

Betydningen av menneskelige forutsetninger ved ansvar for bruk av kunstig intelligens etter den norske culpanormen

1. Nærmere om artikkelens problemstilling og ulike hovedpunkter

Det er få systemer som baseres på eller styres av kunstig intelligens (KI-systemer) som i dag fungerer uten noen form for menneskelig overvåkning og kontroll. Slik selvlærende algoritmer fungerer på noen områder i dag, kan brukeren likevel i begrenset grad hindre risikoen for uforutsigbare skadeutslag. Dette kan skape utfordringer knyttet til grunnforutsetninger for den ulovfestede skyldregelen. Regelen forutsetter at brukeren kan handle annerledes og gripe inn for å unngå skade. For visse handleplikter innebærer det kunnskap om konkret risiko og antakelser bygget på erfaringsbasert påregnelighet. I dag er det lite allmenn kunnskap om KI-systemenes innebygde risikopotensial på ulike bruksområder. I noen situasjoner kan KI-teknologien derfor tenkes å utelukke det ulovfestede skyldansvaret.

Allmenn kunnskap om teknologiens *uforutsigbarhet* kan imidlertid også tale for et strengere ansvar for brukerne. EU-kommisjonen har på denne bakgrunn foreslått et nytt direktiv om KI-ansvar (AI-ansvarsdirektivet) som inneholder regler om bevis tilgang og en årsakspresumsjoner i situasjoner der informasjon etter pålegg ikke fremlegges, eller brukeren har brutt relevante handleplikter.¹ Forslaget forutsetter ansvar basert på skyld, og holder slik sett fast i utgangspunktene i nasjonale regler. Justeringen av skyldregelen i relasjon til KI-ansvaret må derfor uansett ta hensyn til alminnelige erstatningsrettslige utgangspunkter. I den grad systemet kan kontrolleres, kan det uansett tale for nye handleplikter og et skjerpet aktsomhetskrav for brukeren, som på sikt eventuelt kommer i tillegg til direktivets harmonisering av prosessuelle regler.

Ettersom det i norsk teori per i dag er skrevet lite om KI-systemer og alminnelige culpaansvaret,² kan det ha generell interesse å gi en oversikt over enkelte grunnleggende problemstillinger som KI-teknologien reiser. Et gjennomgående spørsmål er om særtrekk ved den nye teknologien medfører at skadelidte i praksis får et svakere erstatningsrettslig vern enn i saker hvor en skade forårsakes av konvensjonell teknologi.

For det første fører enkelte særtrekk ved teknologiens funksjonsmåte til at tilordningen av ansvar til en brukerkategori oftere kan by på vanskeligheter enn ved bruk av konvensjonell teknologi. Noen utgangspunkter for kravene til tilknytningen mellom brukeren og risikoen for skade, skal derfor klargjøres i artikkelens punkt 2.

¹ Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) 28.9.2022 COM (2022) 496 final. Direktivet er foreløpig satt på vent, og det er uklart om EU-organene ønsker å gå videre med forslaget.

² Se imidlertid, A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», *Tidsskrift for erstatningsrett, forsikringsrett og trygderett* (TfE), 2023 s. 7 flg. Herværende working paper er en del av grunnlagsarbeidet for artikkelen.

Gjennomgripende innføring av nye risikokilder setter ofte forklaringsverdien til erstatningsrettens grunnleggende hensyn på prøve. I punkt 3 skal det diskuteres i hvilken grad erstatningsrettens grunnleggende hensyn kan bidra til å understøtte presiserende argumentasjonslinjer for innholdet i ansvaret.

I punkt 4 skal enkelte grunntrekk i skyldregelen undersøkes nærmere. I punkt 4.1 skal det vurderes i hvilken grad sikkerhetsregler og godkjenningsordninger påvirker aktsomhetsvurderingen og aktsomhetsstandarden. I punkt 4.2 skal innholdet i brukerens eventuelle handleplikter behandles nærmere, herunder hvordan aktsomhetsstandarden påvirkes av livsområdet og forventningene til skadelidte selv. Artikkelen punkt 4.3 vil bli viet til enkelte kognitive utfordringer som potensielle skadevoldere kan oppleve i møtet med kunstig intelligens, og i hvilken grad det kan tale for en justering av aktsomhetsstandarden.

2. Forutsetningen om at skadevolderen har en tilstrekkelig tilknytning til skaderisikoen

Det må anses som et grunnleggende rettslig krav både etter den ulovfestede skyldregelen og etter det ulovfestede objektive ansvaret at den påståtte skadevolderen har en tilstrekkelig tilknytning til den ansvarsbetingende risikoen for at vedkommende skal bli ansvarlig.³ I den alminnelige erstatningsretten kommer dette kravet sjelden på spissen. Kjernen i tilknytningskravet kan som fremholdt av Nygaard anses et krav om at den påståtte skadevolderens må være *rette vedkommende til å gripe inn*, og et rettssubjekt som *har et visst herredømme over risikoen* knyttet til den ansvarsbetingende faktoren.

I anbefalingene til Expert Group on Liability and New Technologies, som er et viktig grunnlagsarbeid for valgene i AI-ansvardsdirektivet, fremsettes den rettspolitiske premiss at: «Strict liability should lie with the person who is in control of the risk connected with the operation of emerging digital technologies and who benefits from their operation.»⁴ Dette samsvarer med grunntanken bak tilknytningskravet i norsk rett både etter den ulovfestede skyldregelen og det ulovfestede objektive ansvaret. I Nordsjødykkerdommen i Rt. 2009 s. 1237 avsnitt 70 understrekes det imidlertid at minstekravet til tilknytning skal baseres på en *normativ vurdering*, der «arten og styrken av tilknytningspunkter» er avgjørende for vurderingen. Om tilknytningskravet er oppfylt beror, ifølge Høyesterett, på en bred helhetsvurdering.⁵

Det er generelt lagt til grunn i rettspraksis at *eieren* av en risikokilde har en tilstrekkelig tilknytning.⁶ I noen tilfeller kan det imidlertid stilles spørsmål ved om herredømmet over risikokilden i så stor grad er overtatt av *andre brukere* at disse bør tilordnes ansvaret. Det er

³ Det fremgår blant annet av Nordsjødykkerdommen Rt. 2009 s. 1237.

⁴ Expert Group on Liability and New Technologies New Technologies formation, “Liability for Artificial Intelligence and Other Emerging Digital Technologies”, 2019, DOI: 10.2838/573689, i M. A. Geistfeld/E. Karner/B. Koch/C. Wendehorst (eds.) *Civil Liability for Artificial Intelligence and Software* 2023 s. 327 key findings number 10.

⁵ Se tilsvarende om disse poengene fra dommen i A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», TFE 2023, s. 7 flg. på s. 28.

⁶ N. Nygaard, *Skade og ansvar* 2007 s. 184.

grunn til å tro at dette spørsmålet kommer oftere på spissen ved erstatningskrav for KI-systemer, blant annet fordi produsenten bevarer en viktig del av herredømmet over risikoen selv om KI-systemet er i eierens besittelse. Flere potensielle skadevoldere kan også anses å være brukere på samme tid.

Utover produsenten kan en bruker i praksis typisk være en tilbyder og en sluttbruker. En brukers herredømme over risikoen kan bestå i å bestemme over i hvilke omgivelser algoritmen skal brukes, som samtidig inkluderer hvilke interesser som skal berøres av risikoen. For mange sluttbrukere av gjenstander med integrert kunstig intelligens, vil imidlertid herredømmet være begrenset til å bestemme når systemet skal aktiveres. Dersom KI-systemet er integrert i en gjenstand som skal styres og overvåkes under bruk, vil kontrollsystemer og design sette viktige premisser for brukerens kontroll over risikokilden.

Det sentrale vurderingstemaet for tilknytningskravet synes i rettspraksis å være om brukeren har overtatt kontrollen med sentrale forutsetninger for systemets forsvarlige drift.⁷ Enkelte dommer kan tolkes slik at tilknytning kun foreligger når brukeren kan påvirke risikonivået under driften.⁸ Dersom bruksområdet for KI-systemet forutsetter offentlig godkjenning, og sluttbruker først og fremst bestemmer tidspunktet for når et høyautomatisert system skal aktiveres, vil den sistnevnte typen kontroll først og fremst ligge hos forhandler og produsent. Men eieren kan også få et særlig ansvar for driften på bestemte bruksområder, og slik sett beholde ansvaret. Et eksempel er særreguleringer av selvkjørende biler. Det stilles krav om at eieren sikrer at brukeren har de tilstrekkelige kvalifikasjoner og at det finnes en person som kan overvåke operasjonen.⁹

Hagstrøm og Stenvik anfører at et viktig moment for en tilstrekkelig tilknytning bør være om den påståtte skadevolderen har oppfordring og reell mulighet til å pulverisere potensielle skadekostnader.¹⁰ På samme måte som for en eier kan situasjonen være slik for en bruker av et KI-system med en langvarig leasing eller leieavtale. I lys av grunntanken bak tilknytningskravet fremstår det likevel som mer sentralt om brukeren har innflytelse på *hvor* systemet skal brukes og overfor *hvem og hvor mange* som *eksponeres* for systemets risikoer.¹¹

Det sistnevnte er særlig sentralt for skillet mellom brukerkategoriene *front end user* og *back end operator* i forslaget til EU-parlamentets forordning fra 2017 og i den nevnte

⁷ Se for eksempel Løftekrandommen Rt. 1969 s. 109. Se tilsvarende, A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», TtE, 2023 s. 7 flg. på s. 28.

⁸ V. Hagstrøm/A. Stenvik, *Erstatningsrett*, 2019 s. 211. Se til dette, A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», TtE 2023 s. 7 flg. på s. 28.

⁹ Tilsvarende A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», TtE 2023 s. 7 flg. på s. 29. Se forskrift om utprøving av selvkjørende motorvogn 17. desember 2017 § 11. Se også det generelle kravet for profesjonelle brukere i AIA art. 26, nr. 2 (som er en kategori som både omfatter innehavere/eiere og aktive brukere, se art. 3 nr. 4).

¹⁰ Se om disse momentene generelt i relasjon til tilknytningskravet V. Hagstrøm/ A. Stenvik, *Erstatningsrett* s. 211.

¹¹ Tilsvarende som i dette avsnittet i A.M. Frøseth, «Erstatningsansvar for brukere av kunstig intelligens», TtE 2023 s. 7 flg. på s. 29.

ekspertgruppens anbefalinger.¹² Man har sett behov for nye begrep på KI-området, fordi produsenten eller forhandleren fortsetter å sette sentrale premisser for risikonivået også etter at produktet er tatt i bruk. Den førstnevnte er personen (inkludert juridiske personer) som tar i bruk produktet etter dets formål og som får fordeler av bruken. En *back end operator* er personen som bestemmer de ulike elementene i teknologien, og som gir essensiell og kontinuerlig støtte til systemets funksjoner og bruk.¹³ Støtten kan bestå i nødvendige oppdateringer og digital service. Den sistnevnte rollen vil som regel tilligge produsenten eller forhandleren. Det er grunn til å tro at *en back end operator* vil få kontrollen tilnærmet alene ved fullautomatiserte systemer. Det kan derfor i enkelte tilfeller være tvil om eieren har tilstrekkelig tilknytning til skaderisikoen ved eierposisjonen i slike tilfeller. Produktansvaret vil disse tilfellene få økt betydning.

I denne artikkelen er det særlig sluttbrukerens ansvar det skal fokuseres på, som i ovenfor nevnte terminologi som regel vil være en *front end operator*.

3. Teknologiens møte med erstatningsrettens grunnleggende hensyn

Per i dag finnes det ingen norsk lovfestet særregulering om erstatningsansvar ved bruk av kunstig intelligens.

Dersom det kommer opp en sak for norske domstoler, må domstolen bygge på de alminnelige erstatningsreglene. Vurderingstemaene under disse reglene må tilpasses skadesituasjonene relatert til teknologien i den grad reglene og rettskildene tillater det. Grunnvilkårene for erstatning er i utgangspunktet nokså fleksible, særlig hva gjelder ansvarsgrunnlagene og adekvansreglene. Et kjennetegn ved de erstatningsrettslige vurderingene er imidlertid at de i praksis er svært konkrete. De overordnede rettssetningene vil ligge nokså fast¹⁴, men det vil være få saker som kan sammenlignes, og som derfor har presisert innholdet i reglene for skadesituasjoner som involverer kunstig intelligens.

Legislative hensyn har en sentral rolle når nye tolkninger og presiseringer av reglene skal begrunnes. Skjønsmarginen for domspolitiske vurderinger¹⁵ er særlig romslig når ny samfunnsutvikling setter behovet for rettsutvikling på spissen.¹⁶ Spørsmålet er hvordan de legislative hensynene kan gi bidrag til utvikling av nye argumentasjonslinjer når innholdet i ansvarsgrunnlagene skal relateres til risikoprofilen for AI-teknologien.

Når ny teknologi skaper nye risikoprofiler, påkalles interessen for erstatningsrettens reparasjonshensyn. Et spørsmål kan være om bruk av teknologi med en ny risikoprofil kan

¹² Expert Group on Liability and New Technologies New Technologies formation, Liability for Artificial Intelligence and Other Emerging Digital Technologies, *Civil Liability for Artificial Intelligence and Software*, s. 364-365.

¹³ Se om begrepene, C. Wendehorst, «Strict liability for AI and other Emerging Technologies», *Journal of European Tort Law*, 2020 nr. 2 s. 150 flg. på s. 174-176.

¹⁴ Se i samme retning, N. Nygaard, *Skade og ansvar*, 2007 s. 15.

¹⁵ Se om dette begrepet, N. Nygaard, *Aktløyseavurderinga i norsk rettspraksis*, 1974 s. 30-31.

¹⁶ Se i samme retning, N. Nygaard, *Skade og ansvar*, 2007 s. 16-17.

tale for erstatning i seg selv. Slike synspunkter kan spores i fransk rett, der hensynet til skadelidtes dekningsbehov på enkelte områder fremstår som det sentrale argumentet for et objektivt ansvar.¹⁷ Argumentet kombineres gjerne med synspunkter om at den som skaper risikokilden, og/eller får fordelene av den, bør bære tapet istedenfor den tilfeldige skadelidte.¹⁸ Argumentasjonsmåten var også sentral ved innføringen av det ulovfestede objektive ansvaret for farlige bedrifter, og har vært viktig for skjerping av ansvaret gjennom ansvarspresumsjoner og snudd bevisbyrde i internasjonal sammenheng.¹⁹

Reparasjonshensynet er imidlertid i liten grad konkret artikulert og operasjonalisert i norsk domspraksis. Hensynet kan beskrives som kjernen i erstatningsrettens *funksjon* i rettssystemet, men det sier ikke noe om *når* gjenoppretting bør skje og på *hvilke betingelser*.²⁰ Essensen i reparasjonstanken er i praksis at skaden bør dekkes av den som er årsak til skaden.²¹ Tapet forsvinner ikke, men flyttes fra skadelidte til skadevolderen, fordi sistnevnte har gitt et årsaksbidrag til en påregnelig skade.²² I rettsfilosofiske teorier med utgangspunkt i Aristoteles tanker om korrektiv rettferdighet har erstatningsrettens rettferdighetsgrunnlag fått en ytterligere utdypning. Skadelidtes rettighet til det krenkede godet består også etter skaden, og gir grunnlag for et krav om utlikning av ubalansen.²³ Etter inspirasjon fra Kants teori har Weinrib fremhevet at de erstatningsrettslige ansvarsvurderinger også handler om å balansere skadevolders og skadelidtes generelle handlefrihet i en interesseavveining.²⁴ *Kvaliteten* på skadevolders tilknytning til skaden må imidlertid avgjøres etter sosialpsykologiske identifikasjonskriterier, der faktisk eller rettslig herredømme over risikofaktorene og risikoutviklingen er sentrale momenter.²⁵ Der den påståtte skadevolderen i liten grad har herredømme over risikokilden, vil begrunnelsen for ansvaret svekkes.

Erstatningsretten kan ikke forklares gjennom reparasjonshensynet og gjenopprettingstanken alene, men bygger på en kombinasjon av flere hensyn. Fra gammelt av har prevensjonshensynet vært fremhevet som et sentralt hensyn, og hensynet har en særlig plass når man ønsker å forebygge mot fremtidig skade. I norsk og internasjonal teori er det imidlertid påpekt at prevensjonshensynet isolert sett har en del av de samme begrensninger som reparasjonshensynet.²⁶ Hensynet forteller først og fremst noe om en antatt bieffekt av at

¹⁷ M. Cappelletti, *Justifying Strict Liability*, 2022 s. 73-84.

¹⁸ I norsk rett er dette hensynet blant annet fremhevet av T-L. Wilhelmsen/B. Hagland, *Om erstatningsrett – med utgangspunkt i tekster* av Peter Lødrup 2017 s. 48-49.

¹⁹ Se Expert Group on Liability and New Technologies New Technologies formation, *Liability for Artificial Intelligence and Other Emerging Digital Technologies*, 2019, DOI: 10.2838/573689 i M. A. Geistfeld/E. Karner/B. Koch/C. Wendehorst (eds.) *Civil Liability for Artificial Intelligence and Software* 2023 s. 357.

²⁰ H. Koziol, «Comparative conclusions» i *Basic questions of Tort Law from a Comparative Perspective*, 2015 s. 748. Koziol understreker imidlertid at det underliggende rettferdighetsgrunnlaget for reparasjonshensynet – korrektiv rettferdighet, gir en tydelig base for prinsippet om full erstatning. Se til begrensningene i forklaringsverdien til teorier om korrektiv rettferdighet, A.M. Frøseth, *Skadelidtes egeneksponering for risiko i erstatningsretten*, 2013 s. 58 flg. Se til reparasjonshensynets begrensninger også V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 81.

²¹ Se tilsvarende N. Nygaard, *Skade og ansvar*, 2007 s. 19.

²² F. Stang, *Erstatningsansvar*, 1919 s. 50-51.

²³ E. Weinrib, *The Idea of Private Law* 2012 s. 80 flg.

²⁴ Se nærmere G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», *VersR* 12, 2022 s. flg på s. 722.

²⁵ Se A.M. Frøseth, *Skadelidtes egeneksponering for risiko i erstatningsretten*, 2013 s. 59 og 70-71.

²⁶ G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 722.

Professor dr. juris Anne Marie Frøseth, Det juridiske fakultet, Universitetet i Bergen, Mars 2024, med henvisninger til Artificial Intelligence Act etter ikrafttredelse 1. august 2024.

erstatningsretten håndheves, fordi pålegg om å betale erstatning antas å motivere potensielle skadevoldere til å avstå fra skadevoldende handlinger.²⁷ Det forutsetter at fremtidig skadeforvoldelse kan beregnes med en viss rasjonalitet. Prevensjonshensynet har derfor først og fremst berettigelse, og empirisk belegg, i relasjon til virksomheter og profesjonelle aktører.²⁸ Både gjenopprettingstanken og prevensjonshensynet må forbindes med en mer normativ tankegang som kan indikere når ansvar skal kunne ilegges.²⁹

Når ny teknologi forandrer samfunnet, fremtrer behovet for å begrunne *risikofordelingen* mellom skadevoldere og skadelidte. Den generelle usikkerheten knyttet til konsekvensene av ny teknologi, gir behov for risikostyring gjennom rettslige virkemidler som erstatningsreglene.³⁰ I denne risikofordelingstankegangen er også samfunnsnyttene relevante.³¹ Gjennomgripende innføring av ny teknologi aktualiserer også spørsmål om hvordan økonomiske byrder bør fordeles på flere aktører i samfunnet.³² Behovet for fordelingsrettferdighet i et samfunnsperspektiv bringer tankene hen mot en velferdsøkonomisk tankegang.³³

Under punkt 2 om kravet til tilknytning ble det vist til synspunktet i internasjonal teori om at skade forårsaket av KI-systemer *bør dekkes av det subjektet som har kontroll over risikoen ved teknologien og som får fordelene av den*.³⁴ Tankegangen er en syntese av tankene om korrektiv rettferdighet og rettsøkonomisk teori. Hvis det potensielt er flere skadevoldere, bør skaden dekkes av den som har størst kontroll over bruken og dens risikoeksponering overfor

²⁷ Slik for eksempel Stang, *Erstatningsansvar*, s. 54 flg.

²⁸ Se N. Jansen, *The Structure of Tort Law*, 2021 s. 128-136, for et nyansert bilde av empiriske undersøkelser knyttet til erstatningsreglers evne til å påvirke fremtidig atferd på ulike livsområder. Undersøkelsene viser at erstatningsregler har en begrenset effekt, og dette rammer også rettsøkonomiske teorier som bygger på en målestokk for effektivitet knyttet til «optimal prevensjon» forankret i skadeevneprinsippet, se videre nedenfor under dette punktet.

²⁹ G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 722.

³⁰ Se H. Zech, «Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?», *Gutachten A zum 73. Deutschen Juristentag Hamburg 2020/Bonn 2022*, 2020 s. 87.

³¹ Denne typen tankegang er ikke like anerkjent i alle land, og har nær sammenheng med hva som anses som erstatningsrettens funksjon og om utligningen av urettferdigheten i topartsforholdet mellom skadevolder og skadelidte er det sentrale, eller om man også anser erstatningsrettens å ha en funksjon også med hensyn til «sosialt entreprenørskap». De land som er tilbøyelige å se på erstatningsretten i sistnevnte betydning inkluderer både Frankrike, Italia og Norge. Disse landene er gjerne også tilbøyelige til å begrunne erstatningsansvar med argumenter som skadelidtes behov (sosiale hensyn), behovet for spredning av tapet (fordelingsrettferdighet eller distributive justice), samt behovet for å forebygge mot skade. De samme land er også kjent for sin relative vektlegging av objektive ansvarsmodeller i sine systemer for deliktsansvar. Se til dette, M. Cappelletti, *Justifying Strict Liability*, s. 291-292. Se om de sistnevnte argumentene og teoriene bak disse i T-L. Wilhelmsen, *Om erstatningsrett – med utgangspunkt i tekster av Peter Lødrup* 2017 s. 55-57.

³² G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 722-723.

³³ H. Eidenmüller, «Machine Performance and Human Failure: How Shall We Regulate Autonomous Machines?» i H. Eidenmüller/G. Wagner, *Law by Algorithm*, 2021 s. 29-31.

³⁴ Dette er en av to prinsipielle tilnærminger som i 2017 ble foreslått i en resolusjon av EU parlamentet – den såkalte «risk based approach», se European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics ([2015/2103\(INL\)](#)) avsnitt 55. Se også European Commission, Report on the safety and liability implications of AI, the internet of Things and Robotics, COM (2020) 64 final. Se ellers, G. Wagner, «Robot Liability» i S. Lohsse/R. Schulze/D. Sautenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things*, 2018, s. 27 flg. på s. 31. Synsmåten minner om Calabresi sin rettsøkonomiske modell «The least cost risk avoider», se en kort og pedagogisk god innføring i denne modellen i TL. Wilhelmsen/B. Hagland, *Om erstatningsrett – med utgangspunkt i tekster av Peter Lødrup*, s. 55.

tredjepersoners interesser.³⁵ Dette er i samsvar med normative betraktninger i norsk teori om erstatningsrettens rettferdighetsgrunnlag, også innenfor topartsforholdet. Essensen i begrunnelsen kommer godt frem hos Hagstrøm og Stenvik:

«Må skadelidte bære tapet, blir han rammet av det uforutsette; trygghets- og sikkerhetsfølelsen blir svekket. Flyttes tapet derimot over på skadevolder, elimineres usikkerheten for skadelidte. Dette er for så vidt ikke et argument for at skadevolderen skal bære tapet, hvis resultatet ville bli en tilsvarende usikkerhet for ham. Men skadevolderen har både kunnet forutse og avverge skadetilfellet. (...) Det kan derfor sies at culpanormen gir uttrykk for et rimelig risikofordelingsprinsipp, hvor risikoen for en skade pålegges den som har hatt mulighet til å forutse at den kan inntre, og som er nærmest til, og har mulighet for, å avverge den.»³⁶

En slik tilnærming er også i samfunnets interesse fordi erstatningsansvaret oppfordrer til forebygging. Men det forutsetter at kostnadene ved skade kan forutsees og beregnes. Denne varianten av rettferdighetsgrunnlaget glir dermed over i prevensjonshensynet, og kan anses som en konkretisering av dette. Det er imidlertid ikke maksimal prevensjon man er ute etter, bare optimal prevensjon: Optimal prevensjon oppnås når kostnadene ved å tillate en skadevoldende handling er mindre enn nytten (for skadevolderen) ved å avverge den.³⁷ I vurdering av nytte er ikke bare økonomiske verdimål relevante, men også hvilken beskyttelse og status det truede godet generelt sett har.³⁸ Tankegangen fungerer altså som en målestokk for når man kan kreve at skadevolderen handler annerledes, og dermed standarden for forsvarlig atferd. I denne målestokken er kostnadene/ulempene som følger skadeevnen det sentrale. Skadeevnen er et produkt av sannsynligheten for at skade vil skje sammenholdt med skadeutfallets alvorlighetsgrad.³⁹ Økonomiske betraktninger har også betydning i vurderingen av objektivt ansvar. Dels skyldes dette at skadeevnekriteriet er sentralt i begrunnelsen.⁴⁰ De økonomiske aspektene som er anført i internasjonal teori, har imidlertid vel så nær forbindelse til interessehensynet og pulveriseringshensynet i den avsluttende helhetsvurdering under det objektive ansvaret.⁴¹

Flere forfattere i internasjonal teori om KI-ansvar har advart mot å gi skadeevnekriteriet for dominerende betydning.⁴² Dette også fordi vurderingen av risikoen ved den konkrete teknologien per i dag vanskelig kan beregnes.⁴³ Slikt kunnskapsmateriale er hvert fall vanskelig tilgjengelig. Skyldansvaret og det ulovfestede objektive ansvaret inneholder heller

³⁵ Se blant annet anbefalingene fra den internasjonale ekspertkomiteen oppnevnt av EU kommisjonen, Expert Group on Liability and New Technologies New Technologies formation, «Liability for Artificial Intelligence and Other Emerging Digital Technologies», key findings number 10.

³⁶ V. Hagstrøm/A. Stenvik *Erstatningsrett* 2019 på s. 83.

³⁷ G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 722.

³⁸ L.c.

³⁹ Se N. Nygaard, *Skade og ansvar*, s. 189 flg. og V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 104.

⁴⁰ Se til dette blant annet V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 190 flg. og N. Nygaard, *Skade og ansvar*, s. 207.

⁴¹ Se til de økonomiske aspektene under helhetsvurderingen relatert til ulovfestet objektivt ansvar, V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 212 flg.

⁴² Se blant andre, A. Bertolini/F. Episcopo, «The Expert Group's Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies: a critical assessment», *European Journal of Risk Regulation*, 12 (2012) s. 644 flg. på s. 549-651, B. Schütte/L. Majewski, «Private Liability for AI-Related Harm: Towards More Predictable Rules for the Single Market», *Marked and Competition Law Review* nr. 1 2022 s. 123 flg. på s. 137.

⁴³ A. Bertolini/F. Episcopo, «The Expert Group's Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies: a critical assessment», s. 644 flg. på s. 650.

ingen minstekrav om alvorlig skade isolert sett. Hvilken *type* skade som *kan* inntre, har likevel en viss betydning i den samlede vurderingen av om ansvar skal pålegges.⁴⁴

Oppsummeringsvis kan vi utlede at kontroll over risikokildens skadeevne er det sentrale argumentet for ansvar, både i legitimerende og begrensende forstand. Den begrensende effekten relaterer seg særlig til Stangs betegnelse av det dypeste grunnlag for begge arter av ansvar: «adækvansen»⁴⁵ -- at skadevolderen har muligheter for å forutse og til dels beregne skaderisikoen *ex ante*. En konsekvens av dette er at «...området for culpanormen (er) de konkret påregnelige og avvergelige skader, mens det objektive ansvaret fanger opp de konkret uavvergelige, men generelt påregnelige skader.»⁴⁶

Hagstrøm og Stenvik understreker at den *konkrete skademåten* ikke trenger å være påregnelig for skadevolderen.⁴⁷ Kjernen i *skyldansvaret* er likevel at skadevolderen har mulighet til å handle annerledes.⁴⁸ Dette forutsetter en viss kunnskap om skaderisikoens karakter og et praktisk handlingsalternativ som gir effektiv risikoreduksjon.⁴⁹ I så måte har den praktiske muligheten til å foreta en alternativ handling også betydning som et minstekrav, og begrensning, i aktsomhetsvurderingen.⁵⁰ I noen tilfeller vil kostnadene med en alternativ handling være relevant.⁵¹ Dette momentet gjør seg også for noen typetilfeller gjeldende i vurderingen av ulovfestet objektivt ansvar.⁵²

I de neste punktene blir det en gjennomgripende problemstilling hvor vidt hensynene til forutberegnelighet og kontroll setter materielle grenser for den ulovfestede skyldregelens rekkevidde.

4. Nærmere om skyldansvaret

4.1 Utgangspunkter

I det følgende forutsettes det at den faktiske årsakssammenhengen er kartlagt, og at kravene til faktisk årsakssammenheng er tilfredsstillt. Spørsmålet blir da om det var uforsvarlig av skadevolderen å ta systemet i bruk, påvirke systemet under bruken, eller å ikke gripe inn.

⁴⁴ Se til dette blant annet V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 190 flg. og N. Nygaard, *Skade og ansvar*, s. 210 flg.

⁴⁵ Stang, *Erstatningsansvar*, s. 48.

⁴⁶ V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 81.

⁴⁷ V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 107.

⁴⁸ Se blant annet N. Nygaard, *Skade og ansvar*, s. 193 flg.

⁴⁹ Jeg er enig med Strandberg i at kravet til gjennomføring av en alternativ handling i mange tilfeller best kan beskrives som et krav om *risikoreduksjon* ned til forsvarlig nivå. Nygaard tar derfor litt for hardt i når han formulerer et hypotetisk krav om at skadevolderen gjennom en alternativ handling ville ha hindret skaden eller skadevolderens årsaksbidrag (se Nygaard, *Skade og ansvar* 2007 s. 193) Se om Nygaards teori og omformuleringen av kravet til "hindringssammenheng" til et krav om risikoreduksjon i M. Strandberg, "Teorien om hindringssammenheng som vilkår for erstatningsansvar", *Tidsskrift for Erstatningsrett, forsikringsrett og velferdsrett* 2012 s. 240 flg.

⁵⁰ Se blant annet N. Nygaard, *Skade og ansvar*, s. 199 flg.

⁵¹ Se blant annet V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 110.

⁵² Se om forsikringsmomentet i N. Nygaard, *Skade og ansvar*, s. 279 flg.

Vurderingen av om det er forsvarlig å ta systemet i bruk påvirkes av hva som kan anses som tillatt risiko på det aktuelle bruksområdet.

4.2 Tillatt risiko

Så lenge en virksomhet eller et produkt antas å være nyttig for samfunnet, er det, uavhengig av risikokildens teknologiske særtrekk, en viss toleranse for skaderisiko. Kravet til en potensiell skadevolder er derfor ikke at skaderisikoen kan kontrolleres til enhver tid.⁵³ I en målestokk for forsvarlig atferd må det imidlertid antas å være innebygget i et minstekrav om fornuftig og effektiv risikohåndtering.⁵⁴

For mange systemer basert på kunstig intelligens vil det etter hvert foreligge offentligrettslige godkjenningsordninger, som eventuelt også inkluderer sikkerhetsregler. Et eksempel på dette er EUs Artificial Intelligence Act (AIA)⁵⁵, som gjennom EØS-avtalen også vil gjelde for Norge.⁵⁶ Produktene godkjennes etter forordningen kun for bestemte formål og omgivelser. Forordningen inneholder detaljerte sikkerhetsregler for godkjenning og bruk av høyrisiko-produkter, men ingen krav til lavrisikoprodukter. De sistnevnte produktenes sikkerhetskrav skal baseres på bransjestandarder jf. art. 95.⁵⁷

Utover dette, kan det utledes fra alminnelige erstatningsrettslige prinsipper at det sentrale for tillatt risiko er hvilke omgivelser teknologien skal brukes i, og hvilket formål den har. Det avgjørende er dermed ikke årsaken til risikoen.⁵⁸ Sammen med skaderisikoens frekvens og mulige skaders alvorlighet, utgjør omgivelsenes karakter produktets skadeevne. Disse aspektene er også sentrale kriterier for hva som anses som høyrisiko-produkter etter Artificial Intelligens Act.⁵⁹

Av EUs AIA kapittel II følger det allerede at enkelte høyrisiko-produkter vil bli forbudt per se på grunn av interessene som rammes. Snakkende dukker styrt av kunstig intelligens er et eksempel. Dersom denne teknologien har en latent risiko for å oppmuntre barnet til farlig lek, er produktet forbudt på markedet.⁶⁰ Risikoen anses her som uakseptabel fordi *bruksområdet* strider mot grunnleggende verdier i det europeiske samfunnet.

I EUs Artificial Intelligence Act angis ulike vilkår for å sette et system i drift, og som retter seg mot tilbydere og profesjonelle sluttbrukere. Det er eksempelvis krav om at tilbyder har foretatt en risikoanalyse med tanke på formålet og de interesser som påvirkes. Det er også krav om at systemet har en loggingsfunksjon og lagringsfunksjon som i etterkant kan vise

⁵³ Se blant annet G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 727.

⁵⁴ G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. flg på s. 727.

⁵⁵ Artificial Intelligence Act 13 June 2024 (regulation (EU) 2024/1689).

⁵⁶ Forutsatt beslutning av EØS-komiteen om å innlemme denne forordningen i EØS-avtalen, noe man må regne med at de gjør.

⁵⁷ Det er også særskilte krav om risikoanalyse og transparens til generative KI-modeller som potensielt representerer systemiske risikoer med høy påvirkningskraft overfor omgivelsene.

⁵⁸ G. Wagner, «Verantwortlichkeit im Zeichen digitaler Techniken», s. 727.

⁵⁹ Se AIAs fortale avsnitt 52.

⁶⁰ Se AIAs fortale avsnitt 29.

hvilke prosesser som var virksomme, og om en årsaksfaktor utenfor systemet påvirket resultatene/beslutningene, jf. artikkel 12 og 19.

Samlet sett må man kunne anta at EU, med det totale reguleringsregime, legger opp til at systemer basert på kunstig intelligens skal ha minst samme sikkerhetsnivå som konvensjonelle systemer styrt av mennesker på de samme bruksområdene.⁶¹ Dette er også et naturlig utgangspunkt for handlepliktene og aktsomhetsstandarden for brukere av autonome systemer, og gjelder særlig der skaderisikoen er høy.⁶²

Nedenfor skal ulike faktorer av sentral betydning for innholdet i skyldregelen og aktsomhetsstandarden analyseres videre. Først skal det i punkt 4.3 diskuteres hvilken betydning systemets tekniske design og lovfestede sikkerhetskrav kan tenkes å ha for innholdet i mulige handleplikter og aktsomhetsstandarden. Handlepliktene og aktsomhetsstandarden må i tillegg justeres etter livsområdet og de forventninger skadevolderen kan ha til potensielle skadelidte, jf. punkt 4.4. Disse momentene har generelt betydning for styrken på skadelidtes vern i erstatningsretten.⁶³ Som tidligere vist, har de nevnte faktorene betydning for vurderingen av skadeevnen sett fra skadevolderens perspektiv. I punkt 4.5 stilles det spørsmål om generelle begrensninger i menneskers kognitive funksjon kan tilsi at det bør settes en lavere aktsomhetsstandard for brukerne.

4.3 Objektivisering av handleplikter og betydningen av sikkerhetsregler

Som påpekt ovenfor, vil en sluttbruker være prisgitt produsentens kontrollsystemer og design i sitt samvirke med KI-systemer. Brukerens mulighet til å regulere systemet ex ante vil være nokså begrenset. Flere forfattere har derfor hevdet at kravet til brukerens handlings- og omsorgsplikter bør reduseres i takt med automatiseringsnivået i systemet.⁶⁴

Det kan likevel tenkes at aktsomhetsterskelen er skjerpet på de få områdene der det er mulig å handle. En skjerpelse av kravet til aktsomhet under bruk, kan særlig tenkes i en overgangsfase der utprøving av teknologien på det aktuelle bruksområdet er ny. Skadevolderens kunnskap om teknologiens generelle uforutsigbarhet er altså en kunnskap som skjerper kravet til aktsomhet.⁶⁵ Kravet til aktsomhet og handleplikter kan også følge av produktinformasjonen og skrevne sikkerhetsregler.

⁶¹ European Commission, Report on the safety and liability implications of AI, the internet of Things and Robotics, COM (2020) 64 final, s. 16-17.

⁶² E.Karner, «Liability for Medical Robots and Autonomous Medical Devices» i E. Karner/U. Magnus/J. Spier/P. Widmer (eds.) *Essays in honour of Helmut Koziol*, 2020 s. 57 flg. på s. 61.

⁶³ N. Nygaard, Skade og ansvar 2007 kap. 1, IV, 6 og kap. 4, I, 1-2.

⁶⁴ Slik J. Pehm, «Systeme der Unfallhaftung beim automatisierten Verkehr – Eine rechtvergleichende Analyse der Haftungsrisiken», *Zeitschrift für Internationales Wirtschaftsrecht*, nr. 2 (2018) s. 259 flg. på s. 263-264. Synspunktet har fått tilslutning i M. Geistfeld/E. Karner/B. Koch/C. Wendehorst, *Civil Liability for Artificial Intelligence and Software* 2023 s. 55.

⁶⁵ Se i samme retning, men uten å tale om et skjerpet aktsomhetskrav på grunn av denne kunnskapen, B. Schütte/L. Majewski, «Private Liability for AI-Related Harm: Towards More Predictable Rules for the Single Market», *Marked and Competition Law Review* nr. 1 2022 s. 123 flg. på s. 133.

Utgangspunktet etter norsk rett er at sikkerhetsregler skal tillegges stor vekt i vurderingen av om brukeren har handlet forsvarlig. Dette forutsatt at det er årsakssammenheng mellom et brudd på regler i lov/forskrift og skaden, og at reglene har som formål å regulere brukerens atferd i relasjon til risiko for skade i den typen situasjoner der skaden oppsto.⁶⁶ Men selv om det foreligger brudd på klare handleregler, må det slik Høyesterett påpeker i HR-2018-1234-A avsnitt 42, «... likevel foretas en konkret vurdering i det enkelte tilfellet».

Vekten bruddet får i vurderingen, beror på flere hensyn. Dette gjelder prinsipielt sett også på områder hvor det stilles strenge krav til aktsomhet, og den strenge normen kan utledes av de skrevne reglene. Dette fremgår eksplisitt blant annet av RIB-dommen, HR-2019-318-A avsnitt 57, hvor Høyesterett påpekte at «... selv om aktsomhetskravet er strengt, medfører ikke ethvert avvik fra optimal atferd at det foreligger ansvarsbetingende uaktsomhet». Og selv om sikkerhetsregler er overholdt, kan det foreligge en oppfordring for skadevolderen til å reagere på risikoen i den aktuelle situasjonen som følge av andre forhold enn de som nevnes i skrevne regler.

Selv om dette er utgangspunktene, vil en generell objektivisering av aktsomhetsstandarden⁶⁷ gi lite rom for subjektive vurderinger av skaderisikoen og skadens påregnelighet. Dette blir forsterket dersom det foreligger en detaljregulering av handleplikter gjennom lov og forskrift, som gir mindre rom for konkrete vurderinger. En måte å se dette på er at et brudd på lovregler, og særlig sikkerhetsregler, i mange tilfeller er tilstrekkelig for ansvar.⁶⁸ En annen måte å artikulere samme tendens på, er å si at detaljert spesialregulering på enkelte områder får så stor relevans og vekt, at det i mange tilfeller ikke er stort rom for en alminnelig forsvarlighetsvurdering i tillegg.⁶⁹

Det er ofte en sammenheng mellom en detaljert regulering av risiko, og at skaderisikoen oppfattes som høy på det aktuelle området. Men det kan også være andre hensyn som foranlediger detaljregulering. Det kan for det første være at en teknologi er ny eller et bruksområde er nytt. Motivasjonen for regulering kan derfor være et behov for forutsigbarhet for alle involverte aktører. Det kan videre være at man ønsker en enhetlig regulering for å fremme salg og innovasjon, og som gjør kostnadsberegningen for utviklerne mer presis og enkel.⁷⁰ Ingen av disse begrunnelsene taler i seg selv for å legge vekt på prevensjonshensynet for å begrunne et strengere ansvar. Det er det først og fremst behovet for harmonisering av rettstilstanden innenfor EU som gjør. Og det sistnevnte behovet ivaretar først og fremst håndhevingssynspunkter, og er gjerne omtalt som effektivitetshensyn.⁷¹

Håndhevingshensynet gjør seg gjennomgående gjeldende på områder som er regulert av lovgivning fra EU. Effektivitetsbetraktninger gir imidlertid ikke rom for en slik skjerpelse av ansvaret at brudd på sikkerhetsregler automatisk fører til ansvar. Dette forutsatt at håndheving

⁶⁶ V. Hagstrøm/A. Stenvik, *Erstatningsrett*, s. 85 flg.

⁶⁷ Se A.M. Frøseth, *Skadelidtes egeneksponering for risiko i erstatningsretten*, s. 82-86.

⁶⁸ Se til objektivisering av culpatornormen "den moderne culpaparegel" og betydningen av lovregler for dansk retts vedkomme i A. Bloch Ehlers, *Om adækvantlæren i erstatningsretten*, 2011 s. 59 flg.

⁶⁹ Se T-L. Wilhelmsen/B. Hagland, *Om erstatningsrett: med utgangspunkt i tekster av Peter Lødrup*, s. 111.

⁷⁰ De nevnte hensyn har motivert de europeiske forslagene til reguleringen av systemer for kunstig intelligens.

⁷¹ T-L. Wilhelmsen/B. Hagland, s. 108.

av EU-reguleringen blir håndhevet med minst samme effektivitet som en særnorsk sammenlignbar regulering.

På bakgrunn av blant annet EUs AIA er det grunnlag for å angi enkelte *typiske* handleplikter en bruker av kunstig intelligens vil få etter skyldregelen. Dette følger blant annet av eksisterende, funksjonelle særtrekk ved teknologien. De plikter som kan utledes av AIA korresponderer i stor grad med oppfatninger i sakkyndige utredninger bestilt av EU organene og juridisk teori. De offentligrettslige kravene til tilbyder får dessuten konsekvenser for systemets design, og designet vil i stor grad sette premissene for hvordan det vil være mulig for sluttbrukeren å samvirke med systemet.⁷² I ansvarsvurderingen vil det i så måte være relevant å se på hvilke kontrollsystemer som er integrert i programvaren.⁷³ Dette vil i sin tur få betydning for hvilken grad av sikkerhet brukeren kan forvente at produsenten/forhandleren har ivare tatt i sin konstruksjon av systemet og den innretningen den er integrert i.

Typiske plikter for sluttbruker kan være å oppdatere programvaren når det foreligger varsel fra systemet, krav om å gjennomføre en innbygget selvtest og forbud mot å endre programvaren og systemets forhåndsinnstillinger. I tillegg er det naturlig å stille krav om at driften av systemet skal overvåkes, se art. 6 nr. 2 og nr. 5.⁷⁴ Selv om brukeren ikke kan modifisere systemets handlinger, kan vedkommende i de fleste tilfeller slå systemet av, jf. tekniske krav i AIA art. 14, nr. 4, bokstav e. Og, i motsetning til den frihet som brukeren vanligvis har i relasjon til konvensjonelle systemer, kan det foreligge en plikt til å avstå fra å gripe inn i grunnleggende betingelser for systemets drift.⁷⁵ Dette også om hensikten med intervensjonen er å reparere. Det følger av den generelle rolleforventningen etter skyldregelen at aktsomhetsstandarden blir skjerpet hvis man beveger seg inn på områder som krever særskilt kompetanse.⁷⁶ Det vil følge av dette at en sluttbruker uten særskilt kompetanse bør overlate reparasjon og oppgradering av komplekse systemer til forhandleren. Slike retningslinjer vil som regel også følge av brukermanualene.

Plikten til å *oppdatere* systemet med nye data henger sammen med behovet for ny informasjon for å holde risikoen for feil på et tilfredsstillende nivå. I et utdatert system vil risikoen for feil øke. Plikten til å *oppgradere* er nært forbundet med en plikt til å sørge for at sikkerheten i systemet utvikler seg i samsvar med forventet standard, også etter skyldregelen. Det er grunn til å tro at kravene til systemenes risikohåndtering vil øke etter hvert som teknologien blir mer sikker, og det vil sannsynligvis forventes at brukeren oppgraderer systemet til det som er tilgjengelig normalstandard.⁷⁷ Et slikt krav vil være spesielt aktuelt for

⁷² J. Pehm, «Systeme der Unfallhaftung beim automatisierten Verkehr – Eine rechtvergleichende Analyse der Haftungsrisiken», *Zeitschrift für Internationales Wirtschaftsrecht*, på s. 263-264.

⁷³ Om betydningen av dette, se A. Bertolini/F. Episcopo, «The Expert Group's Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies: a critical assessment», på s. 651.

⁷⁴ Se oversikt over relevante plikter i M. Thöne, *Autonome Systeme und deliktische Haftung*, 2020 s. 270.

⁷⁵ Dette kan også utledes av det foreslåtte AI-ansvarsdirektivet gjennom at det er en (av flere mulige) betingelse for anvendelse av årsakspresumsjonen overfor profesjonelle brukere, se forslaget art. 4, nr. 6 – med det forbehold at inngrepet er vesentlig.

⁷⁶ Se for eksempel N. Nygaard, *Skade og ansvar*, s. 194 - 195.

⁷⁷ E. Karner, «Liability for Medical Robots and Autonomous Medical Devices» i E. Karner/U. Magnus/J. Spier/P. Widmer (eds.) *Essays in honour of Helmut Koziol* 2020 s. 57 flg. på s. 71-72.

profesjonelle brukere, men vil antakelig forutsette at det ikke medfører en betydelig merkostnad.

Et fellestrekk ved kravene er strenge aktsomhetsplikter knyttet til *forutsetningene* for bruken av systemet. Dette gjelder både før oppstart og gjennom vedlikehold.⁷⁸ Det inkluderer blant annet å sette seg inn i sikkerhetskrav i lov, forskrift og brukermanual.⁷⁹ Generelt kan man imidlertid anta at pliktene til overvåking vil reduseres etter hvert som teknologien på ulike bruksområder blir mer automatisert. Skyldregelen vil naturlig også i takt med dette få redusert sitt praktiske anvendelsesområde.⁸⁰

4.4 Betydningen av livsområdet og forventningene til skadelidtegruppen

På flere områder vil bruken av kunstig intelligens kunne medføre en strengere informasjonsplikt for potensielle skadevoldere overfor potensielle skadelidtegrupper enn ved bruk av konvensjonell teknologi. Dette gjelder blant annet der den kunstige intelligensen brukes som beslutningsassistanse, og alternativet er profesjonell faglig rådgivning.

Dette gjelder for eksempel der produktet benyttes til fondsinvesteringer, medisinske diagnoser eller medisinske analyser⁸¹. Eksempler i den siste kategorien er algoritmer som bruker mønstergjenkjenning for å stille diagnoser basert på røntgenbilder av menneskelig vev. Selv om produktet ikke nødvendigvis innebærer en høyere risiko for skader, skal potensielle skadelidte kunne ta et informert valg knyttet til skaderisikoen. Dette inkluderer fordelene ved å bruke produktet og til en viss grad de teknologiske særegenhetene og funksjonaliteten.⁸² Disse kravene følger av saksområdet natur, og foreliggende rettspraksis på området mer generelt. Kravene til informasjon kan også anses som en forpliktelse til å sikre at forutsetningene for å bruke teknologien er til stede. For at bruken skal anses å være innenfor tillatt risiko må det antas å være et krav at resultatene er av minst samme kvalitet som det en fagperson kunne utført.⁸³ Dette følger av grunnleggende legislative hensyn på området og det nevnte skadeevneprinsippet.

Argumentasjonen illustreres i en avgjørelse fra tysk høyesterett som gjaldt informasjonsplikten ved nye behandlingsmetoder.⁸⁴ Saken gjaldt behandlingsmetoder som

⁷⁸ Se til dette Pehm, «Systeme der Unfallhaftung beim automatisierten Verkehr – Eine rechtvergleichende Analyse der Haftungsrisiken», på s. 262, som for selvkjørende biler nevner regelmessige verkstedbesøk og rengjøring av sensorer.

⁷⁹ Det forutsettes i forslaget til AI-ansvarsdirektivet som en betingelse for anvendelse av årsakspresumsjonen overfor profesjonelle brukere, se art. 4, nr. 3, bokstav a.

⁸⁰ Slik J. Pehm, «Systeme der Unfallhaftung beim automatisierten Verkehr – Eine rechtvergleichende Analyse der Haftungsrisiken», på s. 264.

⁸¹ Når det gjelder pasientskade bygger man i Norge på et objektivt ansvar basert på svikt – et avvik fra alminnelig god medisinsk praksis, jf. pasientskadeloven §2. Men ettersom kriteriet «svikt» i stor grad samsvarer med den objektive siden av culpanormen, og informasjonssvikt også er relevant etter ansvarsgrunnlaget som ledd i en behandlingssvikt, er argumentasjonslinjer som føres i land som bygger pasientskade på et culpaansvar i stor grad overførbare til norsk rett. Se om sviktkriteriet blant annet i A. Syse m.fl., *Pasientskaderett*, 2011 s. 106-108.

⁸² Slik E. Karner for østerriksk rett i «Liability for Medical Robots and Autonomous Medical Devices», på s. 60.

⁸³ Slik E. Karner for østerriksk rett i «Liability for Medical Robots and Autonomous Medical Devices», på s. 61.

⁸⁴ BGH 13.6.2006 i NJW 2006 s. 2477.

ennå ikke kunne anses som standard etter anerkjente internasjonale retningslinjer for behandling av den aktuelle diagnosen. Tysk høyesterett la til grunn at det forelå en skjerpet informasjonsplikt med hensyn til risikoer ved behandlingsmetoder som det gjennom forskning ennå ikke var tilstrekkelig kunnskap om.

Det medisinske området illustrer også godt de særlige spørsmål som oppstår når en profesjonell bruker baserer seg på beslutningsassistanse mer generelt. Brukeren må nemlig forventes å være godt informert om systemenes funksjonsmåte og begrensninger. Det vil også være nødvendig å kvalitetssikre informasjonen gjennom andre kilder, eller supplere informasjonen med andre metoder for kartlegging av faktum – for eksempel for å kunne stille en riktig diagnose eller for å velge riktig behandlingsmåte.

Dersom kunnskapsproduksjon ved hjelp av kunstig intelligens generelt blir pålitelig på bestemte områder, vil den teknologiske utviklingen også kunne føre til en økt forventning om å bruke slike systemer for å sikre et tilstrekkelig kunnskapsgrunnlag.⁸⁵

Visse krav til supplerende kunnskapsinnhenting og etterprøving/kvalitetssikring vil også kunne stilles til potensielle skadelidte, som benytter seg av ny teknologi. Dette gjelder tilsvarende dersom informasjonen brukes til private formål. Det er et generelt erstatningsrettslig prinsipp at kravene om å beskytte potensielle skadelidte mot risiko, reduseres dersom skadelidte forventes å beskytte seg selv.⁸⁶ Dersom det er alminnelig kjent at et informasjonsverktøy basert på kunstig intelligens har høy feilfrekvens, kan heller ikke skadelidte basere sine beslutninger på verktøyet uten ytterligere undersøkelser.

Skadelidte må også generelt ta høyde for upåregnelige hendelser ved bevisst eksponering for ny teknologi. Dette følger av generelle rettssetninger i erstatningsretten, og kommer blant annet til uttrykk i RIB-dommen. Høyesterett uttaler i avsnitt 36 at et avtalt formål med en risikofylt aktivitet kan ha betydning for ansvarsvurderingen. Synspunktet forutsetter altså at karakteren av risikoen er synbar for skadelidte. Den konkrete skademåten trenger imidlertid ikke å være påregnelig for skadelidte. Generell kunnskap om risiko har også betydning, så lenge skadevolderen overholder informasjonsplikten om teknologiens særegenheter og usikkerhet. Dommen understreker betydningen av en konkret vurdering på hvert enkelt område.

4.5 Justering av aktsomhetsstandarder som følge av menneskelige begrensninger?

Spørsmålet er her om kognitive og psykologiske begrensninger i menneskets møte med teknologiens funksjonsmåte gjør at man i enkelte tilfeller kan og bør stille mindre krav til reaksjonsevne og avverging av risikosituasjoner. Spørsmålet relaterer seg særlig til de

⁸⁵ Slik E. Karner, "Liability for Medical Robots and Autonomous Medical Devices", på s. 62. Slik også H. Eidenmüller, mer generelt for profesjonsansvaret ved bruk av kunstig intelligens med utgangspunkt i den kjente engelske avgjørelsen *Bolam v. Friern Hospital Management Committee* (1957) 1 WLR 582 (QB), «Machine Performance and Human Failure: How Shall We Regulate Autonomous Machines?» i H. Eidenmüller/G. Wagner, *Law by Algorithm*, 2021 s. 31.

⁸⁶ AM. Frøseth, *Skadelidtes egeneksponering for risiko i erstatningsretten*, s. 87-88.

situasjoner hvor det er forutsatt at en potensiell skadevolder/bruker skal overvåke driften og gripe inn.

Som tidligere nevnt, viser forskning at mennesker har enkelte kognitive utfordringer i samvirke med autonome systemer og kunstig intelligens. Brukerne har vanskeligere for å omstille seg, reagere adekvat og å gripe inn, enn når de styrer hele prosessen aktivt selv. Dette er blant annet et kjent problem knyttet til selvkjørende biler, som på dagens automatiseringsnivå forutsetter at en fører skal overvåke kjøringen og gripe inn dersom det oppstår feil i systemet.⁸⁷ En slik mulighet for overvåkning forutsettes også i EUs AIA i de tekniske krav til KI-systemer som fremgår av art. 14.

Forskning viser også at mennesker ofte reagerer uhensiktsmessig på enkelte av systemenes funksjonelle karakteristika. Som nevnt, forutsetter de fleste systemer i dag at det skjer en interaksjon med mennesker. Dette er både ansett nødvendig for å optimalisere systemets funksjoner, og for å gi varsler til brukeren ved feil. Forskning viser at brukere av automatiserte systemer opplever en «trettthet» knyttet til stadige varsler og signaler som tar oppmerksomheten fra andre ting.⁸⁸ Det er derfor en vanlig atferd at brukere ignorerer varslene eller skrur dem av, selv om de er påkrevd for forsvarlig drift.

Og sist, men ikke minst, kan mennesker ha vanskelig for å forstå komplekse systemer basert på kunstig intelligens. Forskning viser for det første at mennesker har en bias i relasjon til å stole for mye på automatiserte systemene, særlig over tid.⁸⁹ EUs AIA krever at KI-systemer tar høyde for og begrenser den uheldige effekten av denne types bias, se art. 14, nr. 4, bokstav b. Det er også påvist at brukerne har vanskelig for å forstå systemenes kapasitet og begrensninger. Dette gjelder uavhengig av teoretisk kunnskap, og kan særlig komme på spissen i stressede og akutte situasjoner.⁹⁰

Et spørsmål er om kompleksiteten i situasjonen i noen tilfeller kan gjøre manglende risiko-/skadeavverginge unnskyldelig. En høyesterettsdom som tilsynelatende åpner døren for resonnementet, er Avaldsnesdommen fra Rt. 1956 s. 656. Dommen handler om en skipper som i 1944 ble beordret av tyskerne til å ta passasjerer over et sund. Noen av passasjerene tok plass i styrehuset, der det befant seg en bølge med parafin. En av passasjerene kastet en røyk i bøtten, fordi han trodde den inneholdt vann. Dette førte til brann og en av passasjerene ble skadet. Spørsmålet var om skipperen burde fjernet bøtten eller sagt ifra om innholdet. Høyesterett kom til at det måtte anses unnskyldelig at skipperen ikke hadde fjernet risikokilden, fordi det var forståelig at han hadde oppmerksomheten sin om navigasjonen.

⁸⁷ Teknologirådet, *Selvkjørende biler – teknologien bak og veien fremover*, 2018 s. 48.

<https://teknologiradet.no/wp-content/uploads/sites/105/2018/11/Selvkjorende-biler-teknologien-bak-og-veien-fremover.pdf>

⁸⁸ A.D. Selbst, «Negligence and AI's Human User», 100 *B.U.L.Rev.* 2020 s. 1315 flg. på s. 1348, med videre henvisninger.

⁸⁹ I. Wolf, «The Interaction Between Humans and Autonomous Agents» i M. Mauer/J.C. Gerdes/B. Lenz/H. Winner, *Autonomous Driving*, 2019 s. 104-105.

⁹⁰ Se et eksempel på dette og automatiserte systemer i fly i T. Solvang, «Man, machine, and culpa: Or finding a path toward strict liability» i H. Ringbom/E. Røsæg/T. Solvang (eds.) *Autonomous Ships and the Law 2021* s. 98 flg. på s. 99. Se ellers blant annet, Law Commission of England and Wales/Scottish Law Commission, *Automated Vehicles: Joint report*, 26. januar 2022 s. 121.

Skipperen hadde ansvaret for en båt som ikke vanligvis fraktet passasjerer, og han måtte manøvrere i et vanskelig, ukjent farvann.

Dommen baserer seg på en konkret vurdering, og det er usikkert i hvilken grad man kan innfortolke generelle retningslinjer. Dette blant annet fordi skipperen i utgangspunktet var beordret inn i situasjonen av tyskerne, og ikke selv hadde noen egeninteresse i frakten. Så lenge en bruker selv har herredømme over når og hvor KI-systemet blir satt i drift, må det, basert på analysen ovenfor, kunne forventes nokså høye krav til brukeren med hensyn til hans/hennes forutsetninger for å kunne styre bruken på det aktuelle bruksområdet. For enkelte typer teknologi og bruksområder kan det også være aktuelt å kreve opplæring, slik tilfellet er for piloter. Et krav om opplæring vil først og fremst være praktikabelt i mindre skala for profesjonelle brukere. Kravene må imidlertid antas å øke i fremtiden, også i lys av EUs AIA's vektlegging av opplæring ved bruk av KI-systemer, en plikt som etter forordningen påligger tilbydere og profesjonelle brukere, inkludert arbeidsgivere, se art. 4.

Interessant i dette perspektivet er den nye tyske særreguleringen av førerens ansvar ved kjøring av biler på automatiseringsnivå 4 og 5. Loven gir her eksplisitt uttrykk for at føreren lovlig kan ha oppmerksomheten sin på andre ting enn kjøringen, med mindre det automatiserte systemet gir beskjed om at føreren må ta over. Føreren har imidlertid en plikt til å gripe inn ved varsel fra systemet. De spesifikke pliktene suppleres også av en generell aktsomhetsnorm som angir at føreren uansett blir ansvarlig dersom hen burde forstå at forutsetningene for å bruke kjøretøyet ikke var til stede.⁹¹ I den tyske Strassenverkehrsgesetz stilles det imidlertid krav om at systemet selv kan slå seg av.⁹² Slike krav vil naturlig nok også påvirke hva som forventes av føreren.

Pliktene etter den tyske lovgivningen må ses i lys av at man i Tyskland allerede for konvensjonell bilteknologi praktiserer et skyldansvar med en ansvarspresumsjon for føreren.⁹³ I tillegg praktiseres det i tysk rett en ansvarspresumsjon etter culperegelen og/eller en regel om snudd bevisbyrde for årsakssammenheng der en påstått skadevolder har brutt en sikkerhetsregel.⁹⁴ I lys av dette kan eksemplene antakelig ikke tjene som eksempler på hvilken type handleplikter og hvilken standard som generelt bør stilles for autonome høyrisikosystemer i Europa.

Kunnskapen man har om menneskelige begrensninger kan heller enn å lempe på aktsomhetsstandarden tale for å skjerpe aktsomhetsplikten knyttet til *forutsetningene* for å sette systemet i drift. De menneskelige begrensningene tilsier også en streng plikt til å avstå fra å gripe inn i forutsetningene for systemets drift, noe som underbygger et skjerpet aktsomhetskrav knyttet til å sikre kontinuerlige oppdateringer av produktet.

⁹¹ Se om innholdet i den tyske Strassenverkehrsgesetz (StVG) på dette punktet i A.M. Frøseth, «Ansvar for skade forårsaket av kunstig intelligens i selvkjørende kjøretøy – en komparativ oversikt», på s. 585.

⁹² L.c.

⁹³ L.c.

⁹⁴ E. Karner/B. Koch, "Civil Liability for Artificial Intelligence – A Comparative Overview of Current Tort Laws in Europe" i M. A. Geistfeld/E. Karner/B. Koch/C. Wendehorst (eds.), *Civil Liability for Artificial Intelligence and Software*, 2023 s. 49.

I møte med en stadig større grad av automatisering, vil skjerpede aktsomhetskrav likevel ha sin begrensning. Dersom det foreligger strenge og detaljerte offentligrettslige krav til produktets sikkerhet og design, vil ansvaret også i praksis kanaliseres mot produsenten. Og dersom det er lite brukeren faktisk kan gjøre for å påvirke sikkerheten i systemet, vil idealkrav til aktsomhet i praksis innebære et objektivt ansvar.

En slik ansvarsform er ikke i tråd med alminnelige prinsipper for ansvar i norsk rett. Ei heller er det en ansvarsform som ivaretar de grunnleggende begrunnelsene for ansvar som ble diskutert i punkt 3. Det er derfor naturlig å spørre om ansvarsgrunnlaget for brukere av *fullt* automatiserte systemer i fremtiden heller bør være et objektivt ansvar knyttet til ulike bruksområder.

5. Avslutning

Et hovedfunn i arbeidet er at culpanormen stort sett vil være fleksibel nok til å fange opp nye skadetilfeller og fylle sin funksjon på dette området. Denne vurderingen bygger på dagens mest brukte teknologi, og den produktsikkerhetsreguleringen som forventes å gjelde i fremtiden. Konklusjonen må ses i sammenheng med andre ansvarsgrunnlag som kan supplere den alminnelige culpanormen. I så måte fanger det nye produktansvarsdirektivet opp mange praktiske skadetilfeller.⁹⁵ I tillegg kommer potensialet som ligger i det ulovfestede objektive ansvaret. Kravet til bevis for årsakssammenheng kan fremdeles utgjøre betydelige hindringer, men dette er ikke enestående for denne typen saker.

For noen skadetilfeller vil det oppstå store utfordringer knyttet til å påvise den årsaksmessige forbindelsen mellom en skadevolder og feil i KI-systemet. Dette kommer særlig på spissen der mange aktører har innflytelse risikonivået systemet. I slike saker ligger gjerne flere av årsaksspørsmålene i grenselandet mot vurderingene under ansvarsgrunnlagene. Her vil nye ansvarsmodeller med klare elementer av objektivt ansvar, eventuelt hybridvarianter basert på ansvars- eller årsakspresumsjoner, kunne fylle de mest åpenbare lakuner der skaden forårsaker økonomisk tap.

⁹⁵ Teksten vedtatt av Europaparlamentet 12. mars 2024 venter for tiden på formell godkjenning i Rådet. P9_TA(2024)0132. European Parliament legislative resolution of 12 March 2024 on the proposal for a directive of the European Parliament and of the Council on liability for defective products (COM(2022)0495 – C9-0322/2022 – 2022/0302(COD))