



JURIDISKA FAKULTETEN
vid Lunds universitet

Danny Dib

Skadeståndsrätt i den digitala eran - Ansvarsproblematiken för artificiell intelligens

LAGF03 Rättsvetenskaplig uppsats

Kandidatuppsats på juristprogrammet

15 högskolepoäng

Handledare: Olena Bokareva

Termin: VT23

Innehåll

SUMMARY	1
SAMMANFATTNING	2
FÖRKORTNINGAR	3
1 INTRODUKTION	4
1.1 Bakgrund	4
1.2 Syfte och frågeställningar	6
1.3 Avgränsningar	7
1.4 Metod och material	8
1.4.1 Rättsdogmatisk metod och analys.....	8
1.4.2 Material.....	9
1.5 Forskningsläge	9
1.6 Disposition	10
2 ARTIFICIELL INTELLIGENS – TEKNISK ÖVERSIKT OCH RÄTTSLIGA PROBLEM	11
2.1 Artificiell intelligens – en inblick	11
2.2 AI-systemens tekniska funktionssätt och inneboende egenskaper	11
3 FÖRSLAG TILL DIREKTIV OM SKADESTÅNDSANSVAR GÄLLANDE AI OCH NYTT PRODUKTANSVARSDIREKTIV	14
3.1 Direktivens syfte och bakgrund	14
3.2 Nya direktiv – en jämförelse	15
3.3 Direktivens tillämpningsområde	17
3.3.1 AILD-förslaget.....	17
3.3.2 PLD-förslaget.....	19
3.4 Utlösning av ansvarsmekanismen	20
3.5 Nya editionsregler	22
3.5.1 Artikel 3 AILD-förslaget	22
3.5.2 Artikel 8 PLD-förslaget	22
3.5.3 Fördelar, nackdelar och inverkan på ansvarsproblematiken.....	23

3.6	Placering av bevisbördan.....	25
3.6.1	Artikel 4 AILD-förslaget	25
3.6.2	Artikel 9 PLD-förslaget	26
3.6.3	Fördelar, nackdelar och inverkan på ansvarsproblematiken.....	27
4	SLUTSATS OCH REFLEKTIONER.....	30

Summary

In recent years, the rapid development of AI systems has elevated the technology to new heights. The emergence of autonomous vehicles and advanced language models like ChatCPT has brought the discussion around the legal issues of AI systems to the forefront.

The primary aim of tort law is to compensate the injured parties and impose liability on the parties responsible. Due to the complex nature of AI systems, it has become increasingly difficult to identify the liable party. To address this, the European Union has proposed two new directives that take aim at the liability issues caused by AI systems.

Consequently, this paper first aims to examine why it's difficult to identify a liable party when damages are caused by AI systems. Secondly, the objective is to investigate whether the two directives adequately solve these liability issues.

The paper concludes that due to the complexity of how AI systems function, the resulting outputs cannot easily be traced back to a specific internal process. This so-called "black box problem" results in a lack of transparency and unpredictability in these systems. These inherent characteristics can, in combination with the fact that AI systems often incorporate components from different suppliers, complicate the identification of the liable party when damages occur.

The two directives proposed regulations appear to be a step in the right direction to resolve these issues. However, there are concerns about whether these regulations, in practice, achieve a sufficiently balanced distribution of risks and costs.

Sammanfattning

AI-systemens snabba utveckling har på senare år fört teknologin till nya höjder. Framväxten av autonoma fordon och avancerade språkmodeller som ChatGPT har ställt diskussionen kring AI-systemens rättsliga problematik på sin spets.

Bland annat har skadeståndsrättens reparativa syfte blivit svårare att uppnå till följd av att skador orsakade av AI-systemen komplicerar preciseringen av ansvarande part. För att bemöta detta har EU föreslagit två nya direktiv som behandlar skadeståndsansvar i relation till AI-system.

Syftet med denna uppsats är att i första hand undersöka hur skador orsakade av AI-system kan resultera i en ansvarsproblematik, och i andra hand utreda huruvida direktiven kan betraktas som adekvata lösningar på den rättsliga problematiken.

Slutsatsen därav är att komplexiteten av AI-systemens tekniska funktionssätt kan leda till utdata som inte enkelt kan härledas till en specifik intern process. Detta så kallade ”*black box-problem*” leder till att dessa system karaktäriseras av brist på transparens och oförutsägbarhet. Eftersom AI-system ofta bygger på beståndsdelar från olika leverantörer, kan detta tillsammans med systemens inneboende egenskaper komplicera preciseringen av ansvarande part när skador inträffat.

Direktivens föreslagna regleringar framstår vara ett steg i rätt riktning till att lösa ansvarsproblematiken, men det ifrågasätts huruvida regleringarna i praktiken uppnår en tillräckligt balanserad risk- och kostnadsfördelning.

Förkortningar

AILD-förslaget

Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar gällande artificiell intelligens

PLD-förslaget

Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar för produkter med säkerhetsbrister

AI-förordningen

Förslag till Europaparlamentets och rådets förordning om harmoniserade regler för artificiell intelligens

Begreppslista

AI-system

Mjuk- eller hårdvara med inkorporerad artificiell intelligens

Utdata

Det resultat som ett AI-system genererar till följd av sin interna beslutsprocess

Indata

Den information som tillförs ett AI-system

Bias

Systematiska fel i insamlandet och tolkningen av data

Algoritm

En mängd instruktioner som efter de genomförs löser ett givet problem

Språkmodell

AI-system som efterliknar mänsklig kommunikation

1 Introduktion

1.1 Bakgrund

Artificiell intelligens, så kallad ”AI”, har länge varit populärt inom såväl akademien som fiktion. Den snabba utvecklingen de senaste åren har förvandlat delar av det som tidigare enbart var verklighet på filmduken, till något som var och annan person har tillgång till i sin mobiltelefon. Nu har framkomsten av språkmodeller som Microsofts ChatGPT väckt nya frågor om AI-systemens plats i våra liv och följaktligen även i vårt rättssystem.¹

Skadeståndsrätten ämnar försätta en person som lidit skada i samma position som om skadan aldrig skett, men till följd av AI-systemens funktionssätt har detta reparativa syfte blivit svårare att uppnå.² För att en skyldighet att kompensera någons förlust ska uppkomma krävs bland annat att dess orsak ligger inom den skadevållandes sfär.³ Att identifiera en sådan orsak och övertyga domstolen om dess påverkan på händelseförloppet blir försvårat när det är ett AI-system som misstänks ha orsakat skadan, till stor del på grund av systemens komplexitet, brist på transparens och förutsägbarhet. AI-systemens funktionssätt utforskas djupare i avsnitt 2.

Ponera att ett autonomt fordon, såsom en självkörande personbil, krockar med ett annat fordon. Orsaken bakom kollisionen kan vara såväl hårdvaru- som mjukvaru-relaterad. Bromsarna eller sensorerna kan ha slutat fungera, men det kan även ha varit fel i mjukvaran. Felet kan i sin tur ha uppstått av en rad olika anledningar: en miss i programmeringen (s.k. ”bugg”), en uppdatering av mjukvaran som gick snett, eller en angripare som lyckats manipulera mjuk-

¹ Metz, Cade, Schmidt, Gregory, “*Elon Musk and Others Call for Pause on A.I., Citing ‘Profound Risks to Society’*”, The New York Times, 29/3 2023, Tillgänglig: <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html> [Hämtad: 09-05-2023].

² Hellner, Jan, Radetzki, Marcus, “*Skadeståndsrätten*”, Norstedts Juridik (2021), s. 37–38.

³ Hellner, Radetzki (2021), s. 189–193.

varan. Oavsett vad som orsakade kollisionen, så tillför den adderade komplexiteten en rad svårigheter med att fastställa skadevällare och orsakssamband.⁴

Denna problematik belyses i ett fall från 2018 där Elaine Herzberg blir påkörd av ett självkörande fordon och omkommer. Herzberg som på natten korsade två vägbanor med sin cykel träffades av en autonom bil som testkördes av Uber.⁵

I sin rapport framställde det amerikanska trafikverket att kollisionen orsakades av fel i sensorerna som inte klassificerade Herzberg som en fotgängare, dock specificerades det aldrig vad som orsakat felet i sig.⁶ Eftersom fordonets AI-system bygger på både mjuk- och hårdvara kan det vara svårt att sätta fingret på exakt var felet inträffade.

I förevarande fall hölls Uber inte själva ansvariga för dödsfallet i och med att delstatens lagstiftning krävde att en säkerhetsförare skulle vara närvarande och kunna ingripa. Däremot blev Rafaela Vasquez, personen som satt bakom ratten, ansvarig för dödsfallet; men i en framtid där autonoma fordon existerar utan säkerhetsförare belyser detta fall denna potentiella ansvarsproblematiken.⁷

Inom EU har det länge förts diskussioner kring den rättsliga problematik som AI-system medför och ett av Ursula von der Leyens främsta mål som ordförande för EU-kommissionen har varit att genom lagstiftning utveckla en samordnad europeisk plan för artificiell intelligens.⁸

⁴ Karner, Ernst, Koch, Bernhard A., Geistfeld, Mark A., "Comparative Law Study on Civil Liability for Artificial Intelligence", Rapport till Europeiska kommissionen (2021), s. 9 Tillgänglig: <https://op.europa.eu/en/publication-detail/-/publication/8a32ccc3-0f83-11ec-9151-01aa75ed71a1> [Hämtad: 06-05-2023].

⁵ Levin, Sam, "Safety driver charged in 2018 incident where self-driving Uber car killed a woman", The Guardian, 16/9 2020, Tillgänglig: <https://www.theguardian.com/us-news/2020/sep/16/uber-self-driving-car-death-safety-driver-charged> [Hämtad 06-05-2023].

⁶ Levin, The Guardian 16/9 2020.

⁷ Levin, The Guardian 16/9 2020.

⁸ von der Leyen, Ursula, "A Union that strives for more: My agenda for Europe", 2019, Tillgänglig: https://eunec.eu/sites/www.eunec.eu/files/attachment/files/political-guidelines-next-commission_en_kopie.pdf [Hämtad 06-05-2023].

Ett förslag till förordning publicerades i april 2021 som tar sikte på att förebygga säkerhetsriskerna med AI-system genom att bland annat införa förbud mot vissa typer av AI-system samt skapa nya tillsynsmyndigheter i medlemsstaterna.⁹

Detta följdes av ett förslag till direktiv om skadeståndsansvar gällande artificiell intelligens och ett nytt produktansvarsdirektiv i september 2022.¹⁰ Direktiven avser istället bemöta den ovannämnda ansvarsproblematiken och föreslår ett antal lagändringar för detta.¹¹

Hädanefter kommer direktiven benämnas enligt sina engelska förkortningar i syfte att förenkla läsningen: AILD-förslaget och PLD-förslaget.

1.2 Syfte och frågeställningar

Syftet med denna uppsats är i första hand att utforska hur introduktionen av AI-system komplicerar preciseringen av ansvar när skada inträffar i utomobligatoriska förhållanden, för att sedan analysera hur de föreslagna EU-direktiven bemöter denna ansvarsproblematik.

Genom att beskriva AI-systemens tekniska funktionssätt vill jag belysa hur dess inneboende egenskaper kan resultera i svårigheter att fastställa en ansvarande part när skador inträffar i utomobligatoriska förhållanden.

Detta ger en bakgrundsbild som det sedan kan återkopplas till vid analysen av dem två EU-direktivens föreslagna lösningar till den rättsliga ansvarsproblematiken.

⁹ Förslag till Europaparlamentets och rådets förordning om harmoniserade regler för artificiell intelligens (AI Act) COM(2021) 206 final, Art. 5 och 30.

¹⁰ Europeiska kommissionen, ”*New liability rules on products and AI to protect consumers and foster innovation*”, 2022, Tillgänglig: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807 [Hämtad: 06-05-2023].

¹¹ Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar gällande artificiell intelligens (AILD) COM(2022) 496 final, s. 2-3.

Analysen av direktivens föreslagna lagändringar ger en inblick i den potentiella lösning som, givet att direktiven antas, kommer påverka svensk rätt. Det kommer även föras en diskussion kring huruvida dessa kan anses utgöra en adekvat lösning på det rättsliga problemet.

Min avsikt är därmed att besvara följande frågeställningar:

1. På vilket sätt kompliceras preciseringen av ansvarande part när AI-system orsakat skada?
2. Hur bemöter de föreslagna EU-direktiven denna ansvarsproblematik och kan de betraktas som adekvata lösningar till det rättsliga problemet?

1.3 Avgränsningar

I syfte att hålla mig inom uppsatsens ramar kommer inte AI-systemens tekniska funktionssätt beskrivas i större detalj än vad som krävs, för att förstå ansvarsproblematiken när skador orsakas till följd av dessa system.

Eftersom det främst är direktivens lagändringar som diskuteras i denna uppsats kommer AI-förordningen inte heller behandlas i större utsträckning än det som behövs för förståelsen av regleringarna.

I analysen av direktivens lagförslag finns det dessutom aspekter som jag tyvärr inte kan inkludera på grund av uppsatsens gränser. Det hade bland annat varit intressant att djupare utforska vad ett defekt, i relation till utdata som AI-system producerat, faktiskt innebär — bör defekt bedömas utifrån ett mänskligt perspektiv (i.e. hade en människa kunnat göra samma fel) eller ett tekniskt perspektiv (i.e. till vilken grad uppvisar produkter med detta AI-system samma fel)?

Av samma skäl gjordes även ett urval av direktivens viktigaste bestämmelser i relation till ansvarsproblematiken.

1.4 Metod och material

Uppsatsen kommer först utreda hur preciseringen av ansvarande part kompliceras när skada orsakat av AI-system inträffar i utomobligatoriska förhållanden. Därefter analyseras de två förslagen till EU-direktiv som ämnar bemöta denna ansvarsproblematik.

Jag utgår från ett internationellt rättsutvecklingsperspektiv av den orsaken att uppsatsen inte diskuterar direktivens inverkan på någon specifik nationell lagstiftning, utan fokuserar på hur rätten kan utvecklas för att bemöta ansvarsproblematiken. I analysen är utgångspunkten även ett kritiskt perspektiv. Följaktligen har jag använt mig av rättsdogmatisk metod och analys till att besvara mina frågeställningar.

1.4.1 Rättsdogmatisk metod och analys

Åsikterna kring tillämpningen av rättsdogmatisk metod och analys skiljer sig, men jag har i denna uppsats utgått från det som Jan Kleineman skrivit i boken *Juridisk metodlära*.¹²

Rättsdogmatisk metod innebär att gällande rätt undersöks med målet att fastslå vad som gäller i förevarande fall, eller hitta lösningen på ett rättsligt problem.¹³ Denna uppsats utgår visserligen inte från gällande rätt, utan från förslag till EU-direktiv som i senare led kan bli gällande rätt. Eftersom jag i min beskrivning och analys av lagförslagets regleringar utgått från perspektivet att det är gällande rätt, har jag valt att ändå applicera metoden.

I den analys som följer en rättsdogmatisk utredning, särskiljer man på *de lege lata*-argumentation och *de lege ferenda*-argumentation. Den först nämnda är en utredning av hur rättsläget ser ut medan den andra föreslår lösningar på ännu olösta problem.¹⁴

¹² Kleineman (2018), s. 21–41.

¹³ Ibid. s. 21.

¹⁴ Ibid. s. 36–37.

I min analys utgår jag främst från *de lege lata*-argumentation men det förekommer även *de lege ferenda*-argumentation, exempelvis i avsnitt 3.3.1 där jag föreslår ett alternativt synsätt på definitionen av ett AI-system.

1.4.2 Material

Till att besvara min första frågeställning har jag främst använt mig av Jacob Turners bok *Robot Rules: Regulating Artificial Intelligence* (2018). Trots att det kan finnas källor som mer noggrant beskriver hur ett AI-system fungerar, bedömer jag det inte nödvändigt för att förstå varför sådana system kan frambringa en ansvarsproblematik. Turner har dessutom sammanställt denna komplexa information på ett lättförståeligt sätt som ger jurister en tillräcklig inblick i AI-systemens funktionssätt.

Uppsatsen kretsar kring de föreslagna lagändringarna som AILD-förslaget och PLD-förslaget framför, därmed har jag till stor del använt mig av deras angivna skäl och förarbeten. Eftersom AI-förordningens bestämmelser är centrala för många av AILD-förslagets regleringar har även denna brukats, men endast i den mån som krävs.

Slutligen har jag nyttjat ett antal nyhetsartiklar och expertåsikter i syfte att ge exempel och nyansera min analys av direktivens nya regleringar. Expertåsikter jag har använt mig av kommer bland annat från Francesco Paolo Patti¹⁵, Gerhard Wagner¹⁶ och Matt O'Shaughnessy¹⁷.

1.5 Forskningsläge

De två föreslagna EU-direktiven och deras effekt har bland annat tidigare studerats av Samar Abbas Nawas och Philipp Hacker.

¹⁵ Docent i civilrätt vid Università Commerciale Luigi Bocconi.

¹⁶ Professor i civilrätt, affärsjuridik och ekonomi vid Humboldt-Universität zu Berlin och gästprofessor i juridik vid University of Chicago Law School.

¹⁷ Gästforskare i teknologi och internationell politik vid Carnegie Endowment for International Peace.

I sin artikel ”*The Proposed EU AI Liability Rules: Ease or Burden?*” diskuterar Nawas, doktorand vid Peace Research Institute Oslo, effekten av de föreslagna regleringarna på ett potentiellt skadeståndsanspråk. Han kommer fram till att reglerna framförallt gör det lättare för skadelidande att få tillgång till information, men tillför samtidigt bördor som kan vara svåra att överkomma.

Philipp Hacker, professor vid Europa-Universität Viadrina, diskuterar även förslagen i ”*The European AI Liability Directives—Critique of a Half-Hearted Approach and Lessons for the Future*”. Han kritiserar förslagen för deras halvhjärtade regleringsförsök och föreslår därför ett antal ändringar.

Ansvarsproblematiken har diskuterats av Jacob Turner, i sin bok ”*Robot Rules: Regulating Artificial Intelligence*”, och av Wagner i sin uppsats ”*Robot Liability*”.

Andra aspekter av ansvarsproblematiken, såsom dess påverkan på autonoma fordon, har tagits upp av Francesco Paolo Patti och Agnes Juhasz i deras respektive artiklar: ”*The European Road to Autonomous Vehicles*” samt ”*The Legal Framework of Autonomous Driving in Germany*”.

1.6 Disposition

I uppsatsens andra kapitel beskriver jag AI-systemens tekniska funktionssätt, för att belysa hur dess karaktärsdrag kan resultera i svårigheter att precisera en ansvarande part när skador inträffar. På så vis beskriver jag den rättsliga ansvarsproblematiken och besvarar min första frågeställning.

Kapitel 3 kartlägger de två direktivens föreslagna lagändringar och diskuterar dessa utifrån ansvarsproblematiken, samt dess innebörd för skadelidande.

Avslutningsvis reflekterar jag i slutsatsen över direktivens regleringar och till vilken grad de är att anse som adekvata lösningar till det rättsliga problemet.

2 Artificiell intelligens – Teknisk översikt och rättsliga problem

2.1 Artificiell intelligens – en inblick

AI har sedan 1950-talet, när Alan Turing publicerade sin banbrytande artikel ”Computing Machinery and Intelligence”, utvecklats till nya höjder. Hans revolutionerande idéer om maskiner som skulle kunna demonstrera ett intelligent beteende likvärdigt – eller nästintill omöjligt att särskilja från – människor, banade vägen för ytterligare forskning och är än idag grundläggande koncept inom området.¹⁸

Under de följande årtionden utvecklades maskininlärningsmetoder, vilket möjliggjorde för nya framsteg som förde AI-system in i vår vardag. Idag används AI-system i allt från självkörande bilar och mobila röstassistenter till medicinsk diagnostik och vetenskaplig forskning.¹⁹

Det nästkommande avsnittet syftar till att ge en inblick i AI-teknikens funktionssätt samt förklara de inneboende egenskaper som medverkar till den rättsliga ansvarsproblematik uppsatsen ämnar utforska.

2.2 AI-systemens tekniska funktionssätt och inneboende egenskaper

I den föreslagna AI-förordningen definieras ett AI-system bland annat som mjukvara vars utveckling skett genom en eller flera tekniker, däribland maskininläring.²⁰ Definitionen lägger viss vikt vid utvecklingstekniken vilket

¹⁸ Turing, Alan, ”*Computing Machinery and Intelligence*”, Oxford University Press, *Mind*, 59:236 (1950), s. 433-460, Tillgänglig: <https://doi.org/10.1093/mind/LIX.236.433> [Hämtad: 08-05-2023].

¹⁹ Turner, Jacob, ”*Robot Rules: Regulating Artificial Intelligence*”, Springer International Publishing AG (2018), s. 23-28.

²⁰ Förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 3(1) och bilaga 1.

även blir användbart för att förstå ansvarsproblematiken när AI-system orsakar skador.²¹

Maskininlärning, innebär att AI-system ges förmågan att på egen hand vidareutveckla sin prestanda genom att ta lärdom från tidigare erfarenheter. Det krävs med andra ord inte att systemets nya förmågor och prestanda ska programmeras av en människa.²²

Denna förmåga delas i huvudsak in i två kategorier: väglett och icke-väglett lärande. Väglett lärande innebär att AI-systemet ges träningsdata som innehåller facit för rätt och fel svar. Denna data används sedan av systemet för att upprätta hypoteser kring vilka parametrar som utgör rätt och fel, i syfte att i framtiden kunna dra slutsatser och kategorisera okänd data. På så vis kan AI-systemet utveckla sina förmågor och prestanda utan mänskligt ingripande.²³

Icke-väglett lärande skiljer sig genom att algoritmen istället tränas på data som saknar facit. Genom denna metod särskiljer AI-systemet informationen baserat på deras gemensamma karaktärsdrag och grupperar in det i kluster. Av särskilt intresse för denna metod är att systemet på egen hand kan känna igen mönster i träningsdata. Ett av de första försöken på att sätta denna metod i bruk var ett AI-system som tränades till att särskilja katter från andra djur i videor publicerade på Youtube.²⁴

Applikationsmöjligheterna för ett AI-system som på egen hand kan känna igen mönster utifrån träningsdata är givetvis inte begränsade till katter utan denna teknologi kan användas i en mängd andra sammanhang där den rättsliga ansvarsproblematiken får ett större genomslag. Som ett exempel ställdes problematiken på sin spets när Amazon fick skrota sitt rekryteringsverktyg eftersom det visade sig att det diskriminerade mot kvinnor. AI-systemet som tränats på CV inskickade till företaget under en tioårsperiod sållade bland de

²¹ Ibid. s. 12.

²² Turner (2018), s. 70-71.

²³ Turner (2018), s. 71-72.

²⁴ Turner (2018), s. 71-72.

jobbsökande genom att känna igen olika mönster. Av den orsaken att teknologibranschen till stor del domineras av män överfördes detta bias till AI-systemet, vilket varken var deras intention eller ett önskvärt beteende.²⁵

Komplexiteten av AI-systemens tekniska funktionssätt samt dess kapacitet att på egen hand vidareutveckla sina förmågor och prestanda kan därmed leda till utdata som blir svårförklarlig. Eftersom det saknas insyn i de konstant förändrande systemens beslutsprocess kan dessa generera resultat som utvecklarna själva inte nödvändigtvis kan förklara eller koppla till en specifik intern process. Detta benämns som ”*the black box problem*” med anledning av dess bristande transparens.²⁶

Utöver dessa inneboende egenskaper är det inte ovanligt att ett AI-system bygger på beståndsdelar från olika leverantörer. Data som systemet är tränat på kan komma från en part medan algoritmerna kommer från en annan. Är AI-systemet exempelvis en del av ett självkörande fordon kan även hårdvaran som det samspelar med komma från en tredje part.²⁷

Följaktligen blir det problematiskt när komplexa AI-system med inverkan på människors liv tar oförutsägbara beslut som kan orsaka skada. Samspelet mellan AI-systemets olika beståndsdelar samt dess inneboende egenskaper medför att bestämmelserna i gällande skadeståndsrätt försvårar eller rent av omöjliggör för skadelidande att identifiera inte bara den skadevällande parten, men även att uppfylla rekvisiten för en bifallande dom.²⁸

²⁵ Dastin, Jeffrey, “*Amazon scraps secret AI recruiting tool that showed bias against women*”, Reuters 11/10 2018, Tillgänglig: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> [Hämtad: 11-05-2023].

²⁶ Turner (2018), s. 325.

²⁷ Europeiska kommissionen, “*Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*”, 2020, s. 2, Tillgänglig: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064&from=en> [Hämtad: 11-05-2023].

²⁸ Förslag till EU-direktiv (AILD) COM(2022) 496 final, s. 1.

3 Förslag till direktiv om skadeståndsansvar gällande AI och nytt produktansvarsdirektiv

3.1 Direktivens syfte och bakgrund

I en studie utförd av EU-kommissionen, med avsikt att kartlägga användandet av AI-system bland europeiska företag, rapporterade 43% av de som ämnade införa sådana system till deras verksamhet att den rättsliga ansvarsproblematiken var det största hindret. Detta var också den tredje största barriären överlag till varför europeiska företag inte implementerat AI-system i deras verksamheter.²⁹

Målet med införandet av ett direktiv om skadeståndsansvar gällande AI har därmed varit att bemöta denna ansvarsproblematik i syfte att möjliggöra en utökad implementering bland europeiska verksamheter.³⁰

Det gällande produktansvarsdirektivet som infördes 1985 gick, i samband med EU-kommissionens lagstiftningsutredningar kring AI, igenom ett antal utvärderingar med avsikt att klargöra dess tillämplighet vid skador orsakade av AI-system. Utvärderingen fann att lagstiftningen i sin helhet fungerat väl men pekade ut två väsentliga brister: direktivets omoderna definition av en ”produkt” och att dess bevisbörderegler försvårade för skadelidande att söka kompensation i komplexa fall där skadan orsakats av AI-system. Följaktligen introducerades ett förslag till nytt produktansvarsdirektiv i slutet av 2022.³¹

²⁹ Europeiska kommissionen, “*European enterprise survey on the use of technologies based on artificial intelligence – Final report*”, 2020, s. 59, Tillgänglig: <https://data.europa.eu/doi/10.2759/759368> [Hämtad 13-05-2023].

³⁰ Förslag till EU-direktiv (AILD) COM(2022) 496 final, s. 1-2.

³¹ Förslag till EU-direktiv (PLD) COM(2022) 495 final, s. 1-2.

Kommande avsnitt avser därmed redogöra för de två direktivens tillämpningsområde och relevanta lagförslag, samt diskutera deras innebörd i relation till ansvarsproblematiken.

3.2 Nya direktiv – en jämförelse

För att ge en helhetsbild av de föreslagna lagändringarna är det av betydelse att förklara hur direktiven förhåller sig till, samt särskiljer sig från, varandra.

Direktiven är parallellt tillämpliga men täcker olika aspekter, såsom typer av skadevållande och ansvarsgrunder. Skadelidande kan därmed välja det ramverk de vill grunda sitt skadeståndsanspråk på; vilket blir aktuellt eftersom ramverken skiljer sig i typen av skador som omfattas, vem anspråket kan riktas mot, samt tillämplig ansvarsform.³²

PLD-förslaget är tillämpligt på fysiska varor såväl som mjukvara, vilket inkluderar AI-system. Förslaget ämnar introducera en harmoniserad lagstiftning på EU-nivå som medför strikt ansvar för leverantörer av AI-system.³³

AILD-förslaget särskiljer sig här genom att enbart tillämpas på AI-system. Förslaget avser uppnå bättre förutsättningar för att bevisa culpaansvar genom att harmonisera processrättsliga aspekter av medlemsstaternas existerande lagstiftning. Regleringarna kan dessutom appliceras på leverantörer så väl som användare av AI-system.³⁴

De två förslagen skiljer sig även i graden av harmonisering. I enlighet med artikel 3 föreskriver PLD-förslaget fullständig harmonisering³⁵, vilket är en kodifiering av tidigare rättspraxis från EU-domstolen gällande det rådande

³² Europeiska kommissionen, “*Questions and answers on the revision of the Product Liability Directive*”, 2022, s. 2 punkt 9, Tillgänglig: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/qanda_22_5791/QANDA_22_5791_EN.pdf [Hämtad: 13-05-2023].

³³ Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar för produkter med säkerhetsbrister (PLD) COM(2022) 495 final, Art. 1, 3, 4(1), 7.

³⁴ Förslag till EU-direktiv (AILD) COM(2022) 496 final, Art. 1(1), 2(3), 2(4), 3(1).

³⁵ Förslag till EU-direktiv (PLD) COM(2022) 495 final, Art. 3.

produktansvarsdirektivet.³⁶ Domstolen fastställde här i flertalet rättsfall att direktivet skulle tolkas som maximalt harmoniserande inom dess tillämpningsområde.³⁷

I kontrast fastställer AILD-förslaget endast minimal harmonisering, vilket innebär att medlemsstaterna kan implementera mer fördelaktiga regleringar än vad förslaget föreskriver.³⁸

Följaktligen är det möjligt att medlemsstater helt vänder på bevisbördan när det kommer till ansvar baserat på culpa, eller rent av inför strikt ansvar för olika typer av skador orsakade av AI-system.³⁹

Strikt ansvar för skador orsakade av AI-system förväntas redan implementeras i ett antal medlemsstater. Tyskland har exempelvis infört strikt ansvar till förare av autonoma fordon⁴⁰ och införandet av liknande regler är förväntat i länder som använder sig av motsvarande bestämmelser om strikt ansvar i deras trafikregelverk, såsom Frankrike och Italien.⁴¹

Minimiharmoniseringen som AILD-förslaget föreskriver riskerar därmed fragmentera hanteringen av ansvarsproblematiken bland medlemsstaterna, vilket ironiskt nog även kan hindra upptaget av AI-system bland europeiska företag.

³⁶ Se följande: Domstolens dom den 25 april 2002, *Europeiska kommissionen mot Republiken Grekland*, C-154/00, ECLI:EU:C:2002:254; Domstolens dom den 25 april 2002, *María Victoria González Sánchez mot Medicina Asturiana SA.*, C-183/00, ECLI:EU:C:2002:255; Domstolens dom den 25 april 2002, *Europeiska kommissionen mot Franska republiken*, C-52/00, ECLI:EU:C:2002:252.

³⁷ Verheyen, Thomas, "Full Harmonization, Consumer Protection and Products Liability: A Fresh Reading of the Case Law of the ECJ", *European Review of Private Law* 26:1 (2018), s. 119, 121, Tillgänglig: https://www.researchgate.net/publication/323184813_Full_Harmonization_Consumer_Protection_and_Products_Liability_A_Fresh_Reading_of_the_Case_Law_of_the_ECJ [Hämtad: 14-05-2023].

³⁸ Förslag till EU-direktiv (AILD) COM(2022) 496 final, Art. 1(4).

³⁹ Ibid. Skäl 14.

⁴⁰ 1a, 1b, 7 §§ Straßenverkehrsgesetz (StVG); Juhasz, Agnes, "The Legal Framework of Autonomous Driving in Germany", *MultiScience - XXXIII. microCAD International Multidisciplinary Scientific Conference* (2019), s. 2, 4, Tillgänglig: https://www.researchgate.net/publication/342143289_The_Legal_Framework_of_Autonomous_Driving_in_Germany [Hämtad: 14-05-2023].

⁴¹ Patti, Francesco Paolo, "The European Road to Autonomous Vehicles", *Fordham International Law Journal* 43:125 (2019), s. 131-134 Tillgänglig: <https://ir.lawnet.fordham.edu/ilj/vol43/iss1/4> [Hämtad 14-05-2023].

3.3 Direktivens tillämpningsområde

3.3.1 AILD-förslaget

AILD-förslagets tillämpningsområde definieras i artikel 1(1)(a) och 1(1)(b). I artikel 1(1)(a) framförs det att direktivet innehåller nya regler om edition när skador orsakats av AI-system inom kategorin ”hög risk”. Vidare anslår artikel 1(1)(b) att direktivet framställer nya regler om placeringen av bevisbördan i mål där AI-system, oavsett riskkategori, orsakat skador.

Gemensamt för artiklarna är att de endast är tillämpliga i utomobligatoriska förhållanden där skadeståndsanspråket är grundat i culpaansvar.

Det är därmed väsentligt att förstå vad som menas med den anförda riskkategorin samt ”AI-system” i direktivets mening. Artikel 2(1) och 2(2) hänvisar båda vidare till definitionerna utsatta i AI-förordningen.⁴²

Förordningen definierar, som nämnt i avsnitt 2.2, ”AI-system” som mjukvarans utveckling skett genom en eller flera tekniker som förtecknats i bilaga 1, samt att det innehar förmågan att, för mänskligt definierade syften, generera utdata i form av innehåll, förutsägelser, rekommendationer eller beslut som påverkar den miljö systemet interagerar med.⁴³

I skäl 6 framförs det att mjukvarans kapacitet till att generera sådan utdata är av synnerlig vikt för definitionen av ett AI-system.⁴⁴ En sådan bred definition av AI-system omfattar följaktligen en väldigt stor mängd olika typer av komplex mjukvara som traditionellt inte hade ansetts vara ”AI”.⁴⁵ Bilaga 1 till förordningen försöker därför ytterligare avgränsa tillämpningsområdet genom att lista ett antal olika utvecklingstekniker, som visserligen inkluderar

⁴² Förslag till EU-förordning (AI Act) COM(2021) 206 final, Art 3.

⁴³ Ibid. Art. 3(1).

⁴⁴ Ibid. Skäl 6.

⁴⁵ O'Shaughnessy, Matt, “*One of the Biggest Problems in Regulating AI Is Agreeing on a Definition*”, Carnegie Endowment for International Peace (2022), 3:e stycket, Tillgänglig: <https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100> [Hämtad: 15-05-2023].

maskininlärning som ovan diskuterat, men även simplare tekniker som statistiska metoder.⁴⁶

Man kan därför ifrågasätta om det är av vikt att faktiskt definiera vad ett AI-system är istället för att rikta in sig på de karaktärsdrag som orsakar ansvarsproblematiken.⁴⁷ All typ av komplex mjukvara som genererar utdata, vars interna process karaktäriseras av oförutsägbarhet och låg transparens bör rimligen falla under samma regleringar oavsett om de kan klassificeras som ”AI”.

AILD-förslaget lånar som nämnt även definition av ”AI-system med hög risk” från AI-förordningen. I syfte att fokusera diskussionen på typen av system som inkluderas under kategorin hänvisar jag därmed till bestämmelserna i förordningen för definitionen.⁴⁸

Medan det är användbart att behålla samma definition av ”AI-system” i AILD-förslaget som i förordningen, så är det av relevans att diskutera huruvida definitionen av riskkategorin är till lika stor hjälp. Eftersom delar av direktivets föreslagna regleringar endast är tillämpliga på AI-system med hög risk är det av intresse att utforska vilka typer av system denna kategori potentiellt missar.

Som tidigare nämnt belyser skador orsakade av autonoma fordon ansvarsproblematiken väl, men dessa innefattas inte i AI-förordningens klassificering av hög-risk system. Förordningen undantar ett antal områden, däribland förordning (EU) 2018/858 som reglerar autonoma fordon⁴⁹, i syfte att undvika friktion med existerande sektorspecifik lagstiftning.⁵⁰

I enlighet med artikel 84 ska kommissionen, från datumet förordningen träder i kraft, årligen bedöma ifall undantagna områden ska inkluderas under regelverkets klassificering av AI-system med hög risk.

⁴⁶ Förslag till EU-förordning (AI Act) COM(2021) 206 final, Bilaga 1.

⁴⁷ O'Shaughnessy (2022), 14:e stycket.

⁴⁸ Se förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 6, Bilaga 2 och 3.

⁴⁹ Ibid. Art. 2(2).

⁵⁰ Ibid. Skäl 29.

Det är därmed oklart när och till vilken grad autonoma fordon kommer täckas av regleringarna i AILD-förslaget, vilket är anmärkningsvärt med tanke på att dessa i dagsläget redan orsakat både skador och dödsfall.

3.3.2 PLD-förslaget

Till skillnad från AILD-förslaget använder sig PLD-förslaget inte av referenser till AI-förordningen för att avgränsa dess tillämpningsområde.⁵¹ Direktivets regleringar är istället tillämpliga på ekonomiska aktörer när skador orsakade av defekta produkter tillfaller fysiska personer.⁵²

Det är därmed av relevans att utforska hur termerna ”produkt” och ”ekonomisk aktör” definieras för att förstå hur direktivet appliceras på skador orsakade av AI-system.

I kontrast till rådande produktansvarsdirektiv har PLD-förslaget expanderat sin definition av produkt till att inkludera alla typer av mjukvara.⁵³ Artikel 4(1) framför att samtliga lösöre, vare sig det är en del av annat lösöre eller fast egendom, räknas som en produkt. Detta inkluderar uttryckligen även mjukvara.⁵⁴ AI-system, oavsett om de säljs som separat mjukvara eller som del av annan produkt, inkluderas därmed i denna definition.

Produktansvarsdirektivets användning av termen ”tillverkare”⁵⁵ har också ersatts av det mer omfattande begreppet ”ekonomisk aktör”.⁵⁶ En ekonomisk aktör innefattar enligt artikel 4(16) i PLD-förslaget tillverkare av en produkt

⁵¹ Förslag till EU-direktiv (PLD) COM(2022) 495 final, Art. 2, 4.

⁵² Ibid. Art. 1.

⁵³ Det är debatterat huruvida rådande definition av produkt inkluderar mjukvara: Wagner, Gerhard, ”*Robot Liability*”, Humboldt-Universität zu Berlin (2019), s. 10-11 Tillgänglig: <https://www.rewi.hu-berlin.de/de/lf/oe/rdt/pub/working-paper-no-2> [Hämtad: 16-05-2023].

⁵⁴ Ibid. Art. 4(1).

⁵⁵ Rådets direktiv 85/374/EEG av den 25 juli 1985 om tillnärmning av medlemsstaternas lagar och andra författningar om skadeståndsansvar för produkter med säkerhetsbrister, Art. 1, 3.

⁵⁶ Förslag till EU-direktiv (PLD) COM(2022) 495 final, Art. 1, 7.

eller komponent till en produkt, leverantörer av relaterade tjänster⁵⁷, fullmaktshavande representanter, importörer, distributörer eller leverantörer av förverkligande tjänster⁵⁸.

Innebörden av dessa klumpiga termer kan kännas svårförståelig men kontentan är att PLD-förslagets tillämpning av den nya termen ”ekonomisk aktör” resulterar i att flera aktörer inom leverantör- och distributionskedjan kan hållas ansvariga för skador orsakade av en defekt produkt.

Som nämnt i tidigare avsnitt utvecklas AI-system sällan av en enda tillverkare; träningsdata kan komma från ett företag medan algoritmerna kommer från ett annat för att slutligen sättas ihop till en produkt av en tredje part. Denna mjukvara kan sedan vidare integreras i en annan produkt som exempelvis en mobiltelefon eller autonomt fordon.

Det är alltså av stor vikt att PLD-förslaget tilldelar leverantörer av relaterade tjänster skadeståndsansvar, eftersom det innebär att samtliga aktörer i produktionskedjan för AI-system såväl som leverantörer av AI-integrerade produkter kan hållas ansvariga.

3.4 Utlösning av ansvarsmekanismen

Regleringarna i AILD-förslaget tillämpas när skada orsakad av AI-system inträffat i utomobligatoriska förhållanden och ansvarsgrunden baseras på culpa. Eftersom detta diskuterats ovan kommer följande avsnitt fokusera på PLD-förslaget och dess definition av defekt produkt.

PLD-förslaget kräver, på samma vis som det nuvarande produktansvarsdirektivet, att en produkt är defekt för att en part ska hållas ansvarig.⁵⁹

⁵⁷ Digitala tjänster som är integrerade i eller sammankopplade med produkten; Ibid. Art. 4(4).

⁵⁸ Exempelvis speditörtjänster, lagerhantering; Ibid. Art. 4(14).

⁵⁹ Ibid. Art. 5; Rådets direktiv 85/374/EEG av den 25 juli 1985 om produktansvar, Art. 1.

En produkt anses vara defekt om den innehar en säkerhetsbrist som allmänheten inte kan förväntas räkna med, och förslaget listar härnäst ett antal omständigheter som det i bedömningen ska tas hänsyn till.⁶⁰

PLD-förslaget bygger vidare på omständigheterna i det gällande produktansvarsdirektivet med fem nya alternativ, där artikel 6(1)(c) är av störst intresse. Denna framför att en produkts förmåga att vidareutveckla sin prestanda efter den lanserats på marknaden ska tas med vid bedömningen om den är defekt.⁶¹

Ponera att ett AI-system i ett autonomt fordon, givet att det innehar denna förmåga, utvecklar en aggressiv körstil till följd av andra mänskliga förare. För att undvika att deras produkt ska anses defekt måste nu en ekonomisk aktör se till att systemet inte utvecklas i en oönskad riktning.

På samma sätt kan en språkmodell som lär sig av sina användare urarta om en ekonomisk aktör inte kurerar indata till sitt system, vilket var det som inträffade med Microsofts chatrobot Tay. Strax efter AI-systemet lanserades började det, efter flertalet rasistiska interaktioner med användare, upprepa dessa; Microsoft fick stänga ner chatroboten efter bara 16 timmar.⁶²

Tillägget av bestämmelsen i artikel 6(1)(c) bemöter på så vis en potentiell aspekt av ansvarsproblematiken, även om det inte är alla AI-system som har denna förmåga att vidareutveckla sin prestanda.

⁶⁰ Förslag till EU-direktiv (PLD) COM(2022) 495 final, Art. 6(1).

⁶¹ Rådets direktiv 85/374/EEG av den 25 juli 1985 om produktansvar, Art. 6.

⁶² Hunt, Elle, “*Tay, Microsoft’s AI chatbot, gets a crash course in racism from Twitter*”, The Guardian 24/03 2016, Tillgänglig: <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter#> [Hämtad 16-05-2023].

3.5 Nya editionsregler

3.5.1 Artikel 3 AILD-förslaget

I syfte att underlätta för skadelidande att driva en skadeståndstalan, när skada orsakats av AI-system med hög risk, innehåller artikel 3 i AILD-förslaget nya regler om edition och krav på bevarande av bevis.

Reglerna framför att domstolen, på begäran av den som för skadeståndstalan eller överväger att göra det, ska besluta att en aktör⁶³ ska bevara och lämna ut tillgängliga bevis om hög-risk AI-system som misstänks ha orsakat skada.

För att bevilja en sådan begäran måste skadelidande lägga fram tillräckliga bevis till stöd av att sitt anspråk troligtvis inträffat, samt ha vidtagit alla rimliga åtgärder att på egen hand få ut bevisen från den påstådda skadevällaren.

Vid beviljande av editionsföreläggande ska domstolen begränsa detta till det som är nödvändigt och proportionerligt. Bedömningen ska även ta hänsyn till regler om konfidentiell information, såsom företagshemligheter och information som rör nationell säkerhet.

Slutligen ska domstolen, i det fall ett editionsföreläggande inte följs, presumera att den påstådda skadevällaren varit oaktsam och därmed agerat culpöst.

3.5.2 Artikel 8 PLD-förslaget

Artikel 8 i PLD-förslaget motsvarar AILD-förslagets regel om edition men särskiljer sig i ett antal punkter.

Eftersom reglerna i PLD-förslaget appliceras på defekt produkt är editionsplikten även tillämplig på skador orsakade av AI-system som inte kategoriseras som hög-risk.

⁶³ Som i AI-förordningens mening anses vara ansvarig, se förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 24, 28(1).

Vidare behöver den som för skadeståndstalan inte heller ha vidtagit alla rimliga åtgärder att på egen hand få ut bevisen innan domstolen kan bevilja ett editionsföreläggande. Det saknas också en ekvivalent bestämmelse i förslaget som tillåter domstolen att förelägga en potentiell skadevållare att bevara bevis.

Slutligen kan ett editionsföreläggande, i kontrast till AILD-förslaget, inte åberopas förrän huvudförhandling påbörjats.

3.5.3 Fördelar, nackdelar och inverkan på ansvarsproblematiken

De nya regleringarna om editionsföreläggande är i min mening effektiva rättsmedel för att motverka problematiken med AI-systemens brist på transparens. Dessutom möjliggör det för en skadelidande part att fatta ett informerat och rationellt beslut om de vill gå vidare med ett skadeståndsanspråk, samt identifiera korrekt skadevållare⁶⁴. Detta minskar i sin tur kostnaden för de involverade parterna och minskar belastningen på domstolarna — särskilt i relation till AILD-förslagets artikel 3 som kan tillämpas redan under muntlig förhandling.⁶⁵

Det är därmed märkligt att artikel 8 i PLD-förslaget endast är tillämplig efter skadeståndsanspråk tagits upp till huvudförhandling, eftersom det enbart utökar kostnaderna för skadelidande som söker kompensation i medlemsstater där ytterligare avgifter tillkommer för detta. Förslaget anför inte heller något skäl till varför denna skillnad mellan direktiven existerar.⁶⁶

Båda förslagen framför dessutom att editionsplikten endast innefattar de bevis och information som svaranden tillhandahåller. Domstolen, som ska avgränsa editionsföreläggandet till vad som är nödvändigt och proportionerligt, måste därmed röra sig inom gränserna för vad svaranden har tillgängligt. Rimligtvis bör det finnas ett lagstadgat krav på att sådan bevisning och information ska

⁶⁴ Se även: förslag till EU-direktiv (AILD) COM(2022) 496 final, Skäl 17; förslag till EU-direktiv (PLD) COM(2022) 495 final, Skäl 30-31.

⁶⁵ Förslag till EU-direktiv (AILD) COM(2022) 496 final, s. 12.

⁶⁶ Se förslag till EU-direktiv (PLD) COM(2022) 495 final, Skäl 30-31, s. 12-13.

finnas tillgänglig om editionsplikten ska fungera i praktiken. Kapitel 2 i AI-förordningen fastställer visserligen typen av information som ska tillhandhållas för AI-system med hög risk, men sådana krav finns inte för AI-system av annan riskkategori.⁶⁷

PLD-förslagets editionsplikt riskerar därmed att bli verkningslös och man kan ifrågasätta varför AILD-förslaget, givet att det enbart är tillämpligt på hög-risk AI-system, inte fastställer tillhandahållandet av information som ett lagstadgat krav. En skadevållande part kan därmed undvika att presumtionen om culpa blir verksam genom att inte ha den nödvändiga informationen tillgänglig. Detta skulle givetvis leda till påföljder enligt kapitel 10 i AI-förordningen, men det är oklart varför ett sådant kryphål existerar.⁶⁸

Det kan även ifrågasättas varför AILD-förslaget endast är tillämpligt på AI-system med hög risk. Enligt skäl 18 till förslaget är syftet med detta att undvika att ålägga ytterligare krav, än vad AI-förordningen föreskriver, på tillhandahållandet av information för leverantörer av AI-system med lägre riskklassificering.⁶⁹

Avgränsning verkar på så sätt rimlig ur ett proportionalitetsperspektiv, men eftersom PLD-förslagets editionsplikt appliceras på leverantörer av AI-system med lägre risk framstår denna särskiljning som meningslös. En leverantör av AI-system måste följaktligen ändå tillhandahålla åtminstone någon form av information för att inte stå i strid med ett editionsföreläggande.

Slutligen förlitar sig AILD-förslaget dessutom på att det sker en korrekt klassificering av AI-system med hög risk. Som tidigare nämnt innefattar denna kategori för tillfället inte autonoma fordon, vilket innebär att skadelidande i nuläget inte kan avhjälpas genom förslagets regleringar.

⁶⁷ Förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 8-15.

⁶⁸ Se Ibid. Art. 71.

⁶⁹ Förslag till EU-direktiv (AILD) COM(2022) 496 final, Skäl 18.

3.6 Placering av bevisbördan

3.6.1 Artikel 4 AILD-förslaget

Artikel 4 i AILD-förslaget åsyftar att underlätta en skadeståndstalan som grundas i att skada orsakats av ett AI-systems utdata eller underlåtenhet att producera utdata till följd av oaktsamhet. Artikeln innehåller därmed regler om placeringen av bevisbördan för orsakssambandet mellan påstådd skadevållares oaktsamhet och genererad utdata eller underlåtenheten att producera utdata.

Direktivet framför att domstolen under vissa omständigheter ska presumera ett sådant orsakssamband, givet att skadelidande till en början kan visa att motparten varit oaktsam eller att sådan oaktsamhet presumeras till följd av artikel 3(5) AILD-förslaget. Vidare krävs det att genererad utdata eller underlåtenheten att producera utdata sannolikt påverkats av oaktsamheten samt att utdata, eller brist därav, orsakat skadan.⁷⁰

Eftersom förslagets bevisbörderegler inte enbart är tillämpliga på AI-system med hög risk följer bestämmelser med särskilda krav för dessa.⁷¹ För skadeståndsanspråk som avser hög-risk AI-system ställs det ett antal krav för att en leverantörs eller användares oaktsamhet ska anses bevisad eller presumerad.

Enligt artikel 4(2) AILD-förslaget måste skadelidande visa att ett hög-risk AI-system inte uppfyller kraven stadgade i AI-förordningen, i det fall en leverantörs oaktsamhet ska presumeras i enlighet med artikel 4(1). På motsvarande sätt ställs särskilda krav för skadeståndsanspråk mot användare av AI-system med hög risk enligt artikel 4(3).⁷²

⁷⁰ Förslag till EU-direktiv (AILD) COM(2022) 496 final, Art. 4(1).

⁷¹ Ibid. Art. 1(b).

⁷² Se förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 8-29.

Presumtionsregeln ska däremot inte tillämpas givet att den påstått skadeståndsskyldige kan visa att nödvändig bevisning, med rimliga ansträngningar, finns tillgängligt för att fastställa orsakssambandet.⁷³

Slutligen ska domstolen, i fråga om AI-system med lägre risk, endast tillämpa presumtionsregeln om den anser det orimligt svårt att bevisa orsakssambandet. Påstådd skadeståndsskyldig har dessutom rätt till att motbevisa presumptionen.⁷⁴

3.6.2 Artikel 9 PLD-förslaget

Strukturen i artikel 9 PLD-förslaget motsvarar på många sätt artikel 4 i AILD-förslaget. I syfte att åstadkomma en bättre balans mellan skadelidande och produkttillverkare, särskilt i komplicerade mål som exempelvis berör AI-system, föreskriver artikel 9 PLD-förslaget att domstolen under vissa förutsättningar ska presumera att en produkt är behäftad med defekt och att det finns ett orsakssamband mellan skadan och produktens defekt.⁷⁵

Artikel 9(2) framför att ett sådant orsakssamband ska presumeras givet att en påstådd skadevällare inte har följt ett editionsföreläggande, skadelidande bevisar att en produkt inte levt upp till säkerhetskrav eller att skadan orsakats av uppenbart tekniskt fel under normala omständigheter.

Likaså presumeras orsakssambandet i det fall skadan som inträffat är typisk för den fastställda defekten, eller om produktens tekniska komplexitet försvårar för skadelidande att bevisa orsakssamband alternativt produktens defekt. Påstådd skadevällare har, likt artikel 4 AILD-förslaget, rätt att motbevisa presumptionen.⁷⁶

⁷³ Förslag till EU-direktiv (AILD) COM(2022) 496 final, Art. 4(4).

⁷⁴ Ibid. Art. 4(5)-4(7).

⁷⁵ Förslag till EU-direktiv (PLD) COM(2022) 495 final, Skäl 30, 33-34.

⁷⁶ Ibid. Art. 9(3), 9(4).

3.6.3 Fördelar, nackdelar och inverkan på ansvarsproblematiken

Inledningsvis ämnar de nya bevisbördereglerna, liksom reglerna om edition, att bemöta problemet med brist på transparens i vissa typer av AI-system. Genom att införa en presumtion om orsakssambandet mellan skada och antingen oaktsamhet eller defekt, förflyttas bevisbördan till påstådd skadevällare i syfte att uppnå en mer rättvis risk- och kostnadsfördelning. Detta har i min mening en positiv inverkan på den aspekt av ansvarsproblematiken som berör AI-systemens brist på insyn, men det finns dessvärre ett antal anmärkningsvärda svagheter.

På grund av AI-systemens tekniska komplexitet är det inte säkert att en skadelidande på egen hand kan tolka informationen som krävs för att bevisa regleringarnas rekvisit; i många fall krävs det därmed att experter hyrs in.

Exempelvis inträffar presumtionen enligt artikel 4(2) AILD-förslaget, givet att skadelidande kan fastställa att en leverantör inte har levt upp till något av kraven i AI-förordningen. Ett antal av dessa krav kan vara svåra att bevisa, däribland huruvida träningsdata haft en inverkan på skadan, utan genomgående insyn i AI-systemets funktionssätt.⁷⁷

Editionsplikten i artikel 3 bemöter detta genom att tillgängliggöra den nödvändiga informationen, och artikel 4 befriar följaktligen skadelidande från att behöva bekosta experter i syfte att bevisa orsakssambandet mellan skada och oaktsamhet.

Således sänks kostnaden för ett potentiellt skadeståndsanspråk, men denna lättnad sträcker sig endast så långt — en skadelidande måste fortfarande, sannolikt med stöd av inhyrd expertis, bevisa en överträdelse av kraven utsatta i AI-förordningen.

⁷⁷ Se förslag till EU-förordning (AI Act) COM(2021) 206 final, Art. 10(2)-10(4).

Det kan argumenteras att det är i sakens natur att någon form av experthjälp krävs i rättsliga situationer som berör så pass komplexa ämnen som AI-system, men det kommer vara en stor utmaning för skadelidande att bevisa sådana överträdelser.⁷⁸ Inhyrning av sådan expertis ökar inte bara kostnaderna utan slutresultatet förblir osäkert; exempelvis kommer en AI-expert inte kunna fastställa om någon överträdelse skett förrän de har spenderat flertalet timmar med att analysera loggar och träningsdata.⁷⁹ Detta försvåras dessutom ytterligare i fall som berör AI-system med lägre riskklassificering, som exempelvis autonoma fordon, eftersom dessa inte har något lagstadgat krav på den typ av information som ska tillhandahållas. Denna informationsparadox verkar således till skadelidandes nackdel; man vet inte om det föreligger någon överträdelse förrän kostsamt inhyrda experter analyserar informationen.

Ett ytterligare exempel kan vi hitta i artikel 9(2) i PLD-förslaget. En produkts defekt presumeras i det fall: ett editionsföreläggande inte följts, om utsatta säkerhetskrav inte levts upp till, eller skadan orsakats av uppenbart tekniskt fel.

Eftersom PLD-förslaget inte appliceras på hög-risk AI-system saknas det tydligt underlag för vad ett editionsföreläggande ska innehålla. Detta kan leda till incitament för en påstådd skadevållare att dela med sig av stora mängder ofullständig och irrelevant information, i syfte att överväldiga skadelidande med tidskrävande och kostsam analys.

För att undvika dessa kostnader kan skadelidande rikta in sig på de två återstående alternativen, men även dessa kan bli svåra att bevisa. Säkerhetskraven rör en tillverkares interna processer och vad som är ett uppenbart tekniskt fel synes, med tanke på AI-systemens karaktärsdrag, vara svårt att fastställa.

Skadelidandes möjligheter till att uppfylla presumtionskraven förefaller således bli dyrt eller rent av omöjligt. En potentiell lösning skulle vara att tydligt

⁷⁸ Se avsnitt 2.2 om ”*the black-box problem*”.

⁷⁹ Nawaz, Samar Abbas, ”*The Proposed EU AI Liability Rules: Ease or Burden?*”, European Law Blog (2022), Tillgänglig: <https://europeanlawblog.eu/2022/11/07/the-proposed-eu-ai-liability-rules-ease-or-burden/> [Hämtad: 19-05-2023].

fastställa att ett editionsföreläggande enligt artikel 8(1) inte bara innefattar tillhandahållen information, utan specifikt det som behövs för att kunna bevisa rekvisiten i artikel 9(2).

4 Slutsats och reflektioner

Komplexiteten i AI-systemens tekniska funktionssätt kombinerat, med deras förmåga att vidareutveckla sin prestanda, kan leda till utdata som inte enkelt kan härledas till en specifik intern process. Detta så kallade ”*black box-problem*” leder till att dessa system karaktäriseras av en brist på transparens och oförutsägbarhet. Dessutom bygger AI-system ofta på beståndsdelar från olika leverantörer, som tillsammans med dess inneboende egenskaper resulterar i svårigheter att precisera en ansvarande part när skador inträffar.

De två förslagen till EU-direktiv avser att bemöta detta problem genom ett antal nya regleringar, som till synes inte når hela vägen fram. AILD-förslagets föreskrivna miniharmonisering riskerar exempelvis fragmentera lagstiftningen bland medlemsstater, vilket kan hindra upptaget av AI-system bland europeiska verksamheter. Förslagets ändamål kan på så sätt, ironiskt nog, äventyras.

Två huvudsakliga problem identifierades i relation till AILD-förslagets tillämpningsområde. I första hand dess breda definition av ett AI-system, eftersom det i princip innefattar alla typer av komplex mjukvara trots att de traditionellt inte hade betraktats som ”AI”. Därmed kan man ifrågasätta om det är värdefullt att överhuvudtaget definiera ett AI-system. Istället kan det vara fördelaktigt att angripa de egenskaper som karaktäriserar ett AI-system (i.e. komplexitet, brist på transparens och oförutsägbarhet) eftersom det är dessa som ligger till grund för ansvarsproblematiken i sig.

I andra hand identifierades även förslagets definition av hög-risk AI-system. Kategorin innehåller i nuläget inte vissa typer av system som av allt att döma borde klassas som hög risk, däribland autonoma fordon. Dessa system har redan i dagsläget orsakat skador och dödsfall, vilket gör det är anmärkningsvärt att dessa inte täckts under förslagets regleringar.

Till PLD-förslagets fördel kan nu en ekonomisk aktör hållas ansvarig för en produkts defekt. Tillämpningen av den nya termen medför att flera aktörer

inom leverantörskedjan kan hållas ansvariga. På grund av ”*the black box-problem*” är det svårt att fastställa vilken komponent av ett AI-system som kan ha orsakat skadan. Genom att hålla samtliga delar av leverantörskedjan gemensamt ansvariga för en produkts defekt kringgår PLD-förslaget denna aspekt av ansvarsproblematiken.

Vidare föreskriver även PLD-förslaget nya omständigheter som ska tas med i beaktning för att avgöra om en produkt är defekt, däribland en produkts förmåga att vidareutveckla sin prestanda. Tillägget är av betydelse för den typ av AI-system som innehar sådan förmåga och bemöter således en aspekt av ansvarsproblematiken — ett tydligt försök till att skydda skadelidande mot ett AI-systems potentiellt oförutsägbara självutvecklingsförmåga.

Direktivens föreslagna regler om editionsföreläggande innehar främst funktionen att tillgängliggöra information för skadelidande, så att de effektivt kan framställa ett skadeståndsanspråk. På så sätt fastställer direktiven ett transparenskrav på AI-system, men deras tillvägagångssätt skapar dessvärre ett antal hinder för skadelidande att överkomma.

Ett utsatt mål med dessa editions- och bevisbörderegler är att möjliggöra för skadelidande att identifiera korrekt skadevållare, minska kostnaden för de involverade parterna samt minska belastningen på domstolarna.

Det framstår därmed som märkligt att presumtionen i artikel 8 PLD-förslaget inte är tillämplig förrän i huvudförhandling eftersom detta i många fall endast kommer utöka kostnaderna.

Dessutom saknar PLD-förslaget krav på typen av information ett editionsföreläggande ska innehålla, vilket kan motivera en påstådd skadevållare till att lämna över mängder av ofullständig och irrelevant information. Till följd måste skadelidande bekosta tidkrävande och dyr analys av informationen, som inte nödvändigtvis resulterar i tillräckliga bevis för att med framsteg åberopa bevisbördepresumtionen. Till följd skapas en informationsparadox till skadelidandes nackdel, där denna tvingas investera stora summor pengar utan

att veta om det ens finns en grund till ett skadeståndsanspråk — därav kan ett editionsföreläggande bli verkningslöst i praktiken.

Avslutningsvis är direktivens föreslagna regleringar, enligt mig, ett steg i rätt riktning till att lösa ansvarsproblematiken. Det är emellertid oklart om regleringarna i praktiken uppnår en tillräckligt balanserad risk- och kostnadsfördelning, speciellt med avseende på de höga kostnader som skadelidande förväntas bära. Dessa lagförslag är dock fortfarande i den inledande lagstiftningsfasen och kommer sannolikt gå igenom förändringar innan deras slutgiltiga antagande.

Källförteckning

Rapporter

Karner, Ernst, Koch, Bernhard A., Geistfeld, Mark A., “*Comparative Law Study on Civil Liability for Artificial Intelligence*”, Rapport till Europeiska kommissionen (2021), s. 9 Tillgänglig: <https://op.europa.eu/en/publication-detail/-/publication/8a32ccc3-0f83-11ec-9151-01aa75ed71a1> [Hämtad: 06-05-2023]

von der Leyen, Ursula, ”*A Union that strives for more: My agenda for Europe*”, 2019, Tillgänglig: https://eunec.eu/sites/www.eunec.eu/files/attachment/files/political-guidelines-next-commission_en_kopie.pdf [Hämtad 06-05-2023]

EU-dokument

Europeiska kommissionen, ”*New liability rules on products and AI to protect consumers and foster innovation*”, 2022, Tillgänglig: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807 [Hämtad: 06-05-2023]

Europeiska kommissionen, “*Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*”, 2020, Tillgänglig: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064&from=en> [Hämtad: 11-05-2023]

Europeiska kommissionen, “*European enterprise survey on the use of technologies based on artificial intelligence – Final report*”, 2020, Tillgänglig: <https://data.europa.eu/doi/10.2759/759368> [Hämtad 13-05-2023]

Europeiska kommissionen, “*Questions and answers on the revision of the Product Liability Directive*”, 2022, Tillgänglig: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/qanda_22_5791/QANDA_22_5791_EN.pdf [Hämtad: 13-05-2023]

Lagstiftning

Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar för produkter med säkerhetsbrister (PLD) COM(2022) 495 final

Förslag till Europaparlamentets och rådets direktiv om skadeståndsansvar gällande artificiell intelligens (AILD) COM(2022) 496 final

Förslag till Europaparlamentets och rådets förordning om harmoniserade regler för artificiell intelligens (AI Act) COM(2021) 206 final

Rådets direktiv 85/374/EEG av den 25 juli 1985 om tillnärmning av medlemsstaternas lagar och andra författningar om skadeståndsansvar för produkter med säkerhetsbrister,

Straßenverkehrsgesetz (StVG)

Rättsfallsregister

Domstolens dom den 25 april 2002, *Europeiska kommissionen mot Republiken Grekland*, C-154/00, ECLI:EU:C:2002:254

Domstolens dom den 25 april 2002, *Europeiska kommissionen mot Franska republiken*, C-52/00, ECLI:EU:C:2002:252

Domstolens dom den 25 april 2002, *María Victoria González Sánchez mot Medicina Asturiana SA.*, C-183/00, ECLI:EU:C:2002:255

Litteratur

Hellner, Jan, Radetzki, Marcus, "*Skadeståndsrätten*", 11. Uppl., Norstedts Juridik, Visby (2021)

Kleineman, Jan, "*Rättsdogmatisk Metod*" i Nääv, Maria, Zamboni, Mauro (red.), "*Juridisk metodlära*", 2. Uppl., Studentlitteratur, Lund (2018)

Turner, Jacob, "*Robot Rules: Regulating Artificial Intelligence*", 1. Uppl., Springer International Publishing AG, (2018)

Vetenskapliga artiklar

Juhasz, Agnes, "*The Legal Framework of Autonomous Driving in Germany*", MultiScience - XXXIII. microCAD International Multidisciplinary Scientific Conference (2019), Tillgänglig: https://www.researchgate.net/publication/342143289_The_Legal_Framework_of_Autonomous_Driving_in_Germany [Hämtad: 14-05-2023]

Nawaz, Samar Abbas, "*The Proposed EU AI Liability Rules: Ease or Burden?*", European Law Blog (2022), Tillgänglig: <https://europeanlawblog.eu/2022/11/07/the-proposed-eu-ai-liability-rules-ease-or-burden/> [Hämtad: 19-05-2023]

O'Shaughnessy, Matt, "*One of the Biggest Problems in Regulating AI Is Agreeing on a Definition*", Carnegie Endowment for International Peace (2022), Tillgänglig: <https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100> [Hämtad: 15-05-2023]

Patti, Francesco Paolo, “*The European Road to Autonomous Vehicles*”, *Fordham International Law Journal* 43:125 (2019), Tillgänglig: <https://ir.lawnet.fordham.edu/ilj/vol43/iss1/4> [Hämtad 14-05-2023]

Turing, Alan, ”*Computing Machinery and Intelligence*”, Oxford University Press, *Mind*, 59:236 (1950), Tillgänglig: <https://doi.org/10.1093/mind/LIX.236.433> [Hämtad: 08-05-2023]

Verheyen, Thomas, “*Full Harmonization, Consumer Protection and Products Liability: A Fresh Reading of the Case Law of the ECJ*”, *European Review of Private Law* 26:1 (2018), Tillgänglig: https://www.researchgate.net/publication/323184813_Full_Harmonization_Consumer_Protection_and_Products_Liability_A_Fresh_Reading_of_the_Case_Law_of_the_ECJ [Hämtad: 14-05-2023]

Wagner, Gerhard, ”*Robot Liability*”, Humboldt-Universität zu Berlin (2019), Tillgänglig: <https://www.rewi.hu-berlin.de/de/lf/oe/rdt/pub/working-paper-no-2> [Hämtad: 16-05-2023]

Artiklar

Dastin, Jeffrey, “*Amazon scraps secret AI recruiting tool that showed bias against women*”, Reuters 11/10 2018, Tillgänglig: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> [Hämtad: 11-05-2023]

Hunt, Elle, “*Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*”, *The Guardian* 24/03 2016, Tillgänglig: <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter#> [Hämtad 16-05-2023]

Levin, Sam, “*Safety driver charged in 2018 incident where self-driving Uber car killed a woman*”, *The Guardian*, 16/9 2020, Tillgänglig:

<https://www.theguardian.com/us-news/2020/sep/16/uber-self-driving-car-death-safety-driver-charged> [Hämtad 06-05-2023]

Metz, Cade, Schmidt, Gregory, “*Elon Musk and Others Call for Pause on A.I., Citing ‘Profound Risks to Society’*”, The New York Times, 29/3 2023, Tillgänglig: <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html> [Hämtad: 09-05-2023]