omfatter også administrativ kontroll, språkvask og kontroll av klarhet og tilgjengelighet (nivå 1-3). Nivå 1-4 omfatter kun *rapporten* fra datatekniske undersøkelser, mens nivåene 5-7 også omfatter undersøkelser av *datagrunnlaget/speilfilen* for verifisering av funn eller ny analyse. De første tre nivåene krever liten grad av fagkompetanse, og kan utføres av andre enn dataetterforskere.

Nedenfor følger en beskrivelse av de enkelte nivå. De tilhørende sjekklister vil være mer dynamiske og kreve tilpasning til den aktuelle organisasjonen. De vil dermed omfatte flere punkter enn det som inngår i kulepunktene under hvert nivå i den opprinnelige beskrivelsen av kvalitetskontrollhierarkiet, som er inntatt her.



Figur 1: Kvalitetskontrollhierarkiet ⁴

Nivå 1 - Administrativ kontroll

En administrativ kontroll omfatter hvorvidt etterforskningen er gjennomført i tråd med formelle krav og i henhold til oppdraget, altså at avtalte undersøkelser er gjennomført på de respektive beslaglagte enhetene. Sjekken bør omfatte:

- Er mandatet/oppdraget nøyaktig beskrevet i rapporten og svarer rapporten dekkende på mandatet/oppdraget og formålet med undersøkelsen?
- Inneholder rapporten en beskrivelse av hvilken saksinformasjon rapportskriver hadde tilgang til ved analysen av databeslaget?

⁴ Norsk oversettelse. Opprinnelig versjon finnes i Horsman og Sunde (2020) og Sunde og Horsman (2021)

- Har rapportskriver beskrevet egen kompetanse med relevans for oppdraget. Gis en begrunnelse for hvorvidt kompetansen relevant og tilstrekkelig for oppdraget som er utført?
- Er rapporten skrevet i tråd med mal eller rapportstruktur som benyttes ved enheten, og er den signert?
- Er sluttproduktet i tråd med hva man ble enig med oppdragsgiver om?

Nivå 2 - Språkvask

Språkvask omfatter en vurdering av om rapporten inneholder stave- og grammatiske feil som bør rettes opp, og bør omfatte:

- Er det stavefeil i rapporten som må rette opp?
- Er det grammatiske feil som bør rettes opp?
- Om forkortelser er benyttet – er de forklart, og brukes de på en konsistent måte gjennom rapporten?

Nivå 3 - Kontroll av klarhet og tilgjengelighet

Kontroll av klarhet og tilgjengelighet innebærer å vurdere om rapportskriver evner å framstille resultatet av undersøkelsen på en tydelig, forståelig og ryddig måte for en leser uten særskilt teknisk kompetanse. Vurderingen bør omfatte:

- Er språket som brukes i rapporten forståelig?
- Er tekniske ord og uttrykk forklart?
- Har rapportskriver tydeliggjort på hvilket nivå rapporteringen skjer, altså om rapporteringen -
 - o er deskriptiv? (beskrivelser av innhold og metadata basert på hva analyseverktøyet viste)
 - o inneholder vurdering og fortolkning av hva innhold og metadata betyr i lys av sakens omstendigheter som rapportskriver har kunnskap om?
 - o er en foreløpig eller endelig evaluering av funnenes verdi i lys av et sett med hypoteser og definerte omstendigheter/betingelser (eng. evaluative opinion)?
- Framgår det tydelig om rapporten presenterer foreløpige eller endelige resultater

Nivå 4 - Innholdskontroll

Innholdskontroll omfatter en grundig kontroll av rapportens innhold som beskriver resultatet av etterforskningen, men avgrenses mot en verifisering av funn/resultater. Hovedfokus rettes mot det vitenskapelige og logiske fundamentet i rapporten. Vurdering av sammenhengen mellom bevisene som presenteres og konklusjonen er spesielt viktig. Prinsippene om balanse, logisk sammenheng, robusthet, transparens og jus/etikk danner rammen for elementene som skal vurderes, og bør omfatte:

Balanse:

- Er hypotesene som dannet utgangspunkt for undersøkelsen beskrevet?
- Er hensiktsmessige under-hypoteser utarbeidet på et relevant nivå?
- Er resultatet beskrevet i lys av minst to konkurrerende hypoteser?
- Er det intern sammenheng mellom resultatene og konklusjonen?
- Dersom evalueringen av funnenes verdi er gjort, er dette utført i henhold til en gjeldende standard/rammeverk, f.eks. Digital Evidence Certainty Descriptors (Horsman, 2020), C-Scale (Casey, 2002; Casey 2020) eller Case Assessment and Interpretation (CAI) (Jackson et al., 2015), og er den valgte strukturen referert i rapporten?

Logikk:

- Er premissene for slutninger, antakelser og tolkninger begrunnet og forklart?
- Er konklusjonen begrunnet og forklart, og er den balansert?
- Er konklusjonens styrke (konklusjonsgrad) i samsvar med funnene den er basert på?

Robusthet:

- Har den gjennomførte undersøkelsen vært av et tilstrekkelig omfang for å gi grunnlag for konklusjonen?

Transparens:

- Refererer rapporten til prosesser/prosedyrer som er blitt utført? Er avvik fra disse begrunnet og dokumentert?
- Er framgangsmåter beskrevet så nøyaktig at de kan gjentas av andre (med samme resultat)?
- Er funnene presentert på en nøyaktig og detaljert måte?
- Er funnenes beskrevet i lys av den konteksten/sammenhengen de er funnet?
- Er negative funn (det som ble søkt etter men ikke funnet) dokumentert?
- Er forbehold, usikkerhet og begrensninger med metoder, verktøy, funn eller konklusjoner gjort rede for?

Jus/etikk:

- Ble undersøkelsen og rapportskrivningen gjennomført i tråd med straffeprosessuelle regler og relevante etiske retningslinjer?
- Framgår det av rapporten hvorvidt eller i hvilken grad dataetterforsker kan betraktes som uavhengig med hensyn til sine undersøkelser og dokumentasjon?
- Har funn som kan underbygge mistenktes uskyld eller formildende omstendigheter blitt aktivt søkt etter, og er resultatet beskrevet i rapporten?

Nivå 5 - Verifisering av utvalgte spor

Dette nivået omfatter undersøkelser av databeslaget i tillegg til selve rapporten. Her tar fagfellen utgangspunkt i et utvalg av funnene, og undersøker disse med tanke på å verifisere dem. Et viktig prinsipp for å redusere usikkerheten i verifiseringen er å bruke et annet verktøy enn det første dataetterforsker benyttet, og/eller – dersom fagfellen har tilstrekkelig kompetanse – verifisere funnet gjennom manuell undersøkelse av datagrunnlaget. Fagfellevurdering på dette nivået bør omfatte:

- Har dataetterforsker gjort alle funn, samt de data funnene baserer seg på, tilgjengelig for verifisering?
- Kan et utvalg av funnene verifiseres av fagfellen ved å bruke et annet verktøy enn den opprinnelige dataetterforskeren brukte?
- Har fagfellen dokumentert verifiseringsprosessen og hvilke funn som er verifisert på en nøyaktig og detaljert måte?

Nivå 6 - Verifisering av alle spor

Dette nivået omfatter, i likhet med forrige nivå, undersøkelser av databeslaget i tillegg til selve rapporten. Her tar fagfellen utgangspunkt i alle funn, og undersøker disse med tanke på å verifisere

dem. Som i nivå 6 er det viktig å bruke et annet verktøy enn det første etterforsker benyttet, og/eller – dersom fagfellen har tilstrekkelig kompetanse – verifisere funnet gjennom manuell undersøkelse av datagrunnlaget. Fagfellevurdering på dette nivået bør omfatte:

- Har dataetterforsker gjort alle funn, og de data funnene baserer seg på, tilgjengelig for verifisering?
- Kan samtlige funn verifiseres av fagfellen ved å bruke et annet verktøy enn den opprinnelige dataetterforskeren brukte?
- Har fagfellen dokumentert verifiseringsprosessen og hvilke funn som er verifisert på en nøyaktig og detaljert måte?

Nivå 7 - Ny analyse

En ny analyse innebærer å gjennomføre en prosessering, analyse, fortolkning og rapportering/dokumentasjon på nytt av en dataetterforsker som ikke har arbeidet med saken tidligere, og som ikke kjenner framgangsmåten eller utfallet av den opprinnelige undersøkelsen av databeslaget. Dette vil være den mest tids- og ressurskrevende tiltaket sammenlignet med kvalitetskontrollhierarkiets øvrige nivåer, men vil samtidig gi en god mulighet for å utfordre eller den opprinnelige undersøkelsen. Dette nivået vil være effektivt med hensyn til å forebygge kognitive bias, siden dataetterforsker ikke bringer med seg forventninger basert på tidligere undersøkelser. En ny analyse kan føre til de de samme funnene som den opprinnelige analysen, men også nye funn eller nye fortolkninger av funn. Dersom to uavhengige dataetterforskere har kommet fram til samme funn og lik fortolkning av disse, vil dette representere en verifisering med mindre grad av usikkerhet sammenlignet med en verifisering på nivå 6 eller 7. For å forhindre at feil som skyldes implementeringsfeil i analyseverktøyet, bør ny analyse på nivå 7 i likhet med nivå 5 og 6 fortrinnsvis basere seg på andre verktøy enn den opprinnelige analysen.

6. Konfliktløsning ved uenighet mellom dataetterforsker og fagfelle

Dersom fagfelle og rapportskriver er uenige om utfallet av fagfellevurderingen, er det viktig å ha en plan for hvordan dette skal løses. Uenigheten løftes derfor først inn på ledelsesnivå, for å klarlegge hva uenigheten består i. Dersom det har blitt gitt veiledning underveis i dataetterforskningen (se 'the Peer Advisor Stage' s. 4-8 i Sunde og Horsman, 2021), kan eventuell dokumentasjon rundt dette innhentes for å belyse temaet uenigheten handler om. Dersom fagfelle og rapportskriver ikke kommer til enighet sammen med leder etter at premissene for uenigheten er drøftet, vil det være aktuelt å bringe inn en ny intern fagfelle, eller evt. en ekstern fagfelle med høyere kompetanse på området uenigheten dreier seg om. Dokumentasjon og transparenser er viktig i fagfellevurderingsprosessen, og like viktig i prosessen for konfliktløsning.

7. Roller, kompetanse og prosedyre

En fagfelle bør som minimum ha samme eller høyere kompetanse enn rapportskriver. Fagfellen bør også ha kompetanse innenfor veiledning/tilbakemelding.

For å sikre størst mulig uavhengighet og forebygge bias (statuseffekten) er det viktig at fagfellen ikke har hatt noen rolle i etterforskningen fagfellevurderingen omhandler, eller på annen måte kjenner saken. Dersom det er mulig bør fagfellevurderingen skje «blind», dvs. at fagfellen ikke vet hvem som har skrevet rapporten.

En fallgrube i fagfellevurderingsprosessen er vår tilbøyelighet til bekreftelse gjennom å gå fra konklusjon til informasjon, i stedet for fra informasjon til konklusjon. For å forebygge bias og øke

sjansen til å oppdage feil eller villedende informasjon bør fagfellevurdering på nivå 4-6 gjennomføres slik:

1. Fagfellen får tilgang til rapporten som skal vurderes, men konklusjonen er utelatt eller maskert.
2. Fagfellen vurderer rapportens funn og fortolkninger av funn i lys av de angitte hypotesene, og formulerer sin egen konklusjon på bakgrunn av disse.
3. Så sammenligner fagfelle opprinnelige konklusjon opp imot egen konklusjon.

Et samsvar her vil være beheftet med mindre grad av usikkerhet sammenlignet med en prosedyre hvor man går fra konklusjon til informasjon. Prosedyren er i tråd med prinsippene som metoden Linear Sequential Unmasking (LSU) (Dror et al, 2015) bygger på. LSU benyttes for å forebygge beslutningsfeil som skyldes påvirkning fra informasjon uten relevans for oppgaven (kontekstuell bias) i forensiske disipliner som gjør sammenlignende undersøkelser på kilde-nivå (f.eks. DNA, fingeravtrykk).

4. Dokumentasjon av fagfellevurderingen

For å sikre transparens og notoritet i fagfellevurderingsprosessen er det viktig med forutsigbarhet for hva kontrollen skal omfatte, og god dokumentasjon over hva som er vurdert/kontrollert. Det er derfor utarbeidet maler med sjekkpunkter for hvert nivå (til og med nivå 6) i kvalitetskontrollhierarkiet. Denne dokumentasjonen danner grunnlag for kommunikasjonen mellom fagfelle og rapportskriver, og vil gi oversikt over hva som er kontrollert, på hvilket nivå – og like viktig – hva som *ikke* er kontrollert. Dette er verdifull informasjon for den som skal vurdere beviset, og kan forebygge overdreven tro på feilfrihet og kvalitet i de digitale bevisene. Når fagfellevurderingen er gjennomført bør dokumentasjonen lagres slik at den kan legges fram i retten om det skulle oppstå behov for dette.

Det anbefales at det framgår av analyserapporten hvorvidt fagfellevurdering er gjennomført, på hvilket nivå i henhold til kvalitetskontrollhierarkiet, og av hvem.

5. Avsluttende kommentarer

Det finnes mange versjoner av dataetterforskningsprosessen i akademisk litteratur og fagbøker, men de færreste har beskrevet kvalitetskontroll som en del av prosessens steg (Jahren, 2020). Tradisjonelt har man trukket et skille mellom administrativ og teknisk fagfellevurdering, men det finnes ingen tydelige og detaljerte føringer for hva de ulike kontrollene skal omfatte. Formålet med kvalitetskontrollhierarkiet er å øke presisjonsgraden på kontrollen gjennom et mer nivådelt rammeverk. Med dette vil man kunne oppnå økt transparens omkring hva som er kontrollert, i hvilket omfang og av hvem. Ingen rammeverk eller prosedyrebeskrivelser garanterer feilfrihet, og dersom prosessen utføres av fagfeller uten tilstrekkelig kompetanse eller manglende grundighet kan dette tiltaket tjene mot sin hensikt ved å gi et skinn av kvalitet uten at dette faktisk er tilfelle. Et kvalitetssystem fordrer derfor en god kvalitetsledelse, som sørger for tilstrekkelig kvalifisert personell til arbeidet, med rammebetingelser som legger til rette for at arbeidet utføres i tråd med hva de ulike nivåene i kvalitetskontrollhierarkiet krever.

Referanser:

- Ask, K. (2013). Bias: feil og faldgruber i etterforskning. I: Hald, C., Vrist Rønn, K. (Red), *Om at oppdage: Metodiske refleksjoner over politiets undersøgelsespraksis*, s. 149-169. Samfundslitteratur.
- Casey, E. (2002). Error, uncertainty and loss in digital evidence. *International Journal of Digital Evidence*, 1 (2).

- Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers, and the internet*. Academic press.
- Casey, E. (2020). Standardization of forming and expressing preliminary evaluative opinions on digital evidence. *Forensic Science International: Digital Investigation*, 32, 200888. <https://doi.org/10.1016/j.fsidi.2019.200888>
- Doyle, S. (2019). *Quality Management in forensic science*. Academic Press.
- Dror, I. E., Thompson, W.C., Meissner, C.A., Kornfield, I., Krane, D., Saks, M., Risinger, M. (2015). Letter to the editor-context management toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of Forensic Sciences*, 60 (4), 1111-1112. <https://doi.org/10.1111/1556-4029.12805>
- Horsman, G. (2018). "I couldn't find it your honour, it mustn't be there!" Tool errors, tool limitations and user error in digital forensics. *Science & Justice*, 58 (6), 433-440. <https://doi.org/10.1016/j.scijus.2018.04.001>
- Horsman, G. (2019). Tool testing and reliability issues in the field of digital forensics. *Digital Investigation*, 28, 163-175. <https://doi.org/10.1016/j.diin.2019.01.009>
- Horsman, G. (2020). Digital evidence certainty descriptors (DECEDs). *Forensic Science International: Digital Investigation*, 32, 200896. <https://doi.org/10.1016/j.fsidi.2019.200896>
- Jackson, G., Aitken, C., Roberts, P. (2015). *Case Assessment and Interpretation of Expert Evidence. Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*. Practitioner Guide No 4. Royal Statistical Society. <http://www.rss.org.uk/Images/PDF/influencing-change/rss-case-assessmentinterpretation-expert-evidence.pdf>
- Jahren, J. H. (2020). *Is the quality assurance in digital forensic work in the Norwegian police adequate?* (Master's thesis). The Norwegian University of Science and Technology, Gjøvik. <https://hdl.handle.net/11250/2781174>
- Page, H., Horsman, G., Sarna, A., Foster, J. (2019). A review of quality procedures in the UK forensic sciences: what can the field of digital forensics learn? *Science & Justice*, 59 (1), 83-92. <https://doi.org/10.1016/j.scijus.2018.09.006>
- Tully, G., Cohen, N., Compton, D., Davies, G., Isbell, R., Watson, T. (2020). Quality standards for digital forensics: learning from experience in England & Wales. *Forensic Science International: Digital Investigation*, 200905 <https://doi.org/10.1016/j.fsidi.2020.200905>
- SWGDE, 2018. *Establishing confidence in digital and multimedia evidence forensic results by error mitigation analysis*. Version: 2.0, (November 20, 2018).
- Årnes, A. (Ed.), 2018. *Digital Forensics*. Wiley.

Se for øvrig referanselistene i:

- Horsman, G., Sunde, N. (2020). Part 1: The need for peer review in digital forensics. *Forensic Science International: Digital Investigation* 35, 301062. <https://doi.org/10.1016/j.fsidi.2020.301062>
- Sunde, N., Horsman, G. (2021). Part 2: The Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations. *Forensic Science International: Digital Investigation*, 36, 301074. <https://doi.org/10.1016/j.fsidi.2020.301074>

Refleksjonslogg - Fagfelle

Denne loggen vil være til hjelp for deg for å huske hva du har gjort, og for egen refleksjon og læring underveis. Den vil ikke bli samlet inn, men du vil bli oppfordret til å bruke denne som støttedokument under intervju eller spørreundersøkelse i prosjektet.

Her ønsker vi at du reflekterer over hvordan du la opp din fagfellevurdering, dine tanker om hvordan dette fungerte, valgene du tok underveis og hvilke læringspunkter du har med deg videre.

Fagfellevurdering 1:

Fagfellevurdering 2:

Osv.

Refleksjonslogg - Leder

Denne loggen vil være til hjelp for deg for å huske hva du har gjort, og for egen refleksjon og læring underveis. Den vil ikke bli samlet inn, men du vil bli oppfordret til å bruke denne som støttedokument under intervju eller spørreundersøkelse i prosjektet.

Vi ønsker at du reflekterer over valgene du tok når du organiserte fagfellevurderingene, altså hvordan du valgte ut saker til kvalitetskontroll, hva du la vekt på når du pekte ut fagfelle, erfaringer omkring samarbeid med andre distrikter (hvis dere gjennomførte dette), erfaringer med eventuelle uenigheter underveis og hvordan disse ble håndtert. Noter også læringspunkter du tar med deg videre.

Fagfellevurdering 1:

Fagfellevurdering 2:

Osv.

Vil du delta i forskningsprosjektet ”Kvalitetskontrollprosjektet – Fagfellevurdering av datatekniske rapporter”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å teste ut fagfellevurdering og kvalitetskontroll av datatekniske rapporter er egnet til å sikre god kvalitet ved å avdekke feil og villedende informasjon i rapportene samt å gi læring til de involverte (fagfelle/den som mottar tilbakemelding). I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Formålet med prosjektet er å samle inn erfaringer med kvalitetskontroll av datatekniske rapporter i henhold til Kvalitetskontrollhierarkiet og tilhørende sjekklister/støttedokumenter.

Politidistriktene/særorgan som ønsker å gjennomføre piloten vil organisere og gjennomføre kvalitetskontroll på minimum nivå 4 – Innholdskontroll i en periode fra ca. januar – mai 2023. På bakgrunn av denne erfaringen vil vi samle inn data for å belyse ulike aspekter av kvalitetskontroll og fagfellevurdering som omhandler ledelse og organisering av fagfellevurderinger, samt rollene som fagfelle og mottaker av tilbakemeldinger. Prosjektet vil belyse hvorvidt fagfellevurdering og kvalitetskontroll er egnet til sikre god kvalitet på datatekniske rapporter og hvorvidt tilbakemeldinger fra fagfeller gir læring. Datainnsamlingen vil resultere i en rapport som vil publiseres på PHS forskning – og som vil danne et kunnskapsgrunnlag for beslutninger om hvordan kvalitetskontroll av datatekniske rapporter skal gjennomføres i politiet. Forskningen vil også publiseres i vitenskapelige artikler som belyser ulike aspekter av forskningen.

Hvem er ansvarlig for forskningsprosjektet?

Politihøgskolen er ansvarlig for prosjektet.

Prosjektet ledes av Nina Sunde ved Politihøgskolen. Vitenskapelig ansvarlig er professor Johanne Yttri Dahl ved Politihøgskolen. Høgskolelektor Olav Dahl ved Norsk Brannskole og Rune Kenneth Bauge (Vest PD/ masterstudent ved NTNU) deltar også i forskningsprosjektet.

Hvorfor får du spørsmål om å delta?

Du forespørres om å delta fordi du har en av følgende roller:

- *Leder for enhet som utfører datatekniske undersøkelser*
- *Har en rolle som fagfelle i din enhet og har gjennomført fagfellevurdering datatekniske rapporter*
- *Har mottatt tilbakemelding fra fagfelle på datatekniske rapporter*

Hva innebærer det for deg å delta?

Datainnsamlingen vår vil blant annet være å gjøre intervju med deltakere. Intervjuet vil ta mellom 30-60 minutter og vil foregå på Teams. Spørsmålene vil handle om dine erfaringer med fagfellevurderingene som du har hatt en rolle i, og du vil få anledning til å støtte deg til refleksjonsnotater du har skrevet ned i perioden hvor pilotprosjektet ble gjennomført. Taushetspliktig informasjon vil ikke bli omhandlet i intervjuet.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrevet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Det er kun Nina Sunde, Olav Dahl og Rune Kenneth Bauge som vil ha tilgang til de innsamlede data.

For å sikre at uvedkommende får tilgang til dine personopplysninger vil ditt navn og kontaktopplysninger erstattes med en kode som lagres på en egen navneliste adskilt fra øvrige data fra intervjuene. Utskriftene fra intervjuet vil lagres i dokumenter på en passordbeskyttet mappe som kun Nina Sunde, Olav Dahl og Rune Kenneth Bauge har tilgang til.

I publikasjonene vil deltakerne i prosjektet omtales anonymt, og deltakerne skal ikke kunne gjenkjennes på bakgrunn av opplysninger som publiseres. Deltakere vil bli omtalt med alias f.eks. 'For eksempel Rapportskriver 1' 'Fagfelle 1' 'Leder 1', og ved gjengitte uttalelser vil det ikke refereres til hvilke politidistrikt/særorgan deltakeren er ansatt i.

Hva skjer med personopplysningene dine når forskningsprosjektet avsluttes?

Prosjektet vil etter planen avsluttes ved utgangen av 2023. Etter prosjektslutt vil datamaterialet med dine personopplysninger anonymiseres. Det innebærer at kun de anonymiserte transkripsjonene fra intervjuene vil lagres, mens opptak av samtalen og navnelister vil slettes. Siden datamaterialet vil være gjenstand for vitenskapelige publikasjoner vil det anonymiserte datamaterialet lagres til artiklene er publisert – og senest til utgangen av 2025.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Politihøgskolen har Personverntjenester vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke opplysninger vi behandler om deg, og å få utlevert en kopi av opplysningene
- å få rettet opplysninger om deg som er feil eller misvisende
- å få slettet personopplysninger om deg
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger

Hvis du har spørsmål til studien, eller ønsker å vite mer om eller benytte deg av dine rettigheter, ta kontakt med:

- Politihøgskolen ved Nina Sunde – nina.sunde@phs.no eller tlf 91660069
- Vårt personvernombud: Siri Mette Stenstuen Almås – siri.mette.stenstuen.almas@phs.no eller tlf. 90511703

Hvis du har spørsmål knyttet til Personverntjenester sin vurdering av prosjektet, kan du ta kontakt med:

- Personverntjenester på epost (personverntjenester@sikt.no) eller på telefon: 53 21 15 00.

Med vennlig hilsen

Prosjektansvarlig

Nina Sunde

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om Kvalitetskontrollprosjektet, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju
- Jeg samtykker til at mine personopplysninger behandles fram til prosjektet er avsluttet ved utgangen av 2023.
- at mine *anonymiserte* personopplysninger lagres etter prosjektslutt, til vitenskapelige artikler er publisert, senest 31.12.2025.

(Signert av prosjektdeltaker, dato)



POLITIHØGSKOLEN

Politihøgskolen
Slemdalsveien 5
Postboks 2109, Vika
0125 Oslo
Tlf: 23 19 99 00
www.phs.no

ISSN: 0807-1721
ISBN Elektronisk: 978-82-7808-197-6
PHS Forskning 2023: 7