STORTINGET

Stortingets utredningsseksjon

Til: (...)

Dato: 11.04.2024 Utreder: (...) Oppdragsnr: 2024035

Lovregulering av deepfakes

OPPDRAG

Ny teknologi som deepfakes gjør at omfanget av f.eks. krenkende, nedsettende, pornografiske bilder kan produserer enkelt. Dette har skapt en stor debatt i USA (https://www.nrk.no/kultur/ekspertar-trur-2024-blir-aret-for-_deepfakes_-1.16738596).

Jeg lurer på hvordan dette er regulert i Norge? Har det blitt vurdert lovregulert i større grad? Hvilke rettslige dilemmaer kan oppstå i regulering? Er dette regulert i andre land?

INNHOLD

1 Innledning

1.1 Om oppdraget. Presiseringer og definisjoner

Oppdraget reiser flere spørsmål knyttet til «deepfakes». Konkret er utredningsseksjonen bedt om å redegjøre for:

1) hvordan deepfakes er regulert i Norge,
2) om det har vært vurdert mer omfattende regulering av deepfakes,
3) hvilke rettslige dilemmaer som kan oppstå ved regulering av deepfakes, og
4) hvordan deepfakes er regulert i andre land

Med hensyn til spørsmål 4, har det i korrespondanse med oppdragsgiver blitt avklart at det er særlig interessant å få kartlagt hvilket arbeid USA, EU og de nordiske landene foretar på dette området. Vi forstår det slik at både eksisterende reguleringer og pågående lovarbeid er av interesse.

Når det gjelder det nærmere innholdet av begrepet deepfake, finnes det – så vidt vi vet – ingen legaldefinisjon i norsk rett. For denne utredningens formål har vi derfor lagt til grunn definisjonen som er benyttet i KI-forordning som nylig ble vedtatt av Europaparlamentet:

> *"'deep fake' means AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful."*[1]

Den vedtatte teksten er tilgjengelig på nettsidene til Europaparlamentet. Før den blir endelig vedtatt, må den gjennomgå en juridisk og språklig kontroll. I tillegg må den formelt også vedtas av Rådet, se pressemelding på nettsidene til Europaparlamentet. Se også informasjon på nettsidene til Europalov.

Når det gjelder spørsmålet om innlemmelse av KI-forordningen i EØS-avtalen, uttrykte Kommunal- og distriktsdepartementet i en foreløpig vurdering i 2021 at det opprinnelige forslaget fra Kommisjonen var EØS-relevant. Det var imidlertid for tidlig å si hvorvidt det er akseptabelt, se foreløpig posisjonsnotat av 21. juni 2021. Høsten 2023 uttrykte imidlertid tidligere kommunal- og distriktsminister Sigbjørn Gjelsvik at det var nedsatt en arbeidsgruppe som blant annet skulle lage en plan for rettidig og god innlemming av KI-forordningen i norsk lov så snart den er vedtatt i EU, se innlegg gjengitt på regjerningens nettsted.[2] I en nyhetsartikkel i Altinget 14. mars 2024 er det uttrykt at det i Norge «*... forberedes en rask implementering av EUs KI-lovgivning, så snart de siste detaljene er på plass*», se artikkelen «Verdens første KI-lov er vedtatt». Det vises her til at Digitaliserings- og forvaltningsminister Karianne O. Tung i et tidligere intervju uttrykte at dette er en viktig prioritering for regjeringen.

[1] Se artikkel 3 (60).
[2] I Innst. 151 S (2023-2024) er det av flertallet uttrykt at arbeidsgruppen avga sine anbefalinger november 2023. Vi er ikke kjent med anbefalingen.

1.2 Om besvarelsen

Oppdraget reiser vanskelige spørsmål på et område hvor det foreløpig er skrevet lite i juridisk litteratur. Så langt vi kan se, finnes det ingen utredninger som systematisk gjennomgår hvordan deepfake er regulert i norsk rett. Enkelte jurister har imidlertid gitt uttrykk for rettsoppfatninger i ulike nyhetsartikler. For å besvare spørsmålet om hvordan deepfakes er regulert i norsk rett, har vi søkt å kartlegge relevante bestemmelser i personvernlovgivningen, skadeerstatningsretten og strafferetten. Det finnes – så langt vi kan se – ingen reguleringer som spesifikt adresserer deep fakes. Ulike befatningsmåter med deepfakes vil likevel kunne være omfattet av krav i lovgivningen. Gitt utredningens rammer må vi ta forbehold om at relevante omstendigheter kan ha uteblitt fra vår fremstilling, eller gjengitt unøyaktig.

Når det gjelder relevante bestemmelser på EU-nivå, er det på oppdrag fra Europaparlamentets utredningstjeneste utarbeidet en større utredning:

- Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021)

Utredningen til Europaparlamentets utredningstjeneste inneholder en oversikt over det regulatoriske landskapet på EU-nivå. Vi har ikke rukket å foreta selvstendige søk i EU-rettens primærkilder, og vår fremstilling av relevante EU/EØS-bestemmelser er derfor i stor grad basert på denne utredningen.

Når det gjelder reguleringer i andre land, har vi i hovedsak basert oss på informasjon innhentet fra utredningstjenestene i andre lands parlamenter, se nærmere informasjon under punkt 5.1.

For øvrig redegjør vi for vår tilnærming til oppdragets spørsmål i tilknytning til de enkelte punktene.

2 Regulering av deepfake i Norge

2.1 Innledning

Fremstilling og bruk av deepfake-materiale er ikke spesifikt regulert i Norge. Ulike befatningsmåter ved deepfakes vil likevel kunne være omfattet av krav og begrensninger i ulike lovverk. For eksempel vil befatning med deepfakes kunne begrenses av krav i personopplysningslovgivningen, selv om deepfakes ikke er uttrykkelig adressert.

I det følgende gir vi en overordnet oversikt over relevante bestemmelser i personvernlovgivningen, menneskerettighetslovgivningen, erstatningsretten, straffelovgivningen og ulik lovgivning som gjennomfører krav i EØS-direktiver og -forordninger. Vi tar forbehold om at det kan være ytterligere lovbestemmelser med en side til deepfake-problematikken enn de som nevnes i det følgende.

Det er innledningsvis særlig grunn til å si noen ord om EUs KI-forordning. KI-forordningen inneholder regler som spesifikt adresserer deepfakes. Forordningen er nylig vedtatt av Europaparlamentet, men det gjenstår at den blir formelt vedtatt av Rådet. Overordnet innfører KI-forordningen gjennomsiktighetsbestemmelser med hensyn til deepfakes I Norge har departementet uttrykt at man planlegger for rask og rettidig implementering av KI-forordningen i norsk rett. I denne besvarelsen omtales KI-forordningen i **punkt 5.7.2**.

2.2 Personvernbestemmelser og retten til privatliv

2.2.1 Innledning

Fremstilling og bruk av deepfake-materiale kan tenkes å støte an mot lovbestemmelser om personvern og menneskerettigheter. Kjernen i disse vil regelmessig være vern av den enkeltes privatliv. I det følgende gir vi en oversikt over aktuelle skranker etter personopplysningsloven (herunder personvernforordningen), åndsverksloven § 104, EMK artikkel 8 og Grunnloven § 102, samt ulovfestede regler om personlighetsvern. Grovt sett kan dette i hovedsak betraktes som det sivilrettslige sporet – i motsetning til straffesporet, som vi kommer tilbake til senere.

2.2.2 Personopplysningsloven/personvernforordningen

Personvernforordningen (General Data Protection Regulation, «GDPR») fastsetter regler om vern av enkeltpersoner i forbindelse med behandling av personopplysninger. GDPR gjelder som norsk lov, jf. personopplysningsloven § 1.

GDPR stiller flere krav som den «behandlingsansvarlige» må oppfylle ved behandling av personopplysninger.[3] Begrepet «personopplysning» omfatter enhver opplysning om en identifisert eller identifiserbar fysisk person, jf. artikkel 4 nr. 1. Dette omfatter typisk navn, adresse, telefonnummer mv., men også f.eks. bilder og lydopptak. Med «behandling» menes enhver operasjon eller rekke av operasjoner som gjøres med personopplysninger, jf. artikkel 4 nr. 2. Dette omfatter f.eks. innsamling, lagring og forskjellige former for tilgjengeliggjøring.

I forbindelse med produksjon og annen befatning med deepfake-materiale vil det på flere stadier kunne gjøres bruk av opplysninger som kan knyttes til enkeltpersoner. Der et deepfake-bilde har til hensikt å avbilde en konkret person, vil bildet kunne anses som en personopplysning, se utredningen Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) s. 38:

> *"A deepfake that depicts a natural person can be considered personal data, since it relates to an identified or identifiable natural person."*

Der det skjer behandling av personopplysninger, vil GDPR som utgangspunkt komme til anvendelse. Den som befatter seg med deepfakes - enten det dreier seg om opprettelse, deling eller oppbevaring – vil dermed normalt bli behandlingsansvarlig etter forordningen.

Et praktisk viktig unntak er imidlertid at GDPR ikke gjelder for behandling av personopplysninger som utføres av en fysisk person som ledd i «rent personlige eller familiemessige aktiviteter», jf. artikkel 2 nr. 2 bokstav c. Ifølge fortalepunkt 18 innebærer personlige eller familiemessige aktiviteter en avgrensning mot aktiviteter som er knyttet til yrkes- eller forretningsvirksomhet. Det uttrykkes videre at personlige eller familiemessige aktiviteter kan omfatte «*aktiviteter på sosiale nettverk samt aktiviteter på internett i forbindelse med slike aktiviteter*».

Deepfakes har tidvis nettopp blitt spredt på slike sosiale nettverk. Fortalepunkt 18 kan trekke i retning av at slik deling ikke er omfattet av GDPR. Det kan likevel hevdes at deling som skjer på

[3] Mer presist gjelder forordningen for «helt eller delvis automatisert behandling av personopplysninger og på ikke-automatisert behandling av personopplysninger som inngår i eller skal inngå i et register», jf. artikkel 2.

åpent tilgjengelige nettsider og brukerkontoer må anses omfattet. Slik aktivitet kan neppe anses som rent personlig eller familiemessig, som er kjernen i unntaket.[4] At deling av personopplysninger på åpent tilgjengelige sider er omfattet av GDPR, ble bl.a. lagt til grunn av det såkalte Åpenhetsutvalget i NOU 2019: 10. Dette synes også å ha blitt lagt til grunn av det Det Europeiske Personvernråd (EDPB).[5] Til støtte for dette synspunktet viser både NOUen og EDPB til praksis fra EU-domstolen, som under det tidligere personverndirektivet la til grunn at personopplysninger som legges ut på åpent tilgjengelige sider var omfattet av direktivet.[6]

Vanskeligere spørsmål oppstår der hvor en brukerkonto på sosiale medier er lukket, men hvor brukeren likevel har en nokså stor mottakerkrets som får tilgang til det vedkommende deler. I juridisk teori er det argumentert for at også deling på slike kontoer kan anses omfattet av GDPR.[7] I NOU 2019: 10 uttrykte åpenhetsutvalget at det må bero på en konkret vurdering om personopplysningsloven/GDPR kommer til anvendelse i disse tilfellene. Blant annet ble følgende uttrykt:

> *«En kan dermed se for seg at på et sosialt medium, faller deling av personopplysninger som kan nå en stor mengde tilfeldige personer lettere inn under reguleringens anvendelsesområde enn deling til en mindre gruppe personer i ens egen husstand og/eller familie.»*[8]

Vi har imidlertid ikke sett noen autoritative kilder om problemstillingen under gjeldende forordning.

For øvrig drøftes ulike problemstillinger knyttet til GDPR og deepfakes i artikkelen Deepfakes: regulatory challenges for the synthetic society (Bart van der Sloot og Yvette Wagensveld, 2022). En av problemstillingene som drøftes, er nettopp GDPRs unntak for personlige eller familiemessige aktiviteter. I den forbindelse uttrykkes blant annet:

> *"The household exemption raises the following problem. Suppose an ex-partner stores private photographs of his ex-girlfriend on his computer, with which he then produces a deepfake video in which she performs all kinds of perverse sexual acts. He tells his friends about it, who also communicate this to her. This is just one of the many possible examples of deepfake applications that cannot be addressed under the GDPR. The production of compromising material and the possession of it, is not covered by the GDPR. Once the material is on the internet or distributed to large groups of friends it is, but by then it is too late."*[9]

Forfatterne viser for øvrig til at GDPR vil gjelde for de behandlingsansvarlige og databehandlerne som *stiller midler til rådighet* for behandling av personopplysninger i forbindelse med personlige

[4] Se Does the GDPR offer a solution to the 'problem' of sharenting? (Bessant/Schnebbe, 2022). Se dessuten Barns rett til beskyttelse mot at foreldrene offentliggjør personlig informasjon om dem på sosiale medier – Er barn tilstrekkelig beskyttet? (Nina Tøgersen Allstrin, 2016) punkt 3.3 (som riktignok gjelder den tilsvarende formuleringen under det tidligere personverndirektivet).
[5] Se henholdsvis NOU 2019: 10 punkt 5.1.4.2 – 5.1.4.4 og Guidelines 3/2019 on processing of personal data through video devices (EDPB, 2020) s. 7-8.
[6] Se C-101/01 (Lindqvist).
[7] Se *Does the GDPR offer a solution to the 'problem' of sharenting? (Bessant/Schnebbe, 2022)* punkt 3.4.
[8] Se NOU 2019: 10 punkt 5.1.4.4.
[9] Se punkt 3.

eller familiemessige aktiviteter, jf. artikkel 18 tredje punktum. Det er etter vårt syn noe uklart hva forfatterne mener dette innebærer. Der et deepfake-bilde lastes opp på et lukket nettverk på en sosial medieplattform uten samtykke fra den avbildede, kan det imidlertid drøftes om den avbildede har noen rettigheter *overfor plattformen* (som altså stiller midlene til rådighet). Dette synes det i alle tilfelle å være usikkerhet knyttet til.[10]

Samlet sett tilsier rettskildebildet at GDPR får anvendelse på deling av deepfakes på åpent tilgjengelige sider, samt muligens også på deling som skjer på lukkede kontoer dersom mottakerkretsen er tilstrekkelig stor. Kommersiell deling vil i alle tilfelle være omfattet. Deling som skjer i den private sfære, vil derimot neppe være omfattet av GDPR. Grensene for hva som skal anses som den private sfære er imidlertid uklare.

Forutsatt at GDPR kommer til anvendelse, er det flere krav som må overholdes av den behandlingsansvarlige for at behandlingen skal være lovlig. Sentralt er at behandling av personopplysninger krever et *behandlingsgrunnlag*. I artikkel 6 er det angitt seks mulige behandlingsgrunnlag i bokstav a til f. I Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) er det lagt til grunn at det som regel kun er to behandlingsgrunnlag som er aktuelle i forbindelse med behandling av deepfake-materiale, se s. 39:

1. informert samtykke fra den det gjelder, jf. bokstav a
2. berettiget interesse, jf. bokstav f

Behandlingsgrunnlaget «berettiget interesse» krever at behandlingen er nødvendig for å ivareta en legitim interesse som veier tyngre enn hensynet til den enkeltes personvern. Som et eksempel på når dette vil kunne være tilfelle, nevner utredningen ironiske deepfake-fremstillinger av kjente personer. I slike tilfeller kan man tenke seg at ytringsfriheten må veie tyngst, når formålet er satire eller å ytre en politisk kommentar.

Der hvor det ikke foreligger berettigede interesser, vil deling av deepfakes være betinget av samtykke fra den det gjelder. Vi antar dette ikke spesielt praktisk i denne konteksten.

I utredningen fra Europaparlamentet er det vist til at GDPR gir individer rett til å få *korrigert* personopplysninger som ikke stemmer, og også til å få opplysninger om seg *slettet.* Alle medlemsland skal også ha et tilsyn som sørger for at reglene håndheves, Det kan likevel, bemerkes det, være vanskelig for et offer for deepfake-misbruk å håndheve sine rettigheter etter GDRP. Det er også krevende for tilsynsmyndigheten å følge opp dette. Gjerningspersonene er ofte anonyme, og dermed vanskelig å identifisere. For den enkelte vil det kunne være ressurskrevende å sette i verk en sivilrettslig prosedyre.[11]

I Norge er det Datatilsynet som fører kontroll med at personvernregelverket etterleves. Datatilsynet skal medvirke til at enkeltpersoner ikke blir krenket gjennom bruk av opplysninger som kan knyttes

[10] Forfatterne viser til at enkelte datatilsyn vil avvise klager om deling av personopplysninger på sosiale nettverk: *"However, data protection authorities are also experiencing an increasing number of complaints emanating from individuals' personal use of the internet. A typical complaint might be that a pupil has used a social networking site to say post a derogatory, inaccurate or hurtful message about a teacher. Currently some data protection authorities would reject any complaints about the pupil on the grounds that the processing of personal data involved would fall within the personal or household processing exemption."*
[11] Se utredningen s. 39.

til dem. Det synes likevel som at Datatilsynet relativt sjeldent sanksjonerer privatpersoner for krenkelser, se NOU 2019: 10 punkt 5.1.4.4:

> *«I Norge har Datatilsynet fram til nå vært tilbakeholdne med å sanksjonere privatpersoner som bryter personopplysningsloven.»*

Det er ikke angitt hva som er grunnen til dette. Vi antar at det kan ha sammenheng med ressurser, prioriteringer, og/eller at grensene for GDPRs anvendelse på privatpersoner kan være uklare.

For øvrig nevner vi at Norsk senter for informasjonssikring (NorSIS) i 2012 overtok ansvaret for tjenesten "slettmeg.no" fra Datatilsynet, bl.a. under henvisning til at verktøyene som ligger i personopplysningsloven og straffeloven ikke alltid er tilstrekkelig.[12] Tjenesten er utpreget individrettet.

2.2.3 Åndsverksloven § 104

Åndsverksloven § 104 gir privatpersoner en såkalt «rett til eget bilde». Spørsmålet som behandles i det følgende, er om deepfake-bilder er omfattet av denne retten.

Ifølge åndsverkloven § 104 kan ikke fotografi som avbilder en person, gjengis eller vises offentlig uten samtykke av den som er avbildet. Unntak gjelder i følgende tilfeller:

a. avbildningen har aktuell og allmenn interesse
b. avbildningen av personen er mindre viktig enn hovedinnholdet i bildet
c. bildet gjengir forsamlinger, folketog i friluft eller forhold eller hendelser som har allmenn interesse
d. eksemplar av avbildningen på vanlig måte vises som reklame for fotografens virksomhet og den avbildede ikke nedlegger forbud
e. bildet brukes som omhandlet i § 33 andre ledd eller § 37 tredje ledd

Bestemmelsen regnes for å være en personvernbestemmelse, selv om den står i åndsverksloven. I den forbindelse er det sentralt at det etter bestemmelsen ikke er noe krav om originalitet/verkshøyde, slik det ellers er for å ha vern som *åndsverk* etter loven. Det er tilstrekkelig at det dreier seg om et fotografi som avbilder en person.

Når det nærmere gjelder innholdet av begrepet «fotografi», er det gitt følgende definisjon i åndsverksloven § 23 første ledd annet punktum:

> *«Med fotografisk bilde menes bilde som er frembrakt ved bruk av kamera eller ved annen teknikk som kan likestilles med fotografering.»*

I forarbeidene til åndsverksloven er det forklart at det avgjørende er om *«... et bilde er frembrakt på en bestemt måte, ved tradisjonell fotografering gjennom belysning av lysfølsom film eller måter som «likjest fotografering», jf. tidligere § 1 i fotografiloven. Dette omfatter bl.a fotografering ved bruk av digital teknologi, og avbildninger frembrakt med annen teknikk ved tilførsel av lys og annen*

[12] Se artikkel på følgende lenke: https://www.an.no/nyheter/slettmeg-no-flyttes-fra-datatilsynet/s/1-33-5876104.

stråleenergi. Visuelle arbeider som er fremstilt rent elektronisk, f.eks. med databasert tegne- og designsystemer, omfattes dermed ikke.»[13]

For øvrig omfatter fotografibegrepet også film.[14]

Samlet sett er det avgjørende for vern etter åndsverksloven § 104 at bildet er fremstilt ved fotografisk eller lignende teknikk. Foruten bilder tatt med tradisjonelt fotoapparat og digitalkamera, omfattes bilder som på andre måter er frembrakt ved tilførsel av lys eller annen stråleenergi.[15] Så lenge et bilde er fremstilt på denne måten, vil vernet – så vidt vi forstår – også gjelde der det eventuelt foretas digital bearbeidelse eller manipulasjoner av bildet.[16]

Mot denne bakgrunn er spørsmålet mer konkret om deepfake-bilder kan sies å være produsert ved fotografisk eller lignende teknikk, slik at de er omfattet av vernet i åndsverksloven § 104. Som drøftelsen under vil vise, er det etter vår oppfatning ikke et åpenbart svar på dette spørsmålet.

Produksjon av deepfake-materiale synes å bero på nokså avanserte teknikker som vi har noe begrensede forutsetninger for å sette oss inn i. Basert på ulike fremstillinger virker det imidlertid som at det eksisterer flere måter å produsere slikt materiale på.[17] Det kan derfor tenkes at spørsmålet om deepfake-materiale rammes av åndsverkloven § 104 vil avhenge av hvilken teknikk som er benyttet:

- Dersom et deepfake-bilde er produsert ved en teknikk hvor endringer er gjort direkte i et bilde som er fremstilt ved fotografisk eller lignende teknikk, vil det manipulerte bildet etter vår oppfatning fortsatt være vernet § 104. Selv om det er foretatt en bearbeidelse/manipulasjon, må det vernede element – avbildningen av en person – fortsatt sies å være frembrakt ved «tilførsel av lys eller annen stråleenergi», jf. over.

- Dersom et deepfake-bilde i stedet er en *helt ny konstruksjon* basert på informasjon hentet fra en serie fotografier, er det etter vår oppfatning mer usikkert om vernet i åndsverksloven § 104 gjelder. Det kan hevdes at et slikt bilde ikke er frembrakt ved «tilførsel av lys eller annen stråleenergi», men at det ved denne prosessen i prinsippet er skapt et *helt nytt bilde* basert på *sammenstilling og tolkning av informasjon* fra en rekke bilder.

Vårt inntrykk er at produksjon av deepfakes ofte er basert på at personer og steder rekonstrueres i helt nye bilder. I ulike kilder er det forklart at deepfakes lages ved at AI-programvare forsynes med en rekke bilder av personene man ønsker å lage en manipulert fremstilling av.[18] AI-programvaren

---

[13] Se Prop. 204 L (2016-2017) punkt 4.7.1, hvor denne definisjonen er gitt i tilknytning til omtalen av gjeldende rett *før* vedtakelsen av loven. Det fremgår imidlertid at det ikke har vært meningen å endre denne definisjonen, se punkt 14 merknad til § 23.
[14] Se Åndsverksloven med kommentarer, Haakon Aakre m.fl. (2021), s. 499.
[15] Se Åndsverksloven med kommentarer, Haakon Aakre m.fl. (2021), s. 155.
[16] Sml. straffeloven § 267 a
[17] Se f.eks. artikkelen DeepFake: A Deep Learning Approach in Artificial Content Generation (Brij Grupta, 2023) og forskningsartikkelen Deepfakes Generation and Detection: State-of-the-art, open challenges, countermeasures, and way forward (Momina masood m. fl., 2021).
[18] Vår korte redegjørelse er basert på følgende artikler med utfyllende forklaringer: Deepfakes Generation and Detection: State-of-the-art, open challenges, countermeasures, and way forward (Momina Masood m. fl., 2021), Overview Of How To Create Deepfakes - It's Scarily Simple (Lutz Finger, artikkel i Forbes 2022) og Understanding the Technology Behind DeepFakes (Alan Zucconi, 2018).

benytter disse bildene til å kartlegge grunntrekkene i ansiktene og samler denne informasjonen i såkalte «latent representations»/»latent images». Basert på informasjonen i de latente representasjonene «rekonstrueres» de aktuelle personene i nye bilder. Ved hjelp av denne teknologien kan man fremstille bilder hvor personer får andre ansiktsuttrykk enn de opprinnelig hadde, samt sette dem i kontekster de aldri har vært i. Man kan også lage videoer hvor det fremstår som en person sier eller gjør noe vedkommende aldri har sagt eller gjort.

Når et bilde er digitalt konstruert basert på informasjon avledet av et eller flere andre bilder, kan det – som nevnt – hevdes at det ikke er fremstilt ved tilførsel av lys eller annen stråleenergi. Det fremstår som en nærliggende å anse bildet for å være fremstilt elektronisk. Ifølge forarbeidene faller nettopp bilder som er fremstilt «rent elektronisk» utenfor vernet i åndsverkloven § 104, jf. Prop. 204 L (2016-2017) punkt 4.7.1.

På den annen side må retten til eget bilde også kopier av et fotografi.[19] Det kan hevdes at deepfake-rekonstruksjoner ikke bør bedømmes annerledes, og at det avgjørende må være om det er benyttet et underlagsmateriale som er fremstilt ved fotografisk eller lignende teknikk. Det kan i den forbindelse muligens anføres at et deepfake-bilde ikke er fremstilt «rent elektronisk», ettersom det er benyttet informasjon fra fotografiske bilder. Likevel virker det å være en viss forskjell på *å kopiere* et fotografisk bilde og å *rekonstruere et nytt bilde* basert på informasjon avledet av en serie bilder.

Et særlig moment er dessuten at brudd på åndsverkloven § 104 er *straffbart*, jf. § 79. Det kan derfor stilles spørsmål ved om åndsverkloven § 104 i det hele tatt er klart nok utformet til å kunne ramme deepfake-bilder rent strafferettslig.

Klarhetskravet på strafferettens område følger av legalitetsprinsippet, jf. Grunnloven § 96. I juridisk teori er det fremholdt følgende om legalitetsprinsippets innhold:

> *«Foruten fordringen om at straff må forankres i en lovbestemmelse, stiller prinsippet innholdsmessige krav til straffelovgivningen: straffebudene kan verken være for upresist formulert (klarhetskravet) eller anvendes på tilfeller som ligger for fjernt fra ordlyden (analogiforbudet).»*[20]

For spørsmålet om åndsverkloven § 104 er klart nok utformet til å ramme deepfake-bilder, synes Høyesteretts avgjørelse i Rt. 2012 s. 1211 å være relevant. Avgjørelsen er et eksempel på at anvendt teknologi ble tillagt avgjørende betydning ved vurderingen av om en handling var straffbar.

Konkret dreide saken seg om hvorvidt oppfordringer til drap på polititjenestemenn *ytret på en blogg* kunne straffes etter den tidligere straffeloven § 140. Bestemmelsen satt straff for den som offentlig oppfordrer til straffbar handling. Begrepet «offentlig» var nærmere definert i tidligere straffelov § 7 nr. 2. Det aktuelle alternativet å vurdere forholdet etter var «Udgivelse af trykt Skrift». Trykt skrift omfattet skrift «*der mangfoldiggjøres ved Trykken eller paa anden kemisk eller mekanisk Maade*»,

---

[19] Sml. følgende forarbeidsuttalelser i Ot.prp. nr. 54 (1994-1995) punkt 3.1.2.2, hvor det i tilknytning til fotografens enerett til sine bilder uttrykkes: «*På den annen side vil fysiske eksemplar fremstilt på grunnlag av et fotografiske bilde som er skannet inn eller lagret digitalt, fortsatt være et eksemplar av det fotografiske bilde*».

[20] Se artikkelen Nyere praksis om det strafferettslige legalitetsprinsippet (Thomas Frøberg, 2015), med videre henvisning til Asbjørn Strandbakken.

jf. tidligere straffelov § 10. Mot denne bakgrunn var spørsmålet om ytringer på en blogg kunne sies å være mangfoldiggjort ved trykking «eller på anden kemisk eller mekanisk Maade».

Til tross for at handlingene måtte anses som straffverdige, kom Høyesteretts flertall til at ytringene på bloggen ikke var omfattet av straffebudet. Det ble vist til at elektronisk formidling ikke var dekket av definisjonen på «trykt Skrift», se premiss 19. Høyesterett uttrykte videre i premiss 22:

> *«Slik flertallet ser det, omfatter siktelsen handlinger som klart er straffverdige. Formålsbetraktninger tilsier at ytringer på internett likestilles med dem som fremsettes i trykt skrift, og det kan ikke være tvilsomt at lovgiver ønsker å ramme slike forhold. Straffbarheten må imidlertid følge av loven. Forholdet i siktelsen omfattes ikke av ordlyden i straffeloven § 140 jf. § 7 nr. 2 og § 10, og lovgiver har i andre sammenhenger lagt til grunn at formidling på internett faller utenfor definisjonen av trykt skrift. Klarhetskravet etter Grunnloven § 96 og EMK artikkel 7 er ikke tilfredsstilt, og forholdet i siktelsen er dermed etter flertallets vurdering ikke straffbart.»*

Med hensyn til deepfake-bilder kan det på lignende vis reises spørsmål ved om åndsverkloven § 104 er tilstrekkelig klart utformet til å ramme bilder som er digitalt konstruert basert på informasjon hentet fra andre bilder, gitt forarbeidenes betoning av at bildene må være fremstilt på en bestemt måte: ved tilførsel av lys.

Vi bemerker imidlertid at det er visse forskjeller på deepfakes og saksforholdet i Høyesterett. Der § 140 i tidligere straffelov nokså klart ikke omfattet elektronisk formidling, er det ikke like åpenbart at åndsverkloven § 104 ikke omfatter deepfake-bilder konstruert elektronisk basert på informasjon fra andre bilder. Det er, som nevnt, mulig å forstå åndsverkloven § 104 slik at bestemmelsen omfatter bilder hvor fotografier har inngått som en del av produksjonsprosessen, selv om det endelige resultat ikke kan sies å være *direkte* frembrakt ved fotografisk eller lignende teknikk.

Samlet sett fremstår det for oss som noe usikkert om åndsverkloven får anvendelse på deepfake-materiale der dette er bilder konstruert digitalt basert på informasjon fra andre bilder. Vi har i den forbindelse registrert professor Olav Torvund har uttrykt at **straffeloven § 267 a** – som setter straff for deling av krenkende bilder – ikke synes å ramme fiktive bilder.[21] Ifølge forarbeidene gjelder denne bestemmelsen for fotografiske bilder, slik som åndsverkloven § 104. Samtidig ser Justis- og beredskapsdepartementet ut til å legge til grunn det motsatte standpunkt.[22] Vi kommer nærmere tilbake til straffeloven § 267 a under.

Forutsatt at åndsverkloven kommer til anvendelse, ligger det en begrensning i vernet ved at dette etter loven kun gjelder med hensyn til offentlig gjengivelse eller visning. I juridisk teori er det fremhevet at offentlig gjengivelse eller visning *«…omfatter enhver form for rådighet over fotografiet som medfører at det gjøres tilgjengelig for allmenheten, for eksempel ved tilgjengeliggjøring over internett eller andre digitale nettverk…»*[23] Bestemmelsen hindrer dermed ikke produksjon og besittelse av fotografier, og heller ikke deling av fotografier innenfor privat sfære. Denne begrensningen ligner på privat-unntaket i GDPR. Deling av slikt materiale kan naturligvis oppleves

[21] se NRK-artikkelen «Mener det haster med ny lov om falske nakenbilder» (publisert 23. mars 2024). Dette er nærmere omtalt i vår omtale av straffeloven § 267 a i utredningen **punkt 2.4.2.**
[22] Ibid.
[23] Se Åndsverksloven med kommentarer, Haakon Aakre m.fl. (2021), s. 499.

belastende også der delingen skjer innenfor den private sfære. I slike tilfeller vil imidlertid straffeloven § 267 a fortsatt kunne komme til anvendelse, forutsatt at vilkårene der er oppfylt, se under.

2.2.4 Bestemmelser i menneskerettighetslovgivningen

2.2.4.1 EMK artikkel 8

Den europeiske menneskerettighetskonvensjonen (EMK) artikkel 8 gir rett til respekt for privatliv og familieliv. I praksis er det slått fast at bestemmelsen gir rettigheter knyttet til eget bilde, hvilket gjør det relevant å undersøke om dette omfatter deepfakes. Dersom bestemmelsen gir rettigheter i denne sammenheng, er det et eget spørsmål hvordan disse kan håndheves.

EMK artikkel 8 har følgende ordlyd:

> *"1. Everyone has the right to respect for his private and family life, his home and his correspondence.*
>
> *2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."*

EMK gjelder som norsk lov, jf. menneskerettsloven § 2.

Basert på artikkel 8 har den europeiske menneskerettsdomstolen (EMD) i flere saker lagt til grunn at enkeltpersoner har rettigheter knyttet til eget bilde. En oversikt over denne praksisen er gitt i følgende faktaark utarbeidet av EMD:

- Right to the protection of one's image – factsheet (European Court of Human Rights, Press Unit, 2023)

I en av de grunnleggende sakene, Reklos and Davourlis v. Greece, no. 1234/05, er det uttrykt følgende om innholdet av retten til eget bilde (premiss 40):

> *"A person's image constitutes one of the chief attributes of his or her personality, as it reveals the person's unique characteristics and distinguishes the person from his or her peers. The right to the protection of one's image is thus one of the essential components of personal development and presupposes the right to control the use of that image. Whilst in most cases the right to control such use involves the possibility for an individual to refuse publication of his or her image, it also covers the individual's right to object to the recording, conservation and reproduction of the image by another person. As a person's image is one of the characteristics attached to his or her personality, its effective protection presupposes, in principle and in circumstances such as those of the present case (see paragraph 37 above), obtaining the consent of the person concerned at the time the picture is taken and not simply if and when it is published. Otherwise an essential attribute of personality would be retained in the hands of a third party and the person concerned would have no control over any subsequent use of the image."*

EMD slår altså fast at vernet for eget bilde er en av de essensielle komponentene i ens personlige utvikling. Dette forutsetter en rett til å kontrollere bruken av eget bilde. Det slås videre fast at denne retten, i de fleste tilfeller, gjelder adgangen til å nekte publisering av eget bilde, men at den også omfatter rett til å nekte andre personer å ta, lagre og reprodusere bildet. Samtykke må derfor innhentes, og det må i prinsippet innhentes å på tidspunktet hvor bildet blir tatt (ikke bare før det publiseres).

Retten til eget bilde må balanseres mot andre rettigheter, herunder ytringsfriheten, se Von Hannover v. Germany (No. 2), nos. 40660/08 and 60641/08 premiss 100 og HR-2009-547-A.

I utredningen Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) er det bemerket at retten til eget bilde etter EMK artikkel 8 ikke bare gir beskyttelse for portrett, fotografi eller videoer som avbilder en person, men også for «likeness» eller «resemblance of a person», se s. 40. Vi har ikke gjenfunnet disse formuleringene i EMDs praksis, og det er derfor uklart hva de bygger på. I det hele tatt har vi ikke funnet en uttrykkelig definisjon av «image» i EMDs praksis. Det kan imidlertid være at utredningen bygger på EMDs uttalelser om at et bilde inneholder en persons unike karakteristikker og adskiller vedkommende fra andre. Det kan jo et bilde gjøre, selv om det ikke er fremstilt ved fotografisk eller lignende teknikk.

Dersom retten til eget bilde etter EMK artikkel 8 er å forstå såpass vidt at det omfatter «likeness» og «resemblance», vil det – så vidt vi forstår – omfatte deepfake-materiale uavhengig av fremstillingsmetode, se vår drøftelse i **punkt 2.2.3**. Dette tilsier at produksjon og bruk av deepfakes som utgangspunkt er omfattet av krav om samtykke av den avbildede etter EMK, i hvert fall der det er tale om produksjon og bruk med sikte på offentliggjøring. I andre tilfeller er det litt mer uklart. I tillegg kommer, som nevnt, at det må foretas en avveining mot andre rettigheter.

Se for øvrig veiledning til EMK artikkel 8 s. 49.

Forutsatt at EMK artikkel 8 gir rettigheter knyttet til personbilder fremstilt ved deepfake-teknologi, oppstår spørsmålet om hvordan disse rettighetene kan håndheves. Utgangspunktet er at forpliktelsene etter EMK påhviler staten. Staten har både en negativ plikt til å avstå fra menneskerettighetsbrudd *og* en positiv plikt til å sikre menneskerettighetene. Dette innebærer at man ikke kan anlegge sak mot andre privatpersoner med påstand om at disse har krenket EMK.

Mellom private parter kan imidlertid EMK få *indirekte betydning* når andre reglers innhold skal tolkes og fastlegges. Dette gjelder f.eks. for skadeerstatningsregler. Med hensyn til krenkende deepfakes fremstår det som nærliggende at sak om erstatningskrav etter omstendighetene kan anlegges basert på skadeerstatningsloven § 3-5 (erstatning for krenking av privatlivets fred) og § 3-6 a. (erstatning for ærekrenkelser). Ved fastleggingen av disse bestemmelsenes innhold vil domstolen se hen til hvilke krav som kan utledes av EMK artikkel 8.[24] Å anlegge sivilt søksmål vil imidlertid kunne representere en faktisk, rettslig og /eller ressursmessig barriere.

Et spørsmål for seg er om domstolen kan tilstå individer rettigheter med virkning overfor andre individer *direkte basert* på EMK artikkel 8. Nyere rettspraksis kan tyde på det. I HR-2022-847-A kom Høyesterett til at en mann hadde rett til – og skulle ha – samvær med et barn han ikke var

[24] Med hensyn til skadeserstatningsloven § 3-5, se som eksempel Rt. 2008 s. 1089.

biologisk far til. Dette ble forankret direkte i EMK artikkel 8.[25] Tilsvarende kan man kanskje tenke seg at EMK artikkel 8 kan gi grunnlag for en rett til å kreve at andre sletter deepfakes (og lignende). Grensene for når og hvordan EMK artikkel 8 kan gis virkning mellom private fremstår imidlertid som uklare, og vi behandler ikke spørsmålet nærmere.

2.2.4.2 Grunnloven § 102

Retten til respekt for privatliv ble inntatt i Grunnloven § 102 ved grunnlovsrevisjonen i 2014, etter mønster fra EMK artikkel 8.[26] Grunnloven § 102 har følgende ordlyd:

> *«§ 102.*
>
> *Enhver har rett til respekt for sitt privatliv og familieliv, sitt hjem og sin kommunikasjon. Husransakelse må ikke finne sted, unntatt i kriminelle tilfeller.*
>
> *Statens myndigheter skal sikre et vern om den personlige integritet.»*

Om tolkningen av bestemmelsen, uttaler førstvoterende i Rt. 2015 s. 93 at § 102 «skal tolkes i lys av de folkerettslige forbildene», men med det forbehold at det er Høyesterett som har ansvaret for å tolke, avklare og utvikle Grunnlovens rettighetsbestemmelser.[27] Internasjonal praksis om de folkerettslige rettighetsbestemmelsene binder følgelig ikke Høyesteretts tolkning av Grunnlovens rettighetsbestemmelser, men tolkningen vil tillegges vekt. Betraktningene omtalt overfor om hvordan EMK artikkel 8 kan få betydning for deepfakes vil formodentlig også gjelde etter Grunnloven.

Forarbeidene til grunnlovsrevisjonen påpeker særlig behovet for personvernrettigheter på grunnlovsnivå av hensyn til fremtidig teknologisk utvikling, hvor den rettslige reguleringen har blitt hengende etter.[28] Utvalget påpeker at grunnlovsbestemmelsen kan være en viktig rettesnor ved tolkningen av lovverket i situasjoner som ikke var forutsett da lovgivningen ble vedtatt.[29] Kunstig intelligens må klart sies å representere en slik teknologi. Grunnlovsendringen endret derimot ikke den daværende materielle rettstilstanden, utover å grunnlovsfeste det vernet som tidligere gjaldt samlet etter ulovfestet rett, menneskerettsloven, daværende § 102 i Grunnloven og annen ordinær lovgivning.[30]

Utvalget påpeker at brudd på privatlivets fred kan være en trussel for demokratiet, ved å begrense den meningsutveksling demokratiet bygger på.[31] Utvalget trekker frem situasjoner der offentlige myndigheter misbruker personopplysninger, blant annet at forvaltningens håndtering av personopplysninger oppleves som en «straff». I likhet med EMK artikkel 8 kan det likevel ikke utelukkes at bestemmelsen også kan få betydning for forholdet mellom private, enten ved tolkningen av andre bestemmelser eller indirekte via statens positive forpliktelse til å sikre

[25] Avgjørelsen er nærmere omtalt i Tomas Midttun Tobiassen, Når domstolen griper inn (2023), hvor det innledningsvis uttrykkes at *«Flere rettigheter i Grunnloven, EMK og andre konvensjoner kan i dag anvendes mellom private (horisontal virkning)».*
[26] Stub (2021) i *Grunnloven historisk kommentarutgave*, s. 1143 og Dok.nr. 16 (2010-2011) s. 175.
[27] Rt. 2015 s. 93 avsnitt 57.
[28] Dok.nr. 16 (2010-2011) s. 176.
[29] Dok.nr. 16 (2010-2011) s. 176.
[30] Dok.nr. 16 (2010-2011) s. 178
[31] Dok.nr. 16 (2010-2011) s. 176.

rettigheten.[32] Et typisk tilfelle som nevnes er der hvor rettighetene mellom to private må balanseres, er retten til privatliv og ytringsfriheten. På strafferettens område vil bestemmelsen ha begrenset betydning, grunnet strafferettens strenge legalitetsprinsipp, forankret i Grunnloven § 96.

2.2.5 Ulovfestet vern av personligheten

Vi nevner til slutt at det i rettspraksis er utviklet et ulovfestet vern for «personligheten». Det vises gjerne til Rt. 1952 s. 1217 (To mistenkelige personer) som grunnlag for dette vernet.[33]

Vernets nærmere innhold og status i dag er uklart. At det eksisterer et ulovfestet vern, har imidlertid blitt lagt til grunn i rettspraksis så sent som i 2009. I Rt. 2009 s. 1568 fant Høyesterett at en kjent amerikansk snøbrettkjører hadde krav på kompensasjon for kommersiell bruk av et bilde av ham. Dette ble forankret det ulovfestede personlighetsvernet. Det samme ble lagt til grunn i RG 1999 s. 1009. Saken dreide seg om en fotobutikk som hadde laget en radioreklame hvor stemmen til en kjent filmjournalist var etterlignet. Dette fant retten at var i strid med det ulovfestede personlighetsvernet. Lagmannsretten påpekte at det ulovfestede vernet gikk utover beskyttelsen som fulgte av daværende bestemmelser om ærekrenkelse, krenkelse av privatlivets fred og misbruk av åndsverk og fotografier. Ifølge lagmannsretten måtte det foretas en interesseavveining. Fotobutikkens interesser var av utelukkende kommersiell karakter. Reklameinnslaget derfor var et rettstridig inngrep overfor journalisten.

Det har også blitt vist til det ulovfestede personlighetsvernet i lovforarbeider. I Prop. 68 L (2015-2016) omtaler departementet ulike regler knyttet til beskyttelse av personvernet, og viser blant annet til det allmenne vernet som gjelder på ulovfestet grunnlag.[34]

I juridisk teori har Are Stenvik undersøkt vernets innhold i en noe eldre artikkel.[35] Artikkelen tar særlig for seg om det kan oppstilles et vern mot utnyttelse av personlig særpreg. Stenvik omtaler imidlertid kort at vernet kan tenkes aktuelt under kategoriene «inngrep i privatsfæren» og «offentlige meddelelse om personlig forhold», se artikkelen punkt 2.1 og 2.3 henholdsvis. Når det gjelder kategorien «ytringer som setter andre personer i et falskt lys», jf. punkt 2.2 i artikkelen, fremholder Stenvik at lovgivningen i stor grad har vært dekkende. Vi forstår det slik at det på dette området ikke har vært noe særlig behov for supplering ved det ulovfestede vernet. I lys av utviklingen av deepfakes, som er særlig egnet til å sette andre i et falskt lys, kan det muligens diskuteres om det i dag er et større rom for den ulovfestede personlighetslæren også på dette området.

Når det gjelder utnyttelse av personlig særpreg for kommersielle formål, uttrykker Stenvik følgende:

> *«De beste grunner synes å tale for å oppstille en regel om at enhver i utgangspunktet råder over den kommersielle utnyttelsen av sin fremtoning, sine symboler og sine prestasjoner. Vernet bør omfatte også etterligninger, forutsatt at etterligningen er egnet til å skape den forestilling hos betrakterne at det dreier seg om personen selv, eller dersom det dreier seg om en urimelig utnyttelse eller forringelse av den økonomiske verdien som knytter seg til et personlig særpreg.»*

[32] Se Stub (2021) s. 1144.
[33] Are Stenvik (2003) «Rettsbeskyttelse for personlig særpreg» s. 603.
[34] Prop. 68 L (2012-2016) punkt 3.2.
[35] Are Stenvik (2003) «Rettsbeskyttelse for personlig særpreg» s. 603.

For øvrig kan det være grunn til å anta at reguleringen av personvernet i menneskerettsloven og Grunnloven generelt har redusert det praktiske virkeområdet for den ulovfestede læren.[36] Den raske utviklingen av KI og deepfakes kan imidlertid gi grunn til å drøfte det ulovfestede personlighetsvernet på nytt. Spørsmålet er både usikkert og uavklart, og vi går derfor ikke nærmere inn på dette.

2.3 Bestemmelser i skadeserstatningsloven

2.3.1 Ærekrenkelse i skadeserstatningsloven

Skadeserstatningsloven § 3-6 a gir adgang til å fremme sivilrettslig erstatningskrav for ærekrenkelser. For den som utsettes for en ulovlig deepfake vil dette kunne være en sivilrettslig mulighet for forfølgelse. Ærekrenkelser er ikke straffbart etter straffeloven av 2005.[37]

Skadeserstatningsloven § 3-6 a lyder som følger:

> *§ 3-6 a.(erstatning for ærekrenkelser)*
>
> *Den som uaktsomt har satt frem en ytring som er egnet til å krenke en annens æreføIelse eller omdømme, skal yte erstatning for den lidte skade og slik erstatning for tap i fremtidig erverv som retten ut fra den utviste skyld og forholdene ellers finner rimelig. Han kan også pålegges å betale slik erstatning (oppreisning) for skade av ikke-økonomisk art som retten finner rimelig. Dersom den krenkede døde mindre enn 15 år før krenkelsen etter første ledd fant sted, kan krav om oppreisning settes frem av hans nærmeste.*
>
> *En ærekrenkende ytring medfører ikke ansvar etter første ledd dersom den anses berettiget etter en avveining av de hensyn som begrunner ytringsfrihet. Ved denne vurderingen skal det særlig legges vekt på om ytringen hviler på et fyldestgjørende faktisk grunnlag, på ytringens grad av krenkelse, og om hensynet til den krenkede er tilfredsstillende ivaretatt ved for eksempel adgang til imøtegåelse, om allmenne interesser eller andre gode grunner tilsa at den ble satt frem, og om ytreren har vært i aktsom god tro med hensyn til de momenter som kan gjøre ytringen berettiget*

Det er ikke krav om at ytringen faktisk har krenket en persons æresfølelse eller omdømme. Det avgjørende er at ytringen *er egnet til* å få slike følger, jf. LB-2021-76153-2. Den påstått ærekrenkende ytringen skal tolkes etter hvordan en ordinær leser vil oppfatte utsagnene i den konteksten utsagnet er fremsatt, jf. Rt. 2015 s. 746, LA-2021-155004 og LB-2021-76153-2. Vi er ikke kjent med tilfeller der innhold skapt ved bruk av kunstig intelligens er funnet å utgjøre ytringer egnet til å krenke æresfølelse eller omdømme. Det kan imidlertid ikke utelukkes at dette kan være tilfelle, men dette vil altså avhenge av en konkret vurdering i den enkelte sak.

Ytringen må dessuten vurderes mot ytringsfriheten. I forarbeidene påpeker departementet at § 3-6 a må tolkes med ytringsfriheten i Grunnloven § 100 og EMK artikkel 10 som bakgrunn og rettesnor.[38] Dette har fått utslag i andre ledd, som fastsetter at en ellers ærekrenkende ytring ikke

[36] Se i denne retning Kierulf (2022) «Rettslig og folkelig personvern». Se også Stenvik (2003) s. 617 flg. om legalitetsprinsippets betydning for den ulovfestede læren om vern av personlighet.
[37] Se Hovlid (2018) «Erstatningsretten og krenkelser i sosiale medier».
[38] Ot.prp.nr. 22 (2008-2009) s. 488.

medfører erstatningsansvar om ytringsfriheten etter en avveining begrunner det. En rekke av sakene som domstolene har vurdert har vært søksmål mot nyhetsmedier, med påstand om at artikler inneholder ærekrenkende ytringer. Medias ytringer vil, avhengig av sakens innhold, ofte være i kjerneområdet for vernet av ytringsfriheten, jf. Rt. 2015 s. 746 avsnitt 65. Det kan ikke utelukkes at ytringer som inkluderer deepfake-materiale er i kjerneområdet av ytringsfriheten. Samtidig vil det klart ikke være tale om et sterkt ytringsfrihetsvern for en rekke av typetilfellene hvor deepfake-materiale benyttes, for eksempel svindel, pornografisk innhold o.l. Som tidligere nevnt, vil det i disse tilfellene ofte være utfordringer med å etterfølge et sivilprosessuelt søksmål.

2.4 Bestemmelser i straffeloven

2.4.1 Innledning

Forskjellige befatningsmåter med deepfake-materiale kan være straffbart etter ulike bestemmelser i straffeloven. I det følgende har vi begrenset oss til å gi en oversikt over bestemmelser vi antar er særlig relevante. Vi har skilt mellom straffebestemmelser om krenkende fremstillinger og lignende, og straffebestemmelser knyttet til desinformasjon, villedelse og bedrageri.

Vi bemerker at vi ikke er kjent med at det foreligger noen fellende straffedommer for deepfake-handlinger. Vårt søk er begrenset til publiserte avgjørelser på lovdata.no, og det kan ikke utelukkes at det foreligger påtalemessige saker av interesse (forelegg, bøter, tilståelsesdommer etc.) som ikke er offentliggjort.

2.4.2 Krenkende bilder mv.

Når det gjelder deepfake-materiale som viser personer i krenkende situasjoner, synes særlig straffeloven § 267 a å være relevant. Bestemmelsen har følgende ordlyd:

> *«Med bot eller fengsel inntil 1 år straffes den som uberettiget gjør tilgjengelig for en annen bilde, film eller lydopptak av krenkende eller åpenbart privat karakter, for eksempel av noens seksualliv eller intime kroppsdeler, noen som utsettes for vold eller andre ydmykelser, eller noen som befinner seg i en svært sårbar eller utsatt situasjon.*
>
> *Grovt uaktsom overtredelse straffes med bot eller fengsel inntil 6 måneder.»*

Etter ordlyden er det straffbart å *dele* bilder, filmer og lydopptak som er av krenkende eller åpenbart privat karakter. Bestemmelsen rammer imidlertid ikke fremstilling av slikt materiale.[39]

Det ikke krav om at bildene mv. må deles med allmenheten eller lignende for at handlingen skal være straffbar; det er tilstrekkelig at materialet er gjort tilgjengelig for én annen (forutsatt at det er «uberettiget»).

[39] I forarbeidene ble det vurdert om også fremstilling skulle rammes av straffebudet. I Prop. 159 L (2020–2021) punkt 8.6.1 er det redegjort for hvorfor denne befatningsmåten ble holdt utenfor bestemmelsen. Det gis flere begrunnelser. Blant annet vises det til at det er selve delingen som direkte fører til negative konsekvenser for den avbildede og at integritetskrenkende filming og fotografering allerede i dag kan rammes av straffeloven § 266, dersom den fornærmede oppfatter krenkelsen og dette dekkes av gjerningspersonens forsett. Etter departementets syn kunne det likevel tenkes at også voksne har behov for et strafferettslig vern mot fremstilling av seksualiserte bilder og opptak, og at det kunne være grunn til å se nærmere på behovet for et styrket vern som ledd i en helhetlig gjennomgang av straffelovens kapittel om seksuallovbrudd.

Videre stilles det – etter ordlyden – ikke krav om hvordan bildene mv. skal være produsert. I forarbeidene er det uttrykt at bildene mv. ikke behøver å være reelle. Samtidig er det imidlertid fremholdt at det må være tale om fotografisk materiale, se følgende formulering (vår uthevning):

> *«Straffebudet er **avgrenset** til å gjelde deling av «bilde, film eller lydopptak». Med «bilde» og «film» menes **fotografiske bilder og filmopptak**. Animasjoner, tegninger og lignende vil dermed falle utenfor. Det må imidlertid ikke være tale om reelle bilder eller opptak. Manipulasjoner kan derfor også omfattes.»*[40]

Med hensyn til video- og bildemateriale synes bestemmelsen dermed å være avgrenset på samme måte som åndsverkloven § 104 – dvs. for materiale som er produsert ved fotografisk eller lignende teknikk. Det kan det derfor reises spørsmål ved om straffeloven § 267 a får anvendelse på deepfakes som er produsert ved å konstruere nye bilder basert på informasjon avledet fra andre bilder. Spørsmålet er om ikke denne produksjonsmåten i stedet må anses for å være elektronisk/animasjon, med den følge at bildene ikke rammes av straffebudet.

Vi viser her til vår drøftelse av problemstillingen i tilknytning til åndsverkloven § 104 over. Her gjentar vi kun at problemstillingen synes å aktualisere vurderinger opp mot legalitetsprinsippet. For at deepfakes skal være omfattet av § 267 a, må det fremgå tilstrekkelig klart og presist av straffebudet. Klarhetskravet er behandlet av Høyesterett i flere saker. En avgjørelse som synes å ha en viss interesse i denne sammenheng, er Rt. 2012 s. 1211. Her fant Høyesterett at oppfordringer til straffbare handlinger ytret på en blogg ikke var dekket av den tidligere straffeloven § 140. Grunnen var at elektronisk formidling falt utenfor definisjonen på «trykt skrift». Til tross for at handlingen var klart straffverdig, og at lovgiver utvilsomt ønsket å ramme slike handlinger, var ikke klarhetskravet tilfredsstilt.

Når det gjelder spørsmålet om anvendelsen av straffeloven § 267 a på deepfakes, kan det for så vidt hevdes at deepfakes i utgangspunktet er omfattet av ordlyden ved «bilde» og «film». Det er først og fremst forarbeidene som synes å begrense anvendelsesområdet ved å angi at det med bilder og videoer menes *fotografisk materiale*.

Det fremgår av forarbeidene til straffeloven § 267 a at departementet har ønsket at bestemmelsen skal omfatte manipulasjoner. Her er også deepfakes nevnt; dvs. det er vist til *Medietilsynets høringsuttalelse* om behovet for en teknologinøytral bestemmelse, blant annet av hensyn til deepfake. Departementet sluttet seg til at det var behov for en teknologinøytral utforming med hensyn til *deling*.[41] Så langt vi kan se, er det imidlertid ikke gitt noen vurderinger av hvordan ulike deepfake-produksjoner skal bedømmes, gitt avgrensningen for *fotografiske bilder og filmer*.

Vi tar ikke endelig stilling til hvorvidt straffeloven § 267 a kommer til anvendelse på deepfakes. På grunn av forarbeidenes avgrensning for «fotografisk» materiale, anser vi vurderingen som usikker med hensyn til deepfakes som fullt ut er fremstilt ved digital konstruksjon. Formodentlig må spørsmålet avklares av domstolene og i siste instans Høyesterett.

---

[40] Se Prop. 159 L (2020–2021) punkt 12 Merknader til de enkelte bestemmelsene. Se også lignende formuleringer i punkt 8.5.2.
[41] Se Prop. 159 L (2020–2021) punkt 8.5.2.

Vi har imidlertid registrert at Politiet og Kripos, i medieutspill, har uttrykt at straffeloven § 267 a kan ramme deling av deepfakes.[42] Det ser også ut til at Justis- og beredskapsdepartementet har uttrykt at § 267 a er tilstrekkelig klar i denne sammenheng, se NRK-artikkelen «Mener det haster med ny lov om falske nakenbilder» (publisert 23. mars 2024). I samme artikkel uttrykker imidlertid professor Olav Torvund at det ikke er åpenbart at bestemmelsen omfatter fiktive bilder. Etter Torvunds oppfatning fanges ikke slike bilder opp av dagens bestemmelser.

Nylig besvarte justis- og beredskapsministeren et skriftlig spørsmål om behovet for lovregulering av KI-genererte bilder, se svar på skriftlig spørsmål 5. april 2024. I svaret viser ministeren blant annet til at straffeloven § 267 a – i henhold til forarbeidene – omfatter deling av bilder og videoer som ikke er reelle. Det ser ut til at ministeren legger til grunn at deepfakes er omfattet av bestemmelsen. Samtidig uttrykker ministeren avslutningsvis følgende:

> *«I lys av den teknologiske utviklingen må vi fortløpende vurdere om det er utfordringer med bruk og utvikling av kunstig intelligens som ikke fanges opp av dagens regler.»*

Etter vårt syn er det ikke åpenbart at deepfakes omfattes av bestemmelsen. Rettskildebildet er dog tynt. Vi tar derfor ikke endelig standpunkt til spørsmålet.

Dersom deling av krenkende materiale i ulike former for deepfakes ikke skulle dekkes av straffeloven § 267 a, vil det etter omstendighetene likevel kunne være straffbart etter straffeloven **straffeloven § 266** om hensynsløs adferd og **§ 267** om krenkelse av privatlivets fred.[43] Bestemmelsene har imidlertid visse begrensninger i anvendelsesområde, som vi vil fremheve i det følgende.

- **Straffeloven § 266** setter straff for den som «ved skremmende eller plagsom opptreden eller annen hensynsløs atferd forfølger en person eller på annen måte krenker en annens fred». Rettspraksis gir eksempler på at bestemmelsen har blitt anvendt på deling av krenkende bilder. I RG 2009 s. 1153 ble en person dømt etter den tilsvarende bestemmelsen i tidligere straffelov for å ha lagt ut to bilder på internett som viste fornærmede utføre seksuelle handlinger.

  Ettersom § 266 er generelt utformet, vil vi anta at det det kan ramme deepfakes uavhengig av hvilken teknikk som er brukt. En begrensning i straffebudets anvendelsesområde, ligger likevel i at gjerningspersonens forsett må omfatte at den den fornærmede oppfatter handlingen. I HR-2017-1245-A fant Høyesterett at en 16 år gammel gutt som hadde delt krenkende bilder med to kamerater, ikke kunne straffes etter bestemmelsen. Selv om det objektive gjerningsinnholdet var oppfylt, var det ikke bevist at gutten hadde forsett om at fornærmede skulle oppfatte handlingen.

- **Straffeloven § 267** setter straff for den som gjennom «offentlig meddelelse krenker privatlivets fred…» Begrepet «meddelelse» innebærer et krav om formidling av informasjon. Det er generelt utformet og setter ikke krav til måten informasjonen er formidlet. Vi antar derfor at deling av krenkende deepfakes vil kunne rammes av bestemmelsen.

[42] Se Politiets uttalelser i TV2-artikkelen "Flere tilfeller av KI-genererte nakenbilder: – Det er skremmende" (26.08.2023) og Kripos uttalelser i NRK-artikkelen «"Det skjulte nettverket for å kle av kvinner med KI" (19.03.2024).
[43] Se departementets høringsnotat i forbindelse med § 267 a. Fremstillingen i det følgende tar utgangspunkt i høringsnotatet.

> For at handlingen skal være offentlig, må den kunne nå et større antall personer, jf. straffeloven § 10. I forarbeidene som ligger til grunn for denne definisjonen er det uttalt at budskapet må nå mer enn 20-30 personer for at vilkåret skal være oppfylt.[44]

Som det fremgår har både straffeloven §§ 266 og 267 visse begrensninger i anvendelsesområdet som gjør at ulike befatningsmåter med deepfakes ikke vil rammes av disse bestemmelsene. Straffeloven § 266 vil ikke være anvendelig på deling av krenkende materiale der gjerningspersonen ikke hadde forsett om at den fornærmede skulle oppfatte det. Slikt forsett kan etter omstendighetene være vanskelig å bevise. Videre vil ikke § 267 være anvendelig der krenkende materiale deles med for få personer til at handlingen kan kalles «offentlig». Vi forstår det slik at **straffeloven § 267 a** ble innført blant annet på grunn av de nevnte begrensningene, se følgende uttalelse i forarbeidene:

> *«Formålet med forslaget er å sikre at all uberettiget deling av krenkende bilder er straffbar, og at dette kommer tydelig til uttrykk i loven.»*[45]

Spørsmålet er imidlertid om bestemmelsen er klart nok utformet til å fange opp deepfakes produsert ved digital konstruksjon basert på informasjon avledet av andre bilder, jf. over.

For en nærmere omtale av straffbarhetsvilkårene i straffeloven §§ 266, 267 og 267 a viser vi til forarbeidene, der forarbeidene til sistnevnte bestemmelse redegjør for samtlige bestemmelser: høringsnotat om endringer i straffeloven (bilder som er særlig egnet til å krenke privatlivets fred (Lovavdelingen, 2018) og Prop. 169 L (2020-2021) Endringer i straffeloven mv. (deling av krenkende bilder mv.).

Når det gjelder fremstilling og deling av krenkende materiale er det videre grunn til å fremheve **straffeloven § 311**. Denne bestemmelsen setter straff for ulike befatningsmåter med fremstillinger av seksuelle overgrep mot barn eller fremstillinger som seksualiserer barn. Bestemmelsen rammer blant annet den som produserer, anskaffer, besitter, skaffer seg tilgang til, overlater til en annen eller gjør tilgjengelig slike fremstillinger. [46] I forarbeidene er det uttrykt følgende om manipulasjoner/kunstige fremstillinger:

> *«Komiteen vil understreke at med bruken av ordet «fremstilling» mener man at dette rammer enhver fremstilling uansett medium, også tekst. Fremstillinger som er animert, manipulert eller på andre måter kunstig fremstilt, rammes også hvis de viser seksuelle overgrep mot barn eller seksualiserer barn.»*[47]

I henhold til dette vil bestemmelsen ramme den som ved hjelp av deepfake-teknologi lager fremstillinger av seksuelle overgrep mot barn eller som seksualiserer barn. Det er altså ikke en

---

[44] Se Prop. 53 L (20212-2013) punkt 3.6, med videre henvisninger.
[45] Se Prop. 159 L (2020–2021) punkt 1, samt nærmere redegjørelser i punkt 3.
[46] Se Prop. 159 L (2020–2021) punkt 3.6. Kapittelet inneholder en overordnet redegjørelse for bestemmelsens innhold.
[47] Se Innst. 0. nr. 66 (2004-2005) punkt 2. Strengt tatt er dette innstillingen til bestemmelsen som ble inntatt i den tidligere straffeloven av 1902. Ved utarbeidelse av ny straffelov ble bestemmelsen videreført med visse endringer, se 0t.prp. nr. 22 (2008-2009 punkt 7.20.4. Vi kan ikke se at det har vært meningen å endre innholdet av begrepet «fremstilling».

problemstilling om straffebudet er klart nok utformet i denne sammenheng, ettersom animasjon klart er omfattet. For øvrig er det grunn til å fremheve at der straffeloven § 267 a kun rammer deling, rammer straffeloven § 311 også f.eks. produksjon og besittelse.

2.4.3 Desinformasjon, villedelse og bedrageri

Deepfake-materiale kan videre benyttes i forbindelse med ulike former for spredning av falske nyheter og desinformasjon. I utgangspunktet synes det i seg selv ikke straffbart å spre falsk informasjon.[48] I visse tilfeller kan imidlertid visse villedelseshandlinger, samt handlinger med visse skadevirkninger, rammes av bestemmelser i straffeloven.[49] For eksempel setter **straffeloven § 164** straff for den som *«… uhjemlet utøver offentlig myndighet, eller som foretar handlinger som bare kan utøves av offentlige tjenestemenn …»* Uhjemlet myndighetsutøvelse kan tenkes muliggjort gjennom deepfake-teknologi. I så fall vil det kunne straffes etter bestemmelsen. Straffen er bot eller fengsel i inntil ett år. Videre setter **straffeloven § 165** straff for den som misbruker offentlig uniform, kjennetegn mv., og har følgende ordlyd:

> *«Med bot eller fengsel inntil 6 måneder straffes den som*
>
> *a. ved uhjemlet bruk av uniform eller på annen måte offentlig utgir seg for å ha offentlig myndighet på en slik måte at det er egnet til å skape uleilighet for noen eller svekke tilliten til den offentlige myndigheten,*
>
> *b. uhjemlet bruker et norsk eller utenlandsk offentlig våpen, merke eller segl eller noe som lett kan forveksles med slike, eller*
>
> *c. uhjemlet offentlig eller i rettsstridig øyemed bruker norsk eller utenlandsk offentlig tittel.»*

Det tenkelig at deepfakes kan benyttes i forbindelse med bevisfabrikasjon for å pådra noen siktelse eller domfellelse. Dette er straffbart etter **straffeloven § 222**.[50] Det er også tenkelig at deepfakes kan benyttes for å spre falsk informasjon med henblikk på å skape opptøyer. Dette antar vi at – etter omstendighetene – kan rammes av **straffeloven § 182 annet ledd**, som setter straff for den som har *«… fremkalt eller ledet omfattende ordensforstyrrelser med forsett om å øve eller true med vold på person eller skadeverk på eiendom».* Videre kan deepfakes benyttes til å fremkalle, styrke eller utnytte en villfarelse for å skaffe en uberettiget vinning. Dette vil være straffbart på de nærmere vilkårene som fremgår av **straffeloven § 371**.

2.4.4 Andre bestemmelser i straffeloven

Som angitt, er ikke vår redegjørelse for relevante straffebestemmelser nødvendigvis uttømmende. Det er imidlertid grunn til å fremheve at flere former for deepfakes-misbruk antakelig vil rammes av ulike bestemmelser i straffelovgivningen. I den forbindelse viser vi til følgende uttalelser i NOU

[48] Som påpekt av ytringsfrihetskommisjonen, er det betenkelig å straffe «fake news» og desinformasjon, se uttalelsene i NOU 2022: 9 punkt 9.9.2 (nærmere omtalt i denne utredningen punkt 3).
[49] Se Finansavisen-artikkelen Fake news kan straffes (2020), hvor advokat John Christian Elden redegjør for hvordan spredning av falske nyheter kan rammes av ulike straffebestemmelser. Vår fremstilling tar, for en del, utgangspunkt i denne redegjørelsen.
[50] Professor Erling Johannes Husabø har uttrykt at (blant annet) bildemanipulasjoner og deepfakes rammes av straffeloven § 182 (2), se Karnov-kommentar nr. 5 til bestemmelsen (krever innlogging). For øvrig er manipulering av rettsbevis fremhevet som en av truslene som deepfakes representerer i utredningen Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) punkt 7.3.

2022: 9 punkt 9.9.2, hvor det i tilknytning til *feilinformasjon* påpekes at det allerede eksisterer relevante bestemmelser:

> *«Det er de eventuelle skadene feilinformasjon kan bidra til som lovgivningen primært bør rettes mot. Her finnes det allerede lover. Grunnløse angrep på omdømmet til en person, kan forfølges sivilrettslig som ærekrenkelse. Uriktige anklager og oppdiktede straffbare handlinger som kan føre til ubotelig skade på enkeltmenneskers liv, kan rammes av bestemmelser i straffeloven. Feilinformasjon som skaper alvorlig frykt i befolkningen, kan rammes av straffebudet mot terrorhandlinger. Økonomisk skade som skyldes feilinformasjon, kan rammes av paragrafer om bedrageri. For demokratisk skade utløst av påvirkningsaksjoner, er det et mer åpent spørsmål om lovverket er godt nok. Det er også her den aktuelle diskusjonen om rettstilstanden har vært konsentrert (se neste punkt).»*

2.5 Særlig om AMT-direktivet

Vi vil kort nevne at det også finnes ytterligere EU/EØS-regulering som tilsynelatende har en side til deepfakes.

Endringsdirektivet om audiovisuelle medietjenester (AMT-direktivet) stiler krav til tilbydere av fjernsyn og audiovisuelle bestillingstjenester, og innholdet i deres tjenester.[51] Direktivet fastsetter minimumskrav og gir adgang til strengere eller mer detaljert regulering nasjonal.[52] AMT-direktivet har bestemmelser som pålegger statene å iverksette tiltak mot formidling av oppfordring om vold og hat.[53] Barn er særlig beskyttet i direktivet, ved at statene skal iverksette egnede tiltak rettet mot innhold som kan skade mindreåriges fysiske, psykiske eller moralske utvikling.[54]

Europaparlamentets utredningsenhet uttrykker i sin utredning om deepfakes at direktivet har bestemmelser som blant annet får betydning for spredning av for eksempel spredning av pornografisk innhold som er skapt ved deepfakes, uten at dette er samtykket til.[55] Flere av bestemmelsene, som statens plikt til å iverksette tiltak mot formidling av innhold som oppfordrer til hat eller vold, kan få betydning for innhold skapt ved bruk av kunstig intelligens.

Det reviderte AMT-direktivet ble vedtatt i EU 14. november 2018.[56] Det reviderte direktivet endrer direktivet om audiovisuelle medietjenester (2010/13/EU). Det nye AMT-direktivet ble innlemmet i EØS-avtalen i desember 2022 og er gjennomført i norsk rett.[57] Se regjeringen.no for en nærmere beskrivelse av innholdet i det reviderte AMT-direktivet.

---

[51] https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2016/juni/forslag-til-endringer-i-amt-direktivet-/id2503512/
[52] https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2016/juni/forslag-til-endringer-i-amt-direktivet-/id2503512/
[53] ATM-direktivet artikkel 6.
[54] ATM.direktivet artikkel 6a og Europaparlamentet s. 41.
[55] Europaparlamentet side 42.
[56] https://europalov.no/rettsakt/amt-direktivet-om-audiovisuelle-mediatjenester-endringsbestemmelser/id-9128
[57] https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2016/juni/forslag-til-endringer-i-amt-direktivet-/id2503512/ og https://europalov.no/rettsakt/amt-direktivet-om-audiovisuelle-mediatjenester-endringsbestemmelser/id-9128

2.6 Særlig merknad: Utfordringer knyttet til håndhevelse

Som gjennomgått over, kan flere befatningsmåter med deepfakes fanges opp av ulike bestemmelser i lovgivningen. Det faller utenfor vårt oppdrag å kartlegge «hull» i lovgivningen og mulige måter å tette disse på.[58] Det kan likevel være verdt å fremheve at mulighetene for å *håndhevelse* kan representere en vel så stor utfordring som ev. mangel på materielle regler. I den forbindelse viser vi til at det i Nederland er foretatt en kartlegging av hvordan deepfakes er regulert og hvordan det ev. bør reguleres, se følgende rapport i engelsk sammendrag:

- Deepfakes: The legal challenges of a synthetic society (Bert van der Sloot m.fl., 2021)

Rapporten ble utarbeidet på oppdrag fra det vitenskapelige forsknings- og dokumentasjonssenteret under det nederlandske justis- og sikkerhetsdepartementet. Rapporten fant at hovedutfordringen vedrørende deepfakes ikke var knyttet til manglende regulering, selv om ytterligere regulering kunne være ønskelig på visse områder. Hovedutfordringen var i stedet knyttet til mulighetene til å håndheve eksisterende (og ev. nye) reguleringer, se rapporten s. 6-7 (våre uthevninger):

> *"Perhaps the most important insight regarding the current legal framework is that although amendments are possible and perhaps desirable on specific points, such would not tackle the main problem with regard to deepfakes in horizontal relationships and, more generally, to breaches of privacy in horizontal relationships.* ***In the first, second and third place, the problem is one of enforceability****. Producing pornographic material of another person without her consent is already prohibited; generating child pornography of a fictitious child is already prohibited; committing fraud and deception by means of a deepfake is already prohibited; introducing false evidence in a court case is already prohibited; inciting hatred or violence between groups is already prohibited; exploiting someone's image or likeness or creative works without permission is already prohibited; causing (economic) harm by means of identity theft or reputational harm by fake messages can already be dealt with under tort law; etc.*
>
> ***The legal framework applicable to deepfakes is not the primary problem; the problem is the enforcement of the existing and any additional legal rules****. (…)"*

Det kan, som nevnt, se ut til at flere former for befatning med deepfakes helt eller delvis er omfattet av ulike lovbestemmelser i Norge. Det kan derfor spørres om ikke mulighetene for å *håndheve* disse bestemmelsene representerer en vel så stor utfordring som ev. mangel på regulering. Den nederlandske rapporten peker blant annet på følgende hindringer knyttet til regulering og håndhevelse:

- Teknologien utvikler seg raskt. Teknologi-spesifikke regler blir derfor fort blir utdatert.

- Det er vanskelig å definere teknologi for lovgivningsformål. En for smal definisjon av teknologien man ønsker å regulere, vil gi rom for omgåelse. På den annen side vil en for bred definisjon kunne hindre også positiv anvendelse av teknologien.

[58] Vi gjør imidlertid oppmerksom på at en slik vurdering er gjort på EU-nivå i Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021), se punkt 7 («regulatory gaps») og punkt 8 («regulatory options»).

- Datadrevne teknologier er grensekryssende. Involverte parter er ofte underlagt ulike rettslige regimer. Ofte er partene lokalisert i jurisdiksjonen med den laveste regulatoriske byrden.
- Det er vanskelig å håndheve reglene i en jurisdiksjon overfor parter lokalisert i andre land.
- Det er ofte flere parter involvert i produksjon og distribusjon av deepfakes, som alle har et visst ansvar.
- Det er ofte lett å omgå reglene i en spesifikk jurisdiksjon, for eksempel ved å benytte VPN-forbindelse.
- Identiteten til den som utfører deepfake-misbruk er ofte skjult og vanskelig å finne ut av. Dette fordrer samarbeid mellom, noe som kan være vanskelig å få etablert.

3 Politiske og lovgivningsmessig initiativ

Vi har ikke sett offentlige utredninger eller politiske forslag som *spesifikt og utfyllende* behandler lovregulering av deepfake. I den nasjonale strategien for kunstig intelligens drøftes imidlertid regulering av KI (herunder bl.a. deepfake). I punkt 2.3 «regelverk» er det gjort følgende overordnede betraktninger:

> *«Samtidig ser vi ofte at det blir fremsatt ønsker om regulering når nye teknologier fører til problematiske anvendelser. Med kunstig intelligens har vi sett eksempler på dette, blant annet knyttet til valgmanipulasjon i sosiale medier og såkalte «deep fakes». Samtidig er det krevende – og ofte uhensiktsmessig – å regulere en teknologi som fortsatt er i en tidlig fase. For tidlig regulering kan forme utviklingen på en utilsiktet måte, skape skjevheter i markedet og begrense potensialet for innovasjon. I tillegg vil en gitt teknologi oftest både ha positive og negative anvendelser. Den samme underliggende teknologien som gjør det mulig å lage «deep fakes» gjør det for eksempel også mulig å lage syntetiske testdata – en teknologi som bidrar til å beskytte personopplysninger.»*

En ny digitaliseringsstrategi skal, ifølge opplysninger på regjeringens nettsted, foreligge i 2024. Strategien skal blant annet adressere kunstig intelligens. Vi har også registrert at det i fjor ble nedsatt et arbeidsutvalg med medlemmer fra ulike departementer med oppdrag å vurdere behovet for nasjonal regulering av kunstig intelligens. Ifølge uttalelser i Innst. 151 S (2023-2024) avga utvalget rapport i november 2023. Vi kan ikke se at den er offentlig tilgjengelig. Utvalget skulle for øvrig også lage en plan for rettidig og god gjennomføring av KI-forordningen, se nærmere omtale i punkt 5.7.[59]

Vi nevner videre at regjeringen nylig har foreslått endringer i straffeloven om ulovlig påvirkningsvirksomhet, se Prop. 42 L (2023–2024). Konkret er det foreslått å gjøre det straffbart å bidra på vegne av eller etter avtale med en fremmed etterretningsaktør i virksomhet som har til

[59] Se Altinget-artikkelen «Norges svar på EUs lovarbeid om Kunstig Intelligens: Hurtigarbeidende arbeidsgruppe» (30. juni 2023).

formål å påvirke beslutninger eller den allmenne meningsdannelsen, når virksomheten kan skade betydelige samfunnsinteresser. De foreslåtte bestemmelsene har følgende ordlyd:

Ny § 130:

> *«§ 130 Påvirkning fra fremmed etterretning*
>
> *Med bot eller fengsel inntil 3 år straffes den som på vegne av eller etter avtale med en fremmed etterretningsaktør bidrar i virksomhet som har til formål å påvirke beslutninger eller den allmenne meningsdannelsen, når virksomheten kan skade betydelige samfunnsinteresser.»*

Ny § 130 a:

> *«§ 130 a Grov påvirkning fra fremmed etterretning Grov overtredelse av § 130 straffes med fengsel inntil 10 år. Ved avgjørelsen av om overtredelsen er grov, skal det særlig legges vekt på a. overtredelsens karakter og omfang, b. om gjerningspersonen i kraft av sin stilling nyter en særlig tillit, c. om overtredelsen av andre grunner er særlig samfunnsskadelig, og d. om gjerningspersonen har skaffet seg selv eller andre en betydelig vinning.»*

I proposisjonen har departementet fremhevet at deepfake er blant teknologiene som kan benyttes i påvirkningsvirksomhet som nevnt, se s. 39. Proposisjonen behandles i skrivende stund av Justiskomiteen, som skal avgi sin innstilling 23. april 2024.

Deepfakes har videre blitt adressert i ulike NOUer. Spørsmål om regulering av desinformasjon og feilinformasjon ble behandlet i NOU 2022: 9 En åpen og opplyst offentlig samtale – Ytringsfrihetskommisjonens utredning, se kapittel 9 Desinformasjon og feilinformasjon. Deepfakes er særlig fremhevet under punktet knyttet til trender og framtidsutsikter, se punkt 9.6. Det uttrykkes blant annet til at utbredelse av deepfake-teknologi kan medføre en oppfatning om at det er *«umulig for en vanlig bruker å vite om man kan stole på det man ser og hører. Dette kan føre til at folk lettere blir lurt, men også til at usikkerheten øker og tilliten til informasjon synter».* Utvalget fremhever at utviklingen av bildemanipulasjon m.m. reiser etiske og regulatoriske problemstillinger som går utover spørsmålet om desinformasjon og Ytringsfrihetskommisjonens mandat.

Selv om de særlige regulatoriske utfordringene knyttet til deepfake ikke ble behandlet, ga utvalget vurderinger og anbefalinger vedrørende lovregulering av desinformasjon generelt, se punkt 9.9 og punkt 9.10. Til en viss grad antar vi at disse punktene har interesse også med hensyn til deepfakes som benyttes i forbindelse med slik aktivitet. Helt overordnet advarte utvalget mot å innføre straffebud som forbyr spredning av usann informasjon, se punkt 9.9.2. Det ble blant annet vist til at det kan svekke ytringsfriheten, at det i mange tilfeller ikke er noen fasit på hva som er sant eller usant, at det er fare for begrepsutflytning (f.eks. at påstander som beskyldes for å være usanne, ofte er politiske vurderinger eller meninger), samt at slike lover lett kan misbrukes og ha en nedkjølende effekt. I tilknytning til det siste ble det blant annet uttrykt følgende:

> *«Å straffeforfølge offentlige ytringer alene fordi de er usanne, kan stenge eller redusere det offentlige rommet for kritikk og motstemmer. Dette er en reell problemstilling i land med autoritære styresett som gjerne griper til slike harde tiltak.»*

Utvalget påpekte i det videre at det er de eventuelle skadene feilinformasjon kan bidra til som lovgivningen primært bør rettes mot. Til dette bemerket utvalget at det allerede finnes lover:

> *«Grunnløse angrep på omdømmet til en person, kan forfølges sivilrettslig som ærekrenkelse. Uriktige anklager og oppdiktede straffbare handlinger som kan føre til ubotelig skade på enkeltmenneskers liv, kan rammes av bestemmelser i straffeloven. Feilinformasjon som skaper alvorlig frykt i befolkningen, kan rammes av straffebudet mot terrorhandlinger. Økonomisk skade som skyldes feilinformasjon, kan rammes av paragrafer om bedrageri. For demokratisk skade utløst av påvirkningsaksjoner, er det et mer åpent spørsmål om lovverket er godt nok. Det er også her den aktuelle diskusjonen om rettstilstanden har vært konsentrert (se neste punkt).»*[60]

Til sist kan nevnes at deepfakes har vært omtalt i ulike representantforslag om regulering av KI, se f.eks. følgende:

- Representantforslag 46 S (2023–2024), som inneholder flere forslag knyttet til regulering av KI og bruk av data. Blant annet inneholdt dokumentet forslag om å anmode regjerningen om å kartlegge hvilke deler av lovverket som er utydelige i møte med ny teknologi som kunstig intelligens.

- Representantforslag 232 S (2022–2023), som blant annet inneholdt forslag om å anmode regjeringen om å utvikle en konsesjonsordning for innsamling, bruk og lagring av både offentlige data og persondata for kommersielle aktører.

- Representantforslag 146 S (2021–2022), som inneholdt forslag knyttet til bekjempelse av falske nyheter og desinformasjon.

4 Rettslige dilemmaer ved regulering av deepfake

Spørsmålet om hvilke rettslige dilemmaer som kan oppstå i forbindelse med regulering av deepfake er vanskelig å besvare på en uttømmende måte. En eventuell regulering av deepfake kan gripe inn på mange ulike livsområder og ha en side til flere lovverk. Å identifisere rettslige dilemmaer vil nødvendiggjøre et nærmere kartleggings- og analysearbeid som er krevende å utføre innenfor denne utredningens rammer. Et stykke på vei er det også en politisk avveining – ev. en verdivurdering – hvorvidt man oppfatter noe som et dilemma. Mot denne bakgrunnen har vi begrenset oss til å gjengi overordnet enkelte mulige utfordringer.

For det første kan regulering av deepfakes ha en side til ytringsfriheten. Deepfakes vil etter vår oppfatning være ytringer som i utgangspunktet er vernet av ytringsfriheten etter Grunnloven og EMK. Det kan være betenkelig med vidtfavnende reguleringer som f.eks. rammer kunstneriske fremstillinger, politisk satire og lignende. Samtidig kan deepfakes ha skadevirkninger man kan mene det er nødvendig å gripe inn mot. Som et eksempel på en vurdering av utfordringene ved å regulere deepfakes, herunder forholdet til ytringsfriheten, kan det vises til utredningen Regulating

[60] Punkt 9.9.2. Teksten inneholder fotnotehenvisninger som er utelatt fra sitatet.

Deep Fakes in the Artificial Intelligence Act av Mateusz Łabuz (doktorgradsstipendiat v/ Universitetet for teknologi i Chemnitz). Her uttrykkes det blant annet følgende:

> *"In an increasing number of cases, legislators have prohibited or limited the use of deep fakes, but they have also allowed significant exceptions in the form of obvious or evident satire or parody (of a 'demonstrably' fake nature). This 'obvious' or 'evident' nature may be debatable and would have to be assessed on a caseby- case basis because it might depend on contextualisation as well as the cognitive abilities, media knowledge, or social and political awareness of recipients.*
>
> *Unfortunately, overusing the legal exemptions could be seen as a useful tool to circumvent the restrictions. Deep fakes are described as a phenomenon that might benefit from the 'just joking' excuse, making it possible to smuggle illegal content or manipulate the audience 'under the guise of humour', which might even lead to the 'weaponisation of humour'. Satirical context has already been shown to function as 'a cover for spreading' extremist ideologies with respect to fake news. At the same time, the fight against deep fakes might also be used to justify suppressing freedom of speech. This is especially important in the case of non-democratic countries that hide their censorship tendencies under the guise of protecting social stability."*[61]

Se hertil bemerkningene knyttet til lovregulering av desinformasjon i NOU 2022: 9 - kapittel 9 Desinformasjon og feilinformasjon, jf. omtale over. Se også punkt 6.9.1. i Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021), hvor det i tilknytning til lovgivning i California og Texas om manipulerte bilder av politiske kandidater ble fremhevet følgende (vår uthevning):

> *"Under the new California law (AB 730), it is illegal to distribute manipulated content featuring political candidates within a 60-day period before an election that is intended to injure the candidate's reputation or to deceive a voter into voting for or against the candidate. The Texas law is very similar to California's, but only prohibits distribution within a 30-day period.* ***Both laws drew considerable criticism, particularly questioning their compatibility with the right to free speech.****"*

For det annet kan en streng regulering av deepfakes stå i veien for innovasjon og bruk som man ellers måtte anse ønskelig, sml. uttalelsene i den nasjonale strategien for kunstig intelligens punkt 2.3. En redegjørelse for fordeler og risikoer ved deepfakes er for øvrig gitt i Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) punkt 5. Se også høringsinnspillet fra Fakultetet for informasjonsteknologi og elektronikk ved NTNU i forbindelse med ny nasjonal digitaliseringsstrategi, hvor det motsatt betones at regulering kan være nødvendig for å ta i bruk kunstig intelligens:

> *«Den raske utviklingen av digitale teknologier, og spesielt KI, kan føre til store konsekvenser for samfunnet. Her er det essensielt å ha fokus på lover og reguleringer, samt etikk. Mangel på lover og reguleringer kan ikke bare være et problem for individet, men også være et hinder for at norsk arbeidsliv kan ta i bruk kunstig intelligens.*

[61] Teksten inneholder fotnotehenvisninger som er fjernet fra sitatet.

Digitaliseringsstrategien må derfor sørge for at vi ikke kommer på etterskudd med å få på plass et lovverk som kan beskytte brukere og legge til rette for forskning og utvikling.»

For det tredje er det et spørsmål hvilke deler av deep fake-prosessen som skal reguleres: Skal visse teknologier være forbudt? Skal visse former for produksjon være forbudt eller underlagt særlige krav? Er det eventuelt bare visse former for deling av deepfake-materiale som skal være underlagt begrensninger? Hvilket ansvar skal plattformer hvor deling skjer ha? I denne forbindelse viser vi til at en redegjørelse for mulige reguleringsalternativer på ulike stadier i «deepfake-prosessen» er gitt i Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) punkt 8.

For det fjerde kan vanskelige spørsmål oppstå der det er tale om å regulere ansvaret for autonome KI-programvarer – dvs. KI-programvare som produserer og deler innhold uten menneskelig inngripen.[62] I hvilken utstrekning kan og bør de som har laget en KI-programvare være ansvarlig for innhold KI-programvaren produserer «på egenhånd»?

For øvrig viser vi til det som tidligere er fremholdt om utfordringer knyttet til håndhevelse og hvordan krenkelser skal stanses. Å forfølge krenkelser sivilrettslig vil ofte være kostbart og tidkrevende for den enkelte. Datatilsynet har i dag begrensende ressurser og redskaper for å kunne stanse spredning av en deepfake. Straffesporet vil kunne virke allmennpreventivt, men forutsetter klare og forutsigbare straffebestemmelser. I tillegg er bevisbyrden adskillig høyere enn i sivile saker, og både teknologi og grenseoverskridende aktivitet kan gjøre det krevende å oppnå domfellelser.

5 Regulering i andre land og EU

5.1 Innledning

For å besvare spørsmålet om hvordan deepfake er regulert i andre land, innhentet vi informasjon ved å henvende oss til utredningstjenestene i andre lands parlamenter gjennom ECPRD-nettverket.[63] Konkret sendte vi henvendelser til utvalgte europeiske land, USA og Canada, samt Europaparlamentets utredningstjeneste.

Vi mottok svar fra følgende: Danmark, Sverige, Finland, Tyskland, Estland, Canada og Europaparlamentets utredningstjeneste. I **punktene 5.2 – 5.8** redegjør vi enkeltvis for svarene vi mottok. I enkelte sammenhenger har vi supplert svarene med informasjon vi har funnet ved egne søk. I tillegg redegjør vi overordnet i **punkt 5.9** for informasjon vi ved egne søk har funnet om USA, før vi i **punkt 5.10** kort omtaler reguleringer i Kina.

5.2 Danmark

I ECPRD-svar fra Danmark er det opplyst at fremstilling og distribuering av deepfakes er straffbart etter den danske straffeloven § 264 e. Bestemmelsen har følgende ordlyd:

[62] Se hertil drøftelsene knyttet til ytringsfrihet og -ansvar i tilknytning til autonome AI-brukere i Carl Otto Bjerkelund Arnesen, En analyse av lovregulering for å sikre frie valg mot påvirkningsaksjoner gjennom sosiale medier (masteroppgave, 2021) punkt 3.2.4.7.
[63] ECPRD står for European Center for Parliamentary Research and Documentation. Det er et samarbeid for utveksling av informasjon mellom europeiske parlamenter som Stortinget deltar i.

«For identitetsmisbrug straffes med bøde eller fængsel indtil 6 måneder den, der uberettiget

1) anvender oplysninger om en anden person, herunder cpr-nummer, navn og billede, til på utilbørlig vis at udgive sig for at være denne person eller

2) videregiver materiale, hvor der er gjort brug af oplysninger om en anden person, herunder cpr-nummer, navn og billede, til på utilbørlig vis at manipulere denne persons fremtræden.» [64]

Ifølge forarbeidene omfatter bestemmelsen etter omstendighetene tilfeller hvor det gjøres bruk av deepfake. For eksempel kan dette være tilfellet hvor ansiktet til fornærmede er redigert inn i forskjellige former for pornografisk materiale:

«Den foreslåede bestemmelse vil således efter omstændighederne kunne omfatte tilfælde, hvor der gøres brug af såkaldte deepfakes, dvs. manipulation af lyd og video ved hjælp af kunstig intelligens, så en person figurerer i videoer mv., hvor vedkommende gør eller siger noget, som vedkommende ikke har gjort eller sagt i virkeligheden. Det kan f.eks. endvidere være tilfælde, hvor forurettedes ansigt er klippet ind i forskellige former for pornografisk materiale, så det fremstår, som om forurettede har deltaget i de pågældende seksuelle aktiviteter. Det kan desuden f.eks. være tilfældet, hvor forurettedes ansigt er klippet ind i en video med lyd, hvor det fremstår, som om forurettede siger noget, som vedkommende ikke har sagt.»[65]

Det ser for øvrig ut til å eksistere dansk rettspraksis knyttet til spredning av falske pornografiske bilder. Ved egne søk funnet en nyhetsartikkel hvor det redegjøres for straffedommer knyttet til denne typen spredning, se DR-artikkelen «'Kraftige straffe' har ikke stoppet delingen: Falske nøgenbilleder af danske kvinder bliver stadig delt og byttet», publisert 24. juni 2023. Det er imidlertid ikke presist angitt hvilke straffebestemmelser det er dømt etter.

I ECPRD-svaret fra Danmark er det videre vist til at lovgivning vedrørende personvern må overholdes i forbindelse med deepfakes. Det er lagt til grunn at hvis opplysninger i deepfakes kan spores tilbake til den person, må personvernreglene overholdes.

For øvrig opplyst om at det ikke er noen pågående initiativ som adresserer deepfakes.

5.3 Sverige

I ECPRD-svaret fra Riksdagens utredningstjeneste er det opplyst at deepfakes ikke er spesifikt regulert i svensk rett. Det er likevel uttrykt at flere svenske og internasjonale lover kan aktualiseres ved vurderingen av deepfakes. Ikrafttredelsen av EUs nye KI-forordning er imidlertid antatt å kunne medføre at eksisterende nasjonal lovgivning kan bli tilsidesatt/erstattet, siden KI-forordningen vil ha forrang i tilfelle konflikt med nasjonal lovgivning.[66]

[64] Den danske straffeloven er, så vidt vi kan se, tilgjengelig herfra: https://www.retsinformation.dk/eli/lta/2022/1360.
[65] Se L 103 Forslag til lov om ændring af straffeloven s. 11.
[66] I svaret fra Riksdagens utredningstjeneste er det redegjort for relevante bestemmelser i KI-forordningen. Vi gjengir ikke dette her, men viser til omtalen av KI-forordningen i ECPRD-svaret fra Europaparlamentets utredningstjeneste.

Svaret fra Sverige er for øvrig relativt omfattende, og vi har vært nødt til å forkorte noe.

Når det gjelder nasjonal regulering, viser redegjørelsen fra Riksdagens utredningstjeneste til utgangspunktene i de svenske grunnlovene om regjeringsformen («Regeringsformen»), ytringsfrihet («Yttrandefrihetsgrundlagen») og pressefrihet («Tryckfrihetsförordningen»). En oversikt over hva disse grunnlovene inneholder, er gitt på Riksdagens nettsted.

Ytringsfriheten kan, iht. til Regeringsformen, bare begrenses der det gjøres av grunner som er akseptable i et demokratisk samfunn. Når det gjelder grunnlovene om ytringsfrihet og pressefrihet, fremholdes det i svaret fra Riksdagens utredningstjeneste at disse lovene får anvendelse pressefriheten og den korresponderende retten til å ytre seg på radio, tv eller lignende transmisjoner, i film, video- og lydopptak og andre tekniske opptak. Lovene inneholder et detaljert beskyttelsessystem som hviler på en rekke fundamentale prinsipper. Disse prinsippene har til hensikt å tilveiebringe en særlig sterk beskyttelse for trykt materiale og visse andre medieformater. Prinsippene omfatter forbud mot sensur, etableringsfrihet, individuelt ansvar, beskyttelse for personer som kommuniserer informasjon, en spesifisert lite med forbrytelser og særlige straffeprosessuelle bestemmelser. En viktig komponent i denne beskyttelsen, er at ingen restriksjoner på pressefriheten og ytringsfriheten kan gjøres annet enn på bakgrunn av disse lovene. Andre begrensninger vil kreve konstitusjonelle endringer.

Den svenske utredningstjenesten viser deretter til kapittel 4 og 5 i den svenske straffeloven (Brottsbalk (1962:700)). Det uttrykkes at iht. bestemmelser i disse kapitlene kan personer straffes for ulovlige trusler, ærekrenkelser og krenkende språk og oppførsel. Vi oppfatter det slik at Riksdagens utredningstjeneste anser at dette er bestemmelser som kan komme til anvendelse på deepfakes, men det redegjøres ikke nærmere for de enkelte bestemmelsenes innhold. Imidlertid fremheves én bestemmelse som synes å være av særlig relevans i denne sammenheng, nemlig kapittel 4 § 6c. Ifølge Riksdagens utredningstjeneste trådte denne bestemmelsen i kraft 1. januar 2018. Ifølge bestemmelsen kan en person som gjør inngrep i en annens privatliv ved å spre visse typer av blider eller annen informasjon, straffes for urettmessig inngrep i privatlivet. Dette gjelder så fremt det er gjort med intensjon om å krenke den avbildede personen. Ordlyden i bestemmelsen er som følger:

> *«Den som gör intrång i någon annans privatliv genom att sprida*
>
> *1. bild på eller annan uppgift om någons sexualliv,*
>
> *2. bild på eller annan uppgift om någons hälsotillstånd,*
>
> *3. bild på eller annan uppgift om att någon utsatts för ett brott som innefattar ett angrepp mot person, frihet eller frid,*
>
> *4. bild på någon som befinner sig i en mycket utsatt situation, eller*
>
> *5. bild på någons helt eller delvis nakna kropp*
>
> *döms, om spridningen är ägnad att medföra allvarlig skada för den som bilden eller uppgiften rör, för olaga integritetsintrång till böter eller fängelse i högst två år.*

Det ska inte dömas till ansvar om gärningen med hänsyn till syftet och övriga omständigheter var försvarlig.»

Vi forstår svaret fra den svenske utredningstjenesten slik at denne bestemmelsen også omfatter deepfakes.

Riksdagens utredningstjeneste viser også til at det etter kapittel 18 § 5 i den svenske straffeloven er straffbart å utøve ulovlig tvang eller fremsette ulovlige trusler med hensikt å påvirke den offentlige meningsdannelse eller gjøre inngrep i handlingsfriheten innenfor en politisk organisasjon eller yrkes- eller næringssammenslutning, og derigjennom setter ytrings-, forsamlings- eller foreningsfriheten i fare. Vi oppfatter det slik at Riksdagens utredningstjeneste anser at denne bestemmelsen kan ramme deepfakes som benyttes i slike sammenhenger.

Videre vises det til *Lag (1998: 112) om ansvar för elektroniska anslagstavlor*, som gjelder for tjenester for elektronisk formidling av meldinger. Etter denne loven er tilbydere offentlige digitale tjenester hvor personer kan publisere innhold ansvarlige for å overvåke tjenesten, samt for å fjerne og hindre spredningen av innhold som ikke er i samsvar med særlige lover og reguleringer. Dette omfatter blant annet ulovlige trusler, ulovlig inngrep i privatlivet i henhold til den svenske straffeloven kapittel 4 § 6 – 17, barnepornografiforbrytelser etter kapitel 16 § 10 a i samme lov m.m..

For øvrig vises det til at deepfakes vil være omfattet av GDPR og EMK artikkel 8. Det er også vist til at det i den svenske opphavsrettsloven gjelder en rett til eget bilde. Retten innebærer, med noen unntak, at fotografier som avbilder en person ikke kan tilgjengeliggjøres for allmenheten uten samtykke fra den avbildede personen. Det uttrykkes at denne retten får anvendelse på deepfake-materiale basert på fotografier.

Til sist er det opplyst at det, så langt Riksdagens utredningstjeneste kjenner til, ingen pågående arbeid eller initiativ knyttet til deepfakes.

5.4 Finland

I ECPRD-svaret fra Finland er det kort uttrykt at det ikke eksisterer reguleringer som spesifikt gjelder deepfakes. Det er imidlertid fremholdt at eksisterende lovgivning kan komme til anvendelse, for eksempel GDPR, EMK artikkel 8, den finske opphavsrettsloven (Copyright Act (404/1961)) og straffeloven (Criminal Code (37/1889)).

Når det gjelder opphavsrettsloven, er det vist til at loven beskytter kunstneriske verk, deriblant fotografiske verk. Det er imidlertid også fremholdt at publiserte verk kan benyttes i parodier og karikaturer. Når det gjelder straffeloven, er det vist til at forskjellige bestemmelser kan få anvendelse, særlig straffebud relatert til distribusjon av avbildninger av vold eller krenkende bilder. I tillegg er det vist til straffebestemmelser vedørende bedrageri.

For øvrig er det bemerket at det ikke er noe pågående lovgivningsarbeid vedrørende deepfakes, men at man forbereder seg på KI-forordningen.

5.5 Estland

I ECPRD-svaret fra Estland er det fremholdt at det ikke eksisterer lovgivning som gjelder produksjon og bruk av deepfakes. I visse henseender kan imidlertid dette være dekket av § 157 i den estiske straffeloven. Vi fikk oversendt følgende oversettelse av bestemmelsen

> *"§ 157². Illegal use of another person's identity*
>
> *(1) Transmission of personal data that establish or may enable to establish the identity of another person, grant of access to the data or use thereof, without the consent of that person, with the aim to knowingly cause a misconception of that person by means of assuming that person's identity, if damage is caused thereby to the rights or interests of another person that are protected by law, or to conceal a criminal offence, is punishable by a pecuniary punishment or up to three years' imprisonment.*
>
> *(2) The same act, if committed by a legal person, is punishable by a pecuniary punishment.*
>
> *[RT I, 19.03.2019, 3 – entry into force 01.07.2019]"*

For øvrig er det opplyst at det ikke foreligger informasjon om planer om å lovregulere deepfakes.

5.6 Tyskland

I ECPRD-svaret fra Tyskland opplyses det om at det ikke eksister spesifikk lovregulering av deepfakes. Deepfakes kan imidlertid være omfattet av rettsregler på ulike rettsområder. I den grad det benyttes personopplysninger, gjelder kravene» begrensningene i EMK artikkel 8 og GDPR. Det er også vist til at DSA artikkel 35 (1) bokstav k stiller krav om merking av deepfakes (se omtale i denne utredningen **punkt 5.7**). Vi nevner ellers at det tyske svaret har vært noe utfordrende å fremstille, gitt at svaret ble gitt på tysk og at vi har hatt for kort tid til å oversette dette.

Hva som er tillatt i forbindelse med deepfakes, og hvilke grenser som gjelder, er videre forklart å være et resultat av en avveining av grunnleggende rettigheter og interesser, der individers rettsposisjoner står mot hverandre. For de som rammes av deepfakes, dreier det seg først og fremst om den generelle personlighetsretten som er nedfelt i artikkel 12 (1) i den tyske grunnloven (GG), som kan kollidere med ulike grunnleggende rettigheter for teknologileverandører og -brukere. For sistnevnte kan spesielt rettighetene som informasjonsfriheten (GG artikkel 5, avsnitt 1), frihetene forbundet med kunst og vitenskap (GG artikkel 5, avsnitt 1), yrkesfriheten (GG artikkel 12) eller eiendomsrettigheter (GG artikkel 14) berøres.

I ECPRD-svaret vises det videre til at nevnte personlighetsrett ikke direkte gir grunnlag for erstatningskrav eller for å begjære midlertidig forføyning mot personer som sprer deepfakes. Vi forstår det imidlertid slik at dette er noe som kan vurderes i forbindelse med sivilrettslige bestemmelser i BGB (konkret vises det til artikkel 1004 og 823, uten at det forklares hva disse handler om). Så vidt vi forstår, skal den føderale forfatningsdomstolen – i tilknytning til personlighetsretten – ha foretatt en avveining av de motstridende rettighetene som gjør seg gjeldende i forbindelse med spredning av et manipulert bilde. Det vises til følgende dom, som vi har funnet omtalt på engelsk på nettsidene til domstolen: 1 BvR 240/04.

Det er forklart at personlighetsreten gir sivilrettslige beskyttelseskrav, spesielt i henhold til §§ 1004 og 823 i BGB i forbindeelse med GDPR artikkel 6 (1), forutsatt at det foreligger uautorisert bruk av personopplysninger.

Det er videre vist til at den tyske åndsverksloven (KUG) kan gi grunnlag for krav basert på krenkelse av retten til eget bilde. Ifølge §§22 i nevnte lov kan bilder kun spres eller vises offentlig med samtykke av den avbildede. Vi forstår det slik at denne bestemmelsen, etter den tyske utredningstjenestens oppfatning, omfatter deepfakes.

I tillegg til sivilsrettslige krav, viser den tyske utredningstjenesten til at ulike straffebestemmelser kan ramme deepfakes. Det vises til at deepfakes kan straffeforfølges i henhold til § 201 a (2) i den tyske straffeloven (StGB) for krenkelse av den «mest personlige livssfæren» og personlige rettigheter gjennom bildeopptak, § 33 i den tyske åndsverkloven (KUG) for spredning av et bilde uten samtykke.

5.7 EU

5.7.1 Eksisterende lovgivning i EU

I ECPRD-svaret fra Europaparlamentets utredningstjeneste er det fremholdt at deepfakes berøres av både GDPR og forordningen om digitale tjenester («DSA» – Digital Services Act).

Når det gjelder GDPR, er det uttrykt at ettersom personopplysninger benyttes når deepfakes lages, vil slik aktivitet falle innunder det materielle virkefeltet til GDPR. Det er vist til at slik behandling vil kreve rettsgrunnlag, som f.eks. det informerte samtykke fra den registrerte. Det er videre vist til at deepfakes kan fremstille «hvem som helst» som «gjør eller sier hva som helst», og at dette kan innebære at bruk av sensitive personopplysninger. Dette er i prinsippet forbudt under GDPR, men det er unntak i artikkel 9 (2). Som sensitive personopplysninger regnes opplysninger om rasemessig eller etnisk opprinnelse, politisk oppfatning, religion, filosofisk overbevisning eller fagforeningsmedlemskap, samt behandling av genetiske opplysninger og biometriske opplysninger med det formål å entydig identifisere en fysisk person, helseopplysninger eller opplysninger om en fysisk persons seksuelle forhold eller seksuelle orientering.

I ECPRD-svaret vises det videre til enkelte rettigheter en registrert kan ha mot den behandlingsansvarlige i forbindelse med deepfakes. Blant annet nevnes at den registrerte har

1) rett til sletting etter artikkel 17 (1) bokstav d der personopplysninger har blitt behandlet ulovlig,
2) rett til å bli informert om behandlingen, jf. artikkel 12 (1),
3) rett til å få tilgang til informasjonen som behandles, jf. artikkel 15 (1)
4) rett til å få uriktige personopplysninger rettet, jf. artikkel 16

For en nærmere analyse av GDPR og deepfakes viser Europaparlamentets utredningstjeneste til følgende kilder.

- https://www.sciencedirect.com/science/article/pii/S0267364922000632
- https://library.oapen.org/handle/20.500.12657/88178 (pdf)
- https://bartvandersloot.com/onewebmedia/edpl_2020_04-004.pdf

- https://www.mondaq.com/turkey/privacy-protection/863064/deepfake-an-assessment-from-the-perspective-of-data-protection-rules
- https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2021)690039
- https://arno.uvt.nl/show.cgi?fid=156861
- https://rm.coe.int/iris-special-2-2020en-artificial-intelligence-in-the-audiovisual-secto/1680a11e0b

Vi viser for øvrig til vår redegjørelse for GDPRs anvendelse på deepfakes i denne utredningen **punkt 2.2.2**.

Når det gjelder DSA – Digital Services Act – viser Europaparlamentets utredningstjeneste til at forordningen ble vedtatt for å ansvarliggjøre plattformer, og at den vil gi EU større mulighet til å regulere innhold på internett, herunder manipulativ adferd (f.eks. deepfake videoer, botter og falske brukere). DSAs implementering i 2024 er angitt å ville medføre nye tiltak for å begrense risiko.

Vi tilføyer her at forordningen i henhold til formålsangivelsen skal bidra til å styrke det indre marked ved å harmonisere regler som bidrar til et trygt og sikkert miljø på internett.[67] DSA-forordningen omfatter alle slags tilbydere av digitale tjenester, men omfanget av forpliktelsene er avhengig av størrelse og typen av tilbyder.[68] Forordningen har regler som skal forebygge ulovlige og skadelige aktiviteter på internett og spredning av desinformasjon.[69]

Det er i svaret fra Europaparlamentets utredningstjeneste opplyst at DSA ikke eksplisitt nevner deepfakes, men at forordningen inneholder bestemmelser som kan være av relevans i forbindelse med spredning av slikt materiale. Det er konkret vist til artikkel 35 bokstav k.

Etter artikkel 35 har veldig store plattformer plikt til å iverksette rimelige, forholdsmessige og effektive foranstaltninger tilpasset risikoene nevnt i artikkel 34 (i denne bestemmelsen nevnes at det skal foretas en risikovurdering som blant annet skal inkludere systemiske risikoer for utbredelse av ulovlig innhold og enhver aktuell eller forventet innvirkning på grunnleggende rettigheter, samfunnsdebatten og alvorlige negative konsekvenser for personens fysiske velbefinnende). Foranstaltningene kan ifølge bokstav k omfatte:

> *"(...) ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information."*

I svaret fra Europaparlamentets utredningstjeneste er det fremhevet at nevnte plattformer snart må identifisere KI-generert materiale for å beskytte det forestående valget i EU mot desinformasjon, se informasjon om ny veiledning i følgende pressemelding fra Kommisjonen:

---

[67] DSA-forordningen artikkel 1.
[68] Se informasjon på regjeringens nettsted.
[69] Se informasjon på regjeringens nettsted.

- Commision is gathering viewes on draft DSA guidelines for election integrity (8. februar 2024)

Det er fremhevet at dette er de første retningslinjene under artikkel 35 om "best pracises" og mulige foranstaltninger for å begrense systemiske risikoer som kan true integriteten til en demokratisk valgproses.

Vi tilføyer her at DSA – ifølge Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) – også kan være relevant i forbindelse med kravet i artikkel 16 om at digitale plattformer skal ha mekanismer for varsling og fjerning av ulovlig innhold, se utredningen s. 41.[70] DSA-forordningens regler om innhold på digitale plattformer kan derfor få betydning for spredning av deepfakes. Etter reglene skal individer ha mulighet for å varsle om ulovlig innhold, og deretter er det opp til plattformene å foreta nødvendige skritt. I regjeringens posisjonsnotat er det forklart følgende:

> *«Reguleringen innebærer at internettbaserte plattformer og særlig svært store internettbaserte plattformer pålegges en rekke krav. Herunder plikter om å innføre mekanismer som kan effektivisere anmeldelse og fjerning av ulovlig innhold, og gi en begrunnelse til brukeren om hvorfor noe fjernes, gi klageadgang, rapporteringskrav om hvor mye innhold som er blitt fjernet og redigert og sikre at selgere på nett kan spores på handelsplattformer. De største internettbaserte plattformene skal dessuten utarbeide risikovurderinger ifm. systemiske trusler mot samfunnet, og dele data med myndigheter og forskere.»*

Det har blitt innvendt at det ikke er klart hva en plattform skal anse som «ulovlig innhold», samt at reglene har en «reaktiv tilnærming» til deepfakes» ved at de retter seg mot eksisterende ulovlig innhold, og ikke produksjon.[71]

I ECPRD-svaret fra Europaparlamentets utredningstjeneste er det vist til fortalepunkt 87, hvor det er angitt følgende:

> *" (...) Providers of very large online platforms, in particular those primarily used for the dissemination to the public of pornographic content, should diligently meet all their obligations under this Regulation in respect of illegal content constituting cyber violence, including illegal pornographic content, especially with regard to ensuring that victims can effectively exercise their rights in relation to content representing non-consensual sharing of intimate or manipulated material through the rapid processing of notices and removal of such content without undue delay. Other types of illegal content may require longer or shorter timelines for processing of notices, which will depend on the facts, circumstances and types of illegal content at hand. Those providers may also initiate or increase cooperation with trusted flaggers and organise training sessions and exchanges with trusted flagger organisations."*

[70] Se også informasjon på regjeringens nettsted.
[71] Se artikkelen Deepfakes and the Law: Are we protected?, publisert på nettsidene til Times of Malta, 27. August 2023.

Vi tilføyer at DSA-forordningen trådte i kraft i EU 16. november 2022, og gjelder fra 17. februar 2024 alle digitale plattformer.[72] Kommisjonen pekte i april 2023 ut 19 veldig store plattformer og søkemotorer, som har vært underlagt regelverket siden august 2023.[73]

Når det gjelder innlemmelse i EØS-avtalen, har departementet uttrykt at forordningen anses EØS-relevant, og det formelle arbeidet med å vurdere akseptabiliteten og eventuelt behov for tilpasningstekst har startet.74 I vurderingen av forordningen på regjeringens nettsider omtales forordningen positivt, og at det vil være en fordel for Norge at forordningen inntas i EØS-avtalen.[75]Mer informasjon om innholdet i DSA-forordningen er tilgjengelig bl.a. på regjeringen.no og i Stortingets EU/EØS-nytt.

5.7.2 Pågående lovinitiativ. Særlig om KI-forordningen

Når det gjelder pågående lovgivningsinitiativ, vises det i ECPRD-svaret til KI-forordningen og den nylig oppnådde enigheten knyttet til direktivforslaget om å bekjempe vold mot kvinner og vold i hjemmet. Begge adresserer deepfakes.

Når det gjelder KI-forordningen, fremholder Europaparlamentets utredningstjeneste at det ble oppnådd midlertidig enighet mellom Rådet og Europaparlamentet 9. desember 2023. Vi tilføyer Europaparlamentet i mars vedtok KI-forordningen. Forordningen må imidlertid også bli vedtatt av Rådet og bli publisert i den europeiske unions tidende før den kan tre i kraft.[76]

I svaret fra Europaparlamentets utredningstjeneste uttrykkes det at grunntanken bak KI-forordningen er å regulere KI-systemer basert på systemets potensiale til å forårsake skade mot samfunnet. Reguleringen har en risikobasert tilnærming, slik at det gjelder strengere regler jo større risiko et KI-system har for å forårsake skade.

Det uttrykkes videre at deepfakes er gjenstand for gjennomsiktighetsbestemmelser. Så vidt vi kan se, fremgår dette av artikkel 50 (4) i forslaget som ble vedtatt av Europaparlamentet:

> «***Deployers*** *of an AI system that generates or manipulates image, audio or video content* ***constituting a deep fake, shall disclose that the content has been artificially generated or manipulated. This obligation shall not apply where the use is authorised by law to detect, prevent, investigate or prosecute criminal offence. Where the content forms part of an evidently artistic, creative, satirical, fictional analogous work or programme, the transparency obligations set out in this paragraph are limited to disclosure of the existence of such*** *generated or manipulated* ***content in an appropriate manner that does not hamper the display or enjoyment of the work****.*
>
> ***Deployers of an AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest shall disclose that the text has been artificially generated or manipulated. This obligation*** *shall not apply where the use is authorised by law to detect, prevent, investigate or prosecute criminal offences or* ***where***

[72] Se informajson på regjeringens nettsted om DMA og DSA
[73] Se informasjon på regjeringens nettsted om DMA og DSA.
[74] Se informasjon i regjeringens posisjonsnotat vedrørende DSA.
[75] Ibid.
[76] Om forordningens ikrafttredelsesbestemmelser, se pressemelding på nettsidene til Europaparlamentet.

the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the *publication of the content*.»

I neste punkt (50 (5)) er det presisert:

> "*The information referred to in paragraphs 1 to 4 shall be provided to the natural persons concerned in a clear and distinguishable manner at the latest at the time of the first interaction or exposure. The information shall conform to the applicable accessibility requirements.*"

Vi antar at "Deployer" skal tolkes som enhver fysisk eller juridisk person, bortsett fra der hvor AI-systemer brukes i personlig, ikke-profesjonell sammenheng, jf. artikkel 3.

Europaparlamentets utredningstjeneste viser videre til uttalelser i fortalen, som etter Europaparlamentets vedtakelse lyder slik i punkt 134:

> "*Further to the technical solutions employed by the providers of the system, deployers, who use an AI system to generate or manipulate image, audio or video content that appreciably resembles existing persons, places or events and would falsely appear to a person to be authentic (deep fakes), should also clearly and distinguishably disclose that the content has been artificially created or manipulated by labelling the artificial intelligence output accordingly and disclosing its artificial origin The compliance with this transparency obligation should not be interpreted as indicating that the use of the system or its output impedes the right to freedom of expression and the right to freedom of the arts and sciences guaranteed in the Charter, in particular where the content is part of an evidently creative, satirical, artistic or fictional work or programme, subject to appropriate safeguards for the rights and freedoms of third parties. In those cases, the transparency obligation for deep fakes set out in this Regulation is limited to disclosure of the existence of such generated or manipulated content in an appropriate manner that does not hamper the display or enjoyment of the work, including its normal exploitation and use, while maintaining the utility and quality of the work. In addition, it is also appropriate to envisage a similar disclosure obligation in relation to AI-generated or manipulated text to the extent it is published with the purpose of informing the public on matters of public interest unless the AI-generated content has undergone a process of human review or editorial control and a natural or legal person holds editorial responsibility for the publication of the content.*"

Når det gjelder direktivet om å bekjempe vold mot kvinner og vold i hjemmet, ble det nylig kunngjort at det er oppnådd enighet mellom EUs lovgivere. Dette omfattet blant annet å kriminalisere deling av intime bilder, inkludert deepfakes, som en form for "cyber violence". Se følgende fra Kommisjonens pressemelding 6. februar 2024 (vår utheving):

> *"Moreover, the most widespread forms of cyber-violence will be criminalised under the new rules, including the non-consensual sharing of intimate images (**including deepfakes**), cyber-stalking, cyber-harassment, misogynous hate speech and "cyber-flashing."*

Mer informasjon om direktivet er tilgjengelig i Europaparlamentets oversikt over prosessen.

Overordnet krever direktivforslaget at medlemsstatene gjør det straffbart å gjøre tilgjengelig intime bilder, eller videoer eller annet materiale som viser seksuelle aktiviteter, med «a multitude of end-users», uten samtykke fra den det gjelder. I fortalepunkt 19, slik det lyder i kommisjonens forslag, er det uttrykt:

> *"The offence should also include the non-consensual production or manipulation, for instance by image editing, of material that makes it appear as though another person is engaged in sexual activities, insofar as the material is subsequently made accessible to a multitude of end-users, through information and communication technologies, without the consent of that person. Such production or manipulation should include the fabrication of 'deepfakes', where the material appreciably resembles an existing person, objects, places or other entities or events, depicting sexual activities of another person, and would falsely appear to others to be authentic or truthful. In the interest of effectively protecting victims of such conduct, threatening to engage in such conduct should be covered as well."*

I ECPRD-svaret fra Europaparlamentets utredningstjeneste er det uttrykt at det vil regnes som straffbart under artikkel 7 å publisere deepfakes, ettersom det vil innebære å bruke informasjons- og kommunikasjonsteknologier til å gjøre eksplisitt seksuelt innhold tilgjengelig for allmenheten uten samtykke fra de som er involvert. Ifølge utredningstjenesten gjelder dette dersom handlingen er «likely to cause serious harm».

Europaparlamentets utredningstjeneste viser videre til at plattformer er adressert i fortalepunkt 40 om hvilke beskyttelsesmekanismer medlemsstatene bør implementere for å beskytte ofrene i forbindelse med materiale som er delt med flere sluttbrukere:

> "*Those measures should include, in particular, empowering national judicial authorities to issue orders to providers of intermediary services to remove, or also to disable access to, one or more specific items of the material in question. Those orders should be issued upon a sufficiently reasoned and substantiated request of the victim. Considering the speed with which such material can spread online and the time it can take to complete criminal proceedings against the persons suspected of having committed the relevant offences, it is necessary for the effective protection of the victims' rights to provide for the possibility of issuing, subject to certain conditions, such orders by means of interim measures, even prior to the termination of such criminal proceedings*."

Vi er dog usikre på om direktivet kan anses EØS-relevant, ettersom strafferetten normalt faller utenfor EØS-avtalens virkefelt.

5.8 Canada

Vi mottok et nokså utfyllende ECPRD-svar fra Canada. I det følgende gjengir vi det vi anser som mest sentralt, men vi kan kontaktes om det er ønskelig med ytterligere informasjon.

I ECPRD-svaret fra Canada uttrykkes det at det ikke eksisterer lovgivning som spesifikt regulerer deepfakes. Det er imidlertid vist til en publikasjon av parlamentets utredningstjeneste av 2019, som ser på hvordan eksisterende lovregulering kan komme til anvendelse på deepfakes, samt eksisterende «gaps»:

- B.J. Siekierski, Deep Fakes: What Can Be Done About Synthetic Video and Audio?, Publication no. 2019-11-E, Library of Parliament, 8 April 2019.

Se særlig punkt 2 for en oversikt over eksisterende lovgivning.

For øvrig gjør Canadas utredningstjeneste oppmerksom på at provinsene kan ha egen lovgivning. Som eksempel nevnes at myndighetene i British Columbia har vedtatt «Intimate Images Protection Act», som gir individer mulighet til å anmode «a tribunal» om å fjerne uønskede, intime bilder av dem på internett. Det er bemerket at lovgivningen omfatter blant annet deepfakes. 8 provinser har vedtatt lover om intime bilder, men innholdet av disse varierer.

For øvrig er det nevnt to pågående lov-initiativer som kan inkludere regulering av deepfakes: The Online Harms Bill og «Artificial INtelligence and Data Act».

5.9 Kort om USA

Vi mottok ikke svar fra USA. Ved et overordnet søk ser det imidlertid ut til at det arbeides med en ny føderal lov kalt «*Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2024 (DEFIANCE Act)*». Loven er blant annet omtalt i Time-artikkelen «How a New Bill Could Protect Against Deepfakes», publisert 31. januar 2024. Ifølge artikkelen skal ofre for deepfake-misbruk, etter lovforslaget, ha adgang til å anlegge sivilt søksmål mot personene som står bak handlingen:

> *"The Disrupt Explicit Forged Images and Non-Consensual Edits, or DEFIANCE Act, allows victims to sue if those who created the deepfakes knew, or "recklessly disregarded" that the victim did not consent to its making."*

Det konkrete lovforslaget ser ut til å være tilgjengelig på Kongressens nettside her. Så vidt vi forstår, er dette et forslag om å gjøre endringer i *en annen lovs bestemmelser* om adgang til å anlegge søksmål i forbindelse med tilgjengeliggjøring av intime bilder («intimate visual depictions»), se 15 U.S.C. 6851(a). Det ser ut til at sistnevnte lovbestemmelser allerede ga vern i forbindelse med tilgjengeliggjøring av intime bilder, og at endringene ved Defiance Act dreier seg om å innføre lignende bestemmelser for *digitale forfalskninger* («digital forgery»). Hva som regnes som intime bilder/forfalskninger, er nærmere definert i loven.

Lovforslaget vil – så vidt vi kan se – grovt sagt gi adgang til å anlegge søksmål sivilt om erstatning mot personen som har produsert og/eller tilgjengeliggjort deepfakes, samt om at denne skal få pålegg om å bringe til tilgjengeliggjøringen til opphør. Vi forstår det slik at lovforslaget rammer både ulike tilgjengeliggjøringshandlinger knyttet til deepfakes og produksjon av deepfakes.

I nevnte Time-artikkel er det for øvrig påpekt at 10 delstater ser ut til å ha straffebestemmelser som vedrører deepfakes.

5.10 Kort om Kina

I Tackling deepfakes in European policy (Europaparlamentets utredningstjeneste, 2021) er det gitt en redegjørelse for reguleringer i enkelte utvalgte land. Blant annet er det vist til at Kina i 2020 vedtok lovregler om at deepfake-materiale må merkes som sådant av app-tilbydere, se punkt 6.9.3 i utredningen. Plattformtilbydere er forpliktet til å uavhengig verifisere og merke eller fjerne umerket materiale. Ifølge loven er det videre forbudt å produsere og spre falske nyheter, og dette må slettes umiddelbart når det oppdages. [77] I Lov&Data-artikkelen «Rettslig regulering av AI – ulike tilnærminger», av Halvor Manshaus i 2023, er det videre vist til at Kina i februar 2023 innførte regler om at det må innhentes samtykke fra personer hvis persondata blir manipulert i deepfakes.

[77] Ettersom dette er et veldig dynamisk område er det grunn til å fremheve at utredningen er fra 2021.