

Handelshögskolan vid Göteborgs Universitet

Affärsjuridik och Innovationsrätt I, 30 HP

Juridiska Institutionen

Artificiell intelligens – subjekt eller objekt?

EN FUNKTIONALISTISK KRITIK AV DEN ONTOLOGISKA
FIXERINGEN I AI-RÄTTEN

ARTA EKMAN OCH OLIVER MARTENS



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Innehållsförteckning

1	BAKGRUND.....	2
2	METODOLOGISK UTGÅNGSPUNKT.....	3
3	KLASSIFICERINGEN METODOLOGISKA PROBLEM.....	3
3.1	KLASSIFICERING SOM SLUTPUNKT, INTE UTGÅNGSPUNKT	4
3.2	PARALLELLER I CIVILRÄTTENS BEGREPPSBILDNING	4
4	ROSS OCH DE TOMMA MELLANLEDEN	5
5	REDUCTIO I: TELEVISIONENS PARADOX.....	6
6	REDUCTIO II: TURINGTESTETS VILLFARELSE	7
7	BLACK-BOX-ARGUMENTETS BEGRÄNSNINGAR	8
8	AI FÖRORDNINGEN SOM REPRODUCERAD METODFÄLLA	9
9	DEN EPISTEMISKA FÖRKLARINGEN: MINIRÄKNARENS ALLEGORI	10
10	KONSEKVENSERNA AV FELORIENTERAD REGLERING	11
11	EN ALTERNATIV ORIENTERING	12
12	AVSLUTNING	13
	BIBLIOGRAPHY OF REFERENCES / KÄLLFÖRTECKNING	15

1 Bakgrund

Den samtida debatten om reglering av artificiell intelligens (AI) hemsöks av en ontologisk fixering. Från Europaparlamentets tidigare diskussioner om att tillskriva robotar ”elektronisk personlighet” till det spektakulära beviljandet av medborgarskap för roboten Sophia, tycks rättsvetenskapen vara låst i en föreställning om att teknikens *vara* måste fastställas innan dess *böra* kan regleras.¹ Denna fixering vid artefaktens status återfinns även i EU:s nyligen antagna AI-förordning (AI Act), som trots sin uttalade riskbaserade ansats vilar på en binär klassificeringslogik där systemets tekniska egenskaper dikterar de normativa skyldigheterna.² Premissen för denna diskurs är att frågan ”är AI ett rättssubjekt eller ett rättsobjekt?” utgör en nödvändig grundval för rättslig analys.

Vår tes är att denna frågeställning är metodologiskt felställd. Genom att fokusera på AI-systemets status faller lagstiftare och jurister i den fälla som den skandinaviska rättsrealismen varnade för redan vid 1900-talets mitt: Begreppsjurisprudens. I stället för att analysera de reella intressekonflikterna och ansvarsrelationerna mellan de människor som utvecklar, säljer och använder tekniken, tillåts abstrakta begrepp styra rättstillämpningen.

För att dekonstruera denna problematik anlägger vi ett funktionalistiskt perspektiv med rötter hos Alf Ross och Claes Martinson. Ross visade i sin inflytelserika analys av begreppet Tû-Tû hur juridiska termer ofta fungerar som semantiskt tomma mellanled, en presentationsteknik för att binda samman rättsfakta med rättsföljder utan att i sig referera till en verklig egenskap i sinnevärlden.³ På samma sätt riskerar begreppet ”AI” idag att fungera som ett modernt Tû-Tû: ett magiskt mellanled vi fyller med föreställningar om agens och autonomi för att hantera teknikens komplexitet.

I enlighet med den skandinaviska funktionalismen, såsom den beskrivs av Martinson, bör juridiska begrepp som ”äganderätt”, eller i detta fall ”rättssubjektivitet” förstås relationellt och kontextuellt snarare än som statiska storheter.⁴ Genom att applicera detta synsätt syftar denna uppsats till att visa att ”black box”-problematiken och AI:s påstådda autonomi inte kräver en ny ontologi. Det krävs i stället en återgång till en strikt ansvarsallokering där tekniken betraktas som ett verktyg i mänskliga handlingskedjor, oavsett dess komplexitet.

¹ Merima Bruncevic, *föreläsning 2 och 3*, fördjupningskursen ”AI och rätten”, Handelshögskolan vid Göteborgs universitet (2026-01-21; 2026-01-22).

² Martin Ebers, *“Truly Risk-Based Regulation of Artificial Intelligence”*, 2024, s. 686; Daniel Braun, *“Why ‘Artificial Intelligence’ Should Not Be Regulated”*, 2025, s. 6.

³ Alf Ross, *“Tû-Tû”*, 1957.

⁴ Wolfgang Faber och Claes Martinson, *“Can Ownership Limit the Effectiveness of EU Consumer Contract Law Directives?”*, 2019, s. 91–92; Claes Martinson, *“The Scandinavian Approach to Property Law”*, 2014, s. 17.

2 Metodologisk utgångspunkt

Analysen tar sin utgångspunkt i den skandinaviska funktionalismen, ett metodologiskt angreppssätt där rättsliga problem förstås relationellt och kontextuellt. Kärnan i denna tradition är att rättens uppgift inte är att deducera lösningar ur abstrakta begrepp, utan att hantera konkreta intressekonflikter i specifika situationer.⁵ Juridiska begrepp används därför inte som analytiska startpunkter, utan som språkliga verktyg för att kommunicera resultatet av en redan genomförd analys.⁶

Inom detta perspektiv saknar begrepp såsom ”äganderätt” eller ”rättssubjektivitet” självständig förklaringskraft. De fungerar i stället som sammanfattande etiketter för hur rättsordningen, i ett givet sammanhang, har valt att ordna relationerna mellan berörda aktörer.⁷ Den metodologiska poängen är därmed att analysen av de bakomliggande intressena måste föregå begreppsbildningen, inte tvärtom.

Även om denna förståelse har sina historiska rötter i Alf Ross analys av semantiskt tomma mellanled (*se vidare avsnitt 4*), tar vår analys sin primära utgångspunkt i den skandinaviska funktionalismens samtida artikulering. Traditionen har utvecklats från en ren begreppskritik till ett konstruktivt verktyg för att hantera moderna rättsliga fenomen.⁸

Tillämpad på AI-rättens område innebär detta att frågan om AI-systemens rättsliga ”status” inte behandlas som ett ontologiskt problem, utan som ett metodologiskt sådant. Fokus förskjuts från vad tekniken är – till hur ansvar, kontroll och beslutsmyndighet faktiskt är fördelade i de mänskliga handlingskedjor där systemen utvecklas och används.⁹

Utifrån denna metodologiska plattform genomförs analysen i tre steg. Först prövas hållbarheten i att tillskriva AI-system subjektstatus genom *reductio ad absurdum*-resonemang. Därefter analyseras de epistemiska orsaker som ligger till grund för den felställda frågan, med fokus på hur teknisk opacitet (*black box*) felaktigt tolkas som agens. Slutligen förs en normativ diskussion om en alternativ orientering av regleringen, där fokus flyttas från artefaktens klassificering till mänsklig kontroll, ansvarsfördelning och institutionell styrning.

3 Klassificeringen metodologiska problem

Med den metodologiska utgångspunkt som etablerats ovan kan en återkommande tendens i AI-rättslig diskussion identifieras. Juridiska resonemang inleds ofta med en klassificering av tekniken, snarare än med en analys av de relationer och handlingskedjor där rättsliga intressen faktiskt uppstår. I detta kapitel prövas hållbarheten i ett sådant angreppssätt. Genom konkreta

⁵ Faber och Martinson, 2019, s. 91–92.

⁶ Martinson, 2014, s. 17.

⁷ Faber & Martinson, 2019, s. 93; Martinson, 2014, s. 18.

⁸ Martinson, 2014, s. 16–18.

⁹ Jfr Faber & Martinson, 2019, s. 96 om “identifying the real issue”.

exempel och jämförelser med andra rättsområden visas hur klassificeringsfixering riskerar att förskjuta juridisk analys från ansvar och kontroll till begreppsliga etiketter.

3.1 Klassificering som slutpunkt, inte utgångspunkt

Juridisk klassificering framstår ofta som en nödvändig första operation, att fastställa vad fenomenet är för att därefter deducera vilka rättsföljder som ska tillämpas. Men detta sätt att närma sig rättsliga problem är inte neutralt. Klassificering bör förstås som ett hjälpmedel för systematisering och normativ utfyllnad, inte som en metod för att fastställa verklighetens ontologiska natur.

Inom den skandinaviska funktionalismen betonas att begrepp inte ska användas för att styra lösningen genom begreppslig logik (*conceptual logic*). Claes Martinson framhåller att juridiska begrepp visserligen är nödvändiga, men att de måste användas relationellt och kontextuellt. Faran uppstår när begreppen tillåts leva sitt eget liv och diktera lösningar som om de vore fysiska objekt, snarare än verktyg för att kommunicera resultatet av en intresseavvägning.¹⁰

Detta är avgörande för vår kritik av AI-diskursen. Problemet är inte att begreppet ”AI-system” används, utan hur det används. När begreppet behandlas som en ontologisk kategori med inneboende egenskaper som per automatik genererar normativa slutsatser, överskrider det sin funktion som kommunikationsverktyg. Vi ser här en återkomst av den begreppsjurisprudens funktionalismen varnat för, klassificeringen av systemet tillåts styra den rättsliga lösningen, vilket skymmer de faktiska intressekonflikterna mellan de mänskliga aktörerna bakom tekniken.

Att juridiken brottas med denna typ av gränsdragningsproblem är inte nytt. Svante Bergström visade i sin analys av gränsen mellan pantavtal och köp hur diskussionen ofta dominerats av ”formella problem” (klassificering) på bekostnad av de ”reella problemen” (vilka regler som bör gälla). Han varnade för tendensen att betrakta gränslinjerna mellan rättsliga kategorier som mer ”rena” än de faktiskt är eller bör vara.¹¹ Samma metodologiska insikt måste appliceras på AI-diskursen: Den praktiskt relevanta frågan är inte vad AI är i metafysisk mening, utan vilka ansvarsrelationer och kontrollmekanismer som bör gälla för specifika risker.

3.2 Paralleller i civilrättens begreppsbildning

Det metodologiska problemet, att låta ontologiska etiketter styra normativa slutsatser har uppmärksamrats tydligt inom insolvensrätten, specifikt i diskussionen om ”fordrans

¹⁰ Martinson, 2014, s. 16–17; se även Faber och Martinson, 2019, s. 92.

¹¹ Svante Bergström, ”Pantavtal eller köp?”, 1969, s. 312–313.

uppkomst”. Frågan när en fordran uppstår kan vid en första anblick framstå som en sökning efter ett faktiskt ögonblick i tiden då en förpliktelse ”föds”. Jonatan Schytzer har dock visat att detta är en chimär; fordrans uppkomst är inte ett faktum utan en juridisk slutsats som varierar beroende på vilket insolvensrättsligt syfte som ska uppfyllas, exempelvis kvittningsrätt eller rätten att delta i konkurs. Frågan ”har fordran uppkommit?” är i själva verket en omskrivning av den funktionella frågan: ”Bör denna borgenär ges företräde i detta specifika typfall?”.¹² Att behandla uppkomsten som ett ontologiskt faktum döljer de reella intresseavvägningarna.

Ett liknande mönster framträder i Henrik Gams analys av timeshare-institutet. Detta fenomen lät sig svårligen placeras i de traditionella kategorierna, då det utgjorde en hybrid av bruksrätt, äganderätt, bolagsstrukturer och avtalsarrangemang. Försök att tvinga in timeshare i den binära strukturen ”hyra eller ägande” riskerade att dölja de specifika risker och skyddsbehov som konstruktionen medförde. Lösningen låg inte i att fastställa timeshare-institutets ”natur”, utan i att reglera de specifika rättigheter och skyldigheter som situationen krävde.¹³

Relevansen för AI-debatten är påtaglig. Precis som insolvensrätten och fastighetsrätten har tvingats överge sökandet efter begreppens väsen till förmån för funktionella analyser, måste AI-rätten göra detsamma. Oskar Mossberg har i sin analys av den konstruktiva riktningen påpekat att begreppet ”begreppsjurisprudens” ofta använts pejorativt för att avfärda motståndare, men att kärnfrågan handlar om huruvida begreppen tillåts styra lösningen eller om de används som verktyg för att bearbeta det positiva rättsstoffet.¹⁴

I AI-debatten fungerar dikotomin ”subjekt/objekt” som en metodologisk återvändsgränd. Genom att ställa frågan om AI är ett subjekt, riskerar vi att importera en hel katalog av rättsföljder (ansvar, rättigheter och straffbarhet) som inte är anpassade för den tekniska verkligheten, precis som en felaktig klassificering av en fordran eller ett timeshare-avtal leder till materiellt felaktiga resultat. Vår kritik riktar sig därför mot själva utgångspunkten, vi bör inte fråga vad AI är, utan hur risker och ansvar ska fördelas i de komplexa handlingskedjor där tekniken ingår.

4 Ross och de tomma mellanleden

Vi har redan berört hur den skandinaviska funktionalismen betraktar juridiska begrepp som verktyg snarare än avbilder av verkligheten. För att fullt ut förstå hur denna insikt kan demaskera dagens AI-debatt måste vi dock gå tillbaka till källan: Alf Ross analys av begreppet Tû-Tû. Det var här den funktionalistiska kritiken mot metafysiska begrepp fick sin

¹² Jonatan Schytzer, *”Fordrans uppkomst inom insolvensrätten”*, 2020, s. 27–28, 32.

¹³ Henrik Gam, *Timeshare*, 1998, s. 9–10.

¹⁴ Oskar Mossberg, *”Den konstruktiva riktningen och ”begreppsjurisprudensen”*, 2022, s. 289 ff.

skarpaste formulering och analogin erbjuder en kraftfull lins för att se igenom dagens tekniska mystik.

Genom att studera hur den fiktiva Noît-cif-stammen använde ordet tû-tû demonstrerade Ross hur juridiska begrepp ofta fungerar som semantiskt tomma mellanled, en presentationsteknik för att systematisera relationen mellan rättsfakta (tabubrott) och rättsföljder (reningsceremoni) utan att begreppet i sig korresponderar mot en verklig entitet eller egenskap i sinnevärlden.¹⁵ Begreppet fyller en funktion för att skapa överskådlighet i komplexa regler, men det saknar egen kausal kraft.

Ross poäng var inte att förkasta juridiska begrepp som ”äganderätt” eller ”fordran”, utan att demystifiera dem. De ska förstås som verktyg för att kommunicera resultatet av en rättslig analys, inte som metafysiska objekt.¹⁶ Denna insikt är direkt tillämpbar på AI-diskursen. När rättsvetenskapen frågar om AI är ett subjekt eller objekt, riskerar vi att begå samma misstag som stammen: Vi behandlar ett tekniskt presentationsbegrepp som en ontologisk realitet.

Faran uppstår när vi tillskriver mellanledet egen agens. I stället för att fastna i frågan om ”AI-systemets” natur, bör begreppet behandlas som en språklig behållare för att strukturera de reella ansvarsrelationerna mellan utvecklare, data, algoritmer och de risker dessa ger upphov till. Precis som ”äganderätt” endast är en förkortning för en komplex väv av rättsregler mellan människor, är ”AI” en förkortning för en komplex väv av tekniska instruktioner och mänskliga beslut som ger upphov till rättsligt relevanta risker.¹⁷

5 Reductio I: Televisionens paradox

Anta att vi, i likhet med modern teknikreglering, definierar AI genom systemets funktionella¹⁸ egenskaper snarare än genom metafysiska kriterier som medvetande. Enligt EU:s AI-förordning är ett AI-system ett maskinbaserat system konstruerat för att fungera med varierande grad av autonomi, som kan uppvisa anpassningsförmåga efter utplacering och som, för explicita eller implicita mål, härleder hur output ska genereras från den input det mottar.¹⁹

Om vi applicerar AI-förordningens funktionella definition strikt uppstår ett störande motexempel: den moderna TV-apparaten. Den tar emot signaler (input), bearbetar dem genom avancerade interna processer för signalbehandling och kodtolkning som är dolda för

¹⁵ Ross, 1957, s. 468 ff.

¹⁶ Martinson, 2014, s. 17.

¹⁷ Jfr Faber och Martinson, 2019, s. 93.

¹⁸ Termen ”funktionella” avser här lagstiftningstekniken att definiera AI genom dess operationella processer input-output. Detta ska inte förväxlas med den ”skandinaviska funktionalismen” som utgör denna uppsats metodologiska lins. Paradoxen som målas upp nedan är att lagstiftarens ”funktionella” definition ironiskt nog riskerar att leda till den typ av begreppslåsning som den juridiska funktionalismen varnar för.

¹⁹ Europaparlamentets och rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens (AI-förordningen), EUT 2024 L 1689, art. 3(1). Stycket beskriver en förenklad tolkning av rekvisiten som uppställs i artikeln. Detta är något semantiskt förenklat för att tydliggöra poängen.

användaren och genererar audiovisuell output för att uppfylla sitt funktionsmål. Systemet opererar autonomt inom sina parametrar för att optimera bild och ljud. Trots att TV:n uppfyller den grundläggande funktionella strukturen input – process – output, uppstår ingen diskussion om dess rättssubjektivitet eller normativa status. Ingen hävdar att TV:n bör hållas ansvarig för de bilder den visar.

Paradoxen blottlägger att en funktionell beskrivning av hur en maskin bearbetar information inte är tillräcklig för att avgöra om den ska bära rättsligt ansvar. Att TV:n och AI-systemet delar samma funktionella grundstruktur men behandlas fundamentalt olika juridiskt visar att definitionen i sig saknar normativ bärkraft. Det är inte funktionen som avgör statusen, utan vår sociala och juridiska relation till artefakten.²⁰

Svante Bergströms kritik av begreppsjurisprudensen är här direkt tillämplig, precis som jurister historiskt har ”betraktat gränslinjerna mellan avtalstyper såsom mera ’rena’ än de i själva verket är”, behandlar dagens AI-diskurs gränslinjen mellan ”vanlig teknik” (TV:n) och ”AI” som skarpare än den är.²¹ En funktionsdefinition expanderar visserligen kategorin av vad som kan regleras, men den tillför inte den normativa precision som krävs för att motivera varför vissa maskiner ska betraktas som rättssubjekt medan andra förblir möbler.

6 Reductio II: Turingtestets villfarelse

Turingtestet har ofta fungerat som kulturell och intellektuell genväg för ”maskinen som subjekt”. Men testets logik är performativ och epistemisk. Det handlar om huruvida en observatör kan skilja maskinens svar från en människas i en begränsad interaktionskanal. Det är varken ett test för intention, moralisk förmåga eller normativ förståelse.²²

Om vi låter denna logik styra den rättsliga statusen hamnar vi i ett farligt cirkelresonemang. I takt med att språkmodeller blir bättre på att simulera mänskliga skälstrukturer (”den låter som en jurist”) ökar risken att vi förväxlar språklig plausibilitet med normativt ansvarsbärande. Men att en maskin kan producera en text som är omöjlig att skilja från ett mänskligt avtalsförslag innebär inte att maskinen kan bära det juridiska ansvaret för avtalet. Den avgörande distinktionen går mellan språklig imitation och normativ förståelse. Förmågan att manipulera symboler enligt ett program är inte detsamma som att delta i en normativ praktik där skäl kan åberopas, ifrågasättas och sanktioneras.²³

Detta belyser kärnan i Ross kritik av metafysiska begrepp. Precis som stammens medlemmar i Ross allegori kunde förväxla ordet tû-tû med en verklig kraft, riskerar AI-diskursen att

²⁰ Jfr Ross, 1957, s. 468 ff. om begrepp som tomma mellanled utan semantisk referens.

²¹ Bergström, 1969, s. 313.

²² Alan M. Turing, ”Computing Machinery and Intelligence” (1950) 59 Mind 433; J.fr. John R. Searle, ”Minds, Brains, and Programs” (1980) 3 Behavioral and Brain Sciences 417. Luciano Floridi, ”Turing’s Three Philosophical Lessons” (2012) s.21.

²³ Searle, 1980, s. 417 ff.

förväxla systemens språkliga output med normativ handlingsförmåga. När en AI-modell genererar en text är det resultatet av statistisk optimering, inte en normativ handling. Att tillskriva systemet subjektivitet baserat på dess förmåga att imitera mänskligt språk är att projicera egenskaper på verktyget baserat på vår egen oförmåga att skilja kopian från originalet.²⁴ Juridiken riskerar därmed att reagera på en kulturell föreställning om den ”talande maskinen” snarare än på den faktiska ansvarsstrukturen.

7 Black-box-argumentets begränsningar

Den återkommande invändningen mot att behandla AI som en konventionell artefakt är det så kallade ”*black box*”-problemet. Argumentet bygger på att moderna AI-system, särskilt de som baseras på djupinlärning, är epistemiskt ogenomträngliga; vi kan inte rekonstruera exakt varför systemet kom fram till ett visst resultat. Vissa rättsvetare drar ur detta slutsatsen att befintliga juridiska kategorier bryter samman. Yavar Bathaee hävdar exempelvis att eftersom vi inte kan spåra beslutsvägen genom modellens interna matematik, så ”upphör befintliga juridiska doktriner att fungera” (*”cease to function”*), särskilt de som kräver uppsåt eller kausalitet.²⁵ På liknande sätt beskriver Stefan Larsson svårigheten att härleda träningsdata och förstå modellens interna logik som en central utmaning för ansvarsutkrävande och transparens.²⁶

Detta resonemang bygger dock på ett metodologiskt felslut: Antagandet att juridiskt ansvar förutsätter total teknisk insyn i kausalkedjan. Om vi applicerar detta strikt skulle en biltillverkare undgå ansvar för en defekt bromsmekanism så länge den interna processen var tillräckligt komplex för att vara ogenomtränglig för en lekman. Bartosz Brożek och Jerzy Stelmach har övertygande argumenterat för att *”the opacity issue is not a genuine problem”* i den meningen att vi dagligen tillskriver ansvar i system vi inte fullt ut förstår, inte minst det mänskliga psyket, som är den ultimata ”svarta lådan”. Det vi uppfattar som ett ontologiskt särdrag hos AI är i själva verket en psykologisk reaktion på det obekanta, ett *”strangeness problem”* snarare än ett juridiskt hinder.²⁷

Vidare är opacitet ofta ett resultat av medvetna designval snarare än en teknisk naturlag. Cynthia Rudin har visat att det i många fall, även vid högriskbeslut, är möjligt att konstruera *”inherently interpretable models”* som presterar likaväl som *”black box”*-modeller. Att välja en ogenomskinlig modell är därmed ett mänskligt beslut, inte en oundviklighet.²⁸

²⁴ Jfr Bartosz Brożek m.fl., *”The Black Box Problem Revisited”*, 2023, s. 431 om *”the strangeness problem”*.

²⁵ Yavar Bathaee, *”The Artificial Intelligence Black Box and the Failure of Intent and Causation”*, 2018, s. 906–907.

²⁶ Stefan Larsson, *”Mellan juridisk fixering och flexibilitet i den europeiska AI-förordningen”*, 2024, s. 287.

²⁷ Brożek m.fl., 2023, s. 430–431.

²⁸ Cynthia Rudin, *”Please Stop Explaining Black Box Models for High Stakes Decisions”*, 2019, s. 1 ff.; se även Doshi-Velez & Kim, *”Towards a Rigorous Science of Interpretable Machine Learning”*, 2017, s. 2 för en diskussion om tolkningsbarhet.

Ur ett funktionalistiskt perspektiv måste därför Bathaees slutsats förkastas. Juridiskt uppsåt lokaliseras inte i modellens interna viktparametrar, utan i de mänskliga besluten att designa, parametrisera och driftsätta systemet. Att hävda att ansvarsdoktriner upphör att fungera för att verktyget är komplext är att låta teknisk opacitet diktera normativ ansvarsfrihet. Opacitet är inte en egenskap som ger AI särställning; det är en riskfaktor som bör skärpa, snarare än upplösa, kravet på mänsklig kontroll och aktsamhet.

8 AI Förordningen som reproducerad metodfälla

EU:s AI-förordning marknadsförs som en funktionell och riskbaserad lagstiftning. Tanken är att reglerna ska anpassas efter den skada ett system kan orsaka, snarare än tekniken i sig. En analys av förordningens struktur visar dock att den reproducerar det metodologiska felslut som funktionalismen varnar för, den låter statistiska begreppskategorier diktera normativa lösningar.

Detta blir tydligt om vi studerar lagens konstruktion. Förordningen definierar först systemets status i artiklarna 5 och 6 (är det förbjudet? är det högrisk?). Först därefter, om systemet hamnar i rätt fack, aktiveras skyldigheterna för leverantörer och användare i artiklarna 16–27. Ordningen är alltså låst: Klassificeringen av artefakten är en förutsättning för ansvarsutkrävandet.²⁹

Denna ordning inverterar logiken i traditionell skadeståndsrätt. Jämför med produktansvarslagen (1992:18). Där aktiveras skyldigheterna av skadan och den bristande säkerheten, inte av vilken produktkategori föremålet tillhör. Frågan är inte, vilken typ av produkt är detta? Utan har en skada uppstått och vem kontrollerade den? AI-förordningen flyttar fokus från den skadeorsakande handlingen till systemets etikett. Daniel Braun har påpekat att detta bygger på det felaktiga a priori-antagandet att en viss uppsättning tekniker ("AI") i sig själva bär på risken, vilket leder till att regleringen missar målet när tekniken ändrar form.³⁰

Ett tydligt exempel på det "taktproblem" detta skapar är hur generativ AI tvingade fram en panikartad omreglering sent i processen eftersom tekniken inte passade i de färdiga facken.³¹ Lagstiftaren tvingas jaga tekniken med nya definitioner i stället för att reglera de bakomliggande intressekonflikterna.

Svante Bergströms analys av lex commissoria-förbudet i 37 § 2st. avtalslagen visar på det funktionalistiska alternativet. Lagstiftaren valde där att reglera själva handlingen (att förverka säkerhet) oavsett om parterna kallade det "pant" eller "säkerhetsöverlåtelse". Lagstiftaren

²⁹ Europaparlamentets och rådets förordning (EU) 2024/1689; jfr Martin Ebers, "Truly Risk-Based Regulation...", 2024, s. 688 som kritiserar att förordningen är produktbaserad snarare än riskbaserad.

³⁰ Braun, s. 2–3.

³¹ Larsson, 2024, s. 287 ff.

riktade in sig på funktionen (kredit mot säkerhet) och förbjöd missbruket, i stället för att fastna i en klassificering av avtalstypen.³² AI-förordningen gör tvärtom, den reglerar "AI-systemet" (formen) i stället för den riskfyllda behandlingen (funktionen), vilket gör att samma riskfyllda handling kan vara tillåten om den utförs med "dum" teknik men förbjuden om den utförs med AI.

Resultatet är att begreppet "AI-system" blir "*an alien lump, undigested*" en begreppslik klump som stör den juridiska matsmältningen genom att tvinga samman vitt skilda problem under en och samma etikett.³³ Begreppet ges självständig normativ tyngd, trots att det egentligen bara borde vara en sammanfattande beteckning.³⁴

9 Den epistemiska förklaringen: Miniräknarens allegori

Förklaringen till varför subjekt- och agensspråket uppstår i teknologiska sammanhang är epistemisk snarare än ontologisk. När de tekniska kausalkedjorna är synliga för oss framstår systemen som vad de är, mänskligt konstruerade kontrollmekanismer. När samma kedjor är dolda eller alltför komplexa för att överblicka, aktiveras en psykologisk mekanism där vi fyller kunskapsluckan med agens. Bartosz Brożek och Jerzy Stelmach beskriver detta som ett "*strangeness problem*", vi tenderar att tillskriva mänskliga egenskaper till artefakter som uppvisar ett beteende vi inte omedelbart kan härleda till en mekanisk orsak.³⁵

Fenomenet illustreras tydligt i diskussionen om så kallade "smarta kontrakt". Forskare som Kevin Werbach och Nicolas Cornell har argumenterat för att algoritmiska kontrakt kan utgöra nya avtalskategorier där koden agerar självständigt. Detta bygger dock på en begreppslik illusion där automatiserade exekveringsmekanismer misstas för normativ autonomi.³⁶ Oskar Mossberg har med hänvisning till Hagerups konstruktiva metod visat att juridisk analys måste ta sin utgångspunkt i "*selve det positive retsstof*" och inte i intuitiva reaktioner på nya fenomen.³⁷ Att intuitivt uppfatta koden som en "aktör" är en psykologisk reaktion, inte en rättskällemässig analys.

Situationen kan liknas vid en miniräknare placerad i ett samhälle utan tidigare teknologisk referensram. Utan förståelse för beräkningsprocessen skulle apparaten möjligen uppfattas som ett fenomen i särställning och dess förmåga att leverera korrekta svar skulle kunna tolkas som en form av insikt eller vilja. Kunskapsbristen hos betraktaren förändrar dock inte vad

³² Bergström, 1969, s. 318.

³³ Karl N. Llewellyn, *The Common Law Tradition*, 1960, s. 93.

³⁴ Jfr Jakob Heidbrink, Föreläsningar i fördjupningskursen "Avancerad kontraktsrätt" VT 2026, i diskussion om hur begreppet "obehörig vinst" ofta fungerar på detta sätt: Som ett skenbart förklarande mellanled som ger intryck av en sammanhållen princip, trots att lösningarna i praktiken hämtas från andra rättsliga strukturer. På samma sätt riskerar "AI" att bli en etikett som döljer analysen snarare än förklarar den.

³⁵ Brożek m.fl., 2023, s. 431.

³⁶ Eliza Mik, "Smart Contracts: Terminology...", 2017, s. 269.

³⁷ Mossberg, 2022, s. 289.

artefakten är, en uppsättning logiska grindar utan endast hur den tolkas. Alf Ross poäng är här högst tillämplig, stammens förklaringar till tû-tû avslöjade mer om stammens föreställningsvärld än om verkligheten.³⁸

Den avgörande slutsatsen är att en epistemisk begränsning inte får leda till en ontologisk omdefiniering. Att vi inte kan se in i ett neuralt nätverk (epistemik) är inte ett argument för att omklassificera systemet från verktyg till rättssubjekt (ontologi). Genom att använda agensbegrepp för att beskriva komplex kausalitet skapar vi en begreppslik dimma där vi, i strid med Hagerups metodideal, låter vår intuition styra analysen snarare än det positiva rättsstoffet.

10 Konsekvenserna av felorienterad reglering

När normativ uppmärksamhet riktas mot systemet i stället för människan sker en ansvarsförskjutning. Utvecklare, vidareanvändare (*deployers*) och institutioner riskerar att uppfatta ”systemet” som den primära riskbäraren.³⁹ Resultatet blir en ansvarsdiffusion där mänskliga beslut gällande design och implementering framstår som tekniska nödvändigheter snarare än normativa val.

Detta är regulatoriskt farligare än att inte reglera alls, eftersom det skapar en illusion av kontroll samtidigt som den normativa kedjan upplöses. Claes Martinson har, med hänvisning till Sjur Brækhus, beskrivit hur juridiska begrepp som äganderätten kan ha ”en stark makt över tanken” (*a strong influence over the mind*). Denna begreppslika makt riskerar att leda till att jurister förbiser de reella intressekonflikterna till förmån för en logik dikterad av begreppet självt.⁴⁰

Samma fenomen riskerar att ske i AI-regleringen, begreppet ”AI-system” får en makt över tanken som förskjuter analysen från de reella ansvarsrelationerna. Klassificeringen producerar därmed en normativ illusion. Den får det att framstå som om risk och påverkan emanerar ur artefakten, snarare än ur de strukturer av designval, parametrisering och användning där normativ kontroll faktiskt kan utövas. Det är inte bara en begreppslik förskjutning utan en regleringsmässig sådan, ansvarsdimensionen försvagas när artefakten behandlas som analysens primära enhet.

³⁸ Ross, 1957, s. 812 ff.

³⁹ Termen *deployer* översätts i den officiella svenska versionen av AI-förordningen till ”användare”, vilket riskerar att skapa begreppsförvirring gentemot den enskilda individ som utsätts för systemet. För att undvika denna otydlighet används här Stefan Larssons term ”vidareanvändare” för att beteckna den aktör, exempelvis en myndighet eller ett företag som implementerar och använder ett AI-system i sin verksamhet under eget ansvar, till skillnad från den drabbade ”slutanvändaren” eller fysiska personen; se Stefan Larsson, ”*Mellan juridisk fixering och flexibilitet i den europeiska AI-förordningen*”, 2024, s. 288.

⁴⁰ Martinson, 2014, s. 24.

11 En alternativ orientering

Vår tes är inte endast att AI-regleringen bör justeras, utan att själva premissen för en separat ”AI-rätt” är metodologiskt tveksam. I stället för att försöka definiera och reglera tekniken som en egen ontologisk kategori, bör den normativa orienteringen utgå från principen om teknikneutralt aktörsansvar. Juridikens uppgift är att identifiera de subjekt som har faktisk kontroll över händelseförloppet, oavsett om verktyget de använder är en hammare, en algoritm eller en organisation.

Detta synsätt är redan inbyggt i centrala delar av rättsordningen. Produktansvarslagen (1992:18) tillhandahåller exempelvis en struktur för att hantera skada orsakad av defekta produkter utan att behöva besjåla produkten. Frågan är inte om AI är en ”produkt” i metafysisk mening, utan vilka aktörer, tillverkare, importörer och distributörer som har kontroll över de egenskaper som orsakar skada. Att placera en AI-komponent i en gräsklippare förändrar inte ansvarskedjan; det specificerar endast var i kedjan kontrollen över den skadebringande egenskapen ligger.⁴¹

På samma sätt hanterar Dataskyddsförordningen (GDPR) automatiserat beslutsfattande i artikel 22 utan att förutsätta en separat AI-personlighet. Fokus ligger strikt på den personuppgiftsansvariges skyldigheter och den registrerades rättigheter. Lagstiftaren reglerar här processen (automatiseringen) och ansvaret (hos den juridiska personen), med administrativa sanktioner och skadeståndsansvar direkt knutna till den personuppgiftsansvariges kontroll snarare än verktygets ”tankar”.⁴²

Inom immaterialrätten blir behovet av denna orientering kanske tydligast. Att ställa frågan ”kan AI ha upphovsrätt?” är som att fråga om Eulers formel kan ha patent på sina lösningar. Formeln tar in variabler och producerar output; AI-system gör samma sak med fler dimensioner och större datamängder. Det finns inget kreativt val i den tekniska processen, endast transformation enligt parametrar som människor definierat. Att ett resultat ser ut som något en människa kunde ha skapat förändrar inte dess natur. Den normativa konsekvensen är tydlig: Det är människan som använder verktyget, som väljer, kombinerar och modifierar som är det enda tänkbara rättssubjektet. Att tillskriva verktyget skapandeförmåga är att förväxla beräkningskraft med intention.

Poängen med dessa exempel är inte att befintliga regelverk är perfekta, utan att de redan innehåller den normativa struktur som krävs, ansvarsallokering baserad på kontroll. Att skapa en parallell ”AI-rätt” där systemet görs till primär analysenhet riskerar att fragmentera dessa ansvarskedjor och skapa regulatoriska hål där aktörer kan hävda att ”systemet”, inte de själva,

⁴¹ Jfr Direktiv (EU) 2024/2853 om skadeståndsansvar för produkter med säkerhetsbrister, som moderniserar detta ansvar för digitala produkter utan att överge aktörsfokus.

⁴² Europaparlamentets och rådets förordning (EU) 2016/679, art. 22; jfr Sandra Wachter m.fl., ”Why a Right to Explanation...”, 2017, s. 76 ff. om hur fokus ligger på mänsklig intervention och ansvar.

bar ansvaret. Rättens uppgift är att peka ut det mänskliga subjektet bakom den tekniska ridån, inte att ge ridån juridisk status.

12 Avslutning

Vår analys visar att ontologiseringen av AI inte bara är onödig, utan utgör ett metodologiskt felslut. AI är inte ett rättssubjekt i vardande, utan en statistisk process inbäddad i mänskliga maktstrukturer. Den funktionalistiska traditionen erbjuder här mer än bara ett perspektiv; den tillhandahåller en nödvändig demystifiering. Faber och Martinson har visat att begrepp som äganderätt måste förstås relationellt, de används för att lösa en specifik intressekonflikt och strukturera hur ansvar och rättsföljder knyts till aktörer, inte för att fastställa en absolut ontologisk status som per automatik kan överföras från en situation till en annan.⁴³ Samma logik måste tillämpas på begreppet "AI-system": Det är ett begrepp som bör användas för att strukturera ansvarsförhållanden, inte ett ontologiskt objekt med inneboende normativa egenskaper.

Denna insikt har sin historiska förankring i Ross analys av Tû-Tû. Han visade som tidigare nämnt att juridiska begrepp fungerar som tomma mellanled utan egen semantisk referens, men med en viktig funktion att strukturera rättsfakta och rättsföljder.⁴⁴ Faran i dagens AI-diskurs är att vi, likt stammens medlemmar, har börjat tro att "AI" refererar till en magisk kraft i maskinen snarare än till en uppsättning tekniska processer. Detta leder till att vi felaktigt söker svaret på ansvarsfrågor i artefaktens status ("är den ett subjekt?") i stället för i de mänskliga handlingskedjorna.

Att inrätta en separat "AI-rätt" där systemet görs till analysens primära enhet är därför att institutionalisera detta metodfel. Det riskerar att skapa en ansvarsdiffusion där juridiken jagar tekniska skuggor medan de faktiska riskerna lämnas oreglerade. Lösningen är i stället att tillämpa en teknikneutral ansvarsrätt där fokus ligger på kontroll och riskallokering. Produktansvar, marknadsföringsrätt och dataskyddsregler tillhandahåller redan de strukturer som krävs för att hålla mänskliga aktörer ansvariga för de verktyg de sätter i drift.

Detta innebär inte att rätten är statisk eller att lagstiftningen inte behöver uppdateras i takt med teknik- och samhällsutvecklingen. Specifika justeringar kommer att krävas för att hantera nya fenomen, exempelvis skärpta krav på transparens kring träningsdata eller tydligare regler för hur upphovsrättsligt material får processas.⁴⁵ Men sådana uppdateringar bör ske inom ramen för befintliga rättsliga strukturer och rikta sig mot de mänskliga aktörernas beteenden, inte genom att tillskriva mjukvaran en ny ontologisk status.

⁴³ Faber och Martinson, 2019, s. 92 och 95)

⁴⁴ Ross, 1957, s. 468 ff.

⁴⁵ Jfr Kasia Söderlund, "AI Transparency in Trustworthy AI...", 2025, s. 16–17 om behovet av transparens för ansvarsutkrävande.

Frågan ”är AI subjekt eller objekt?” är därmed samma typ av felställd fråga som att fråga om tû-tû är farligt. Svaret är att frågan saknar vetenskaplig mening. Den normativa relevansen ligger inte i begreppets referens utan i de relationer det strukturerar. I AI-sammanhang innebär detta relationer mellan utvecklare, vidareanvändare och drabbade tredje män. Det är i dessa mänskliga relationer, inte i artefaktens klassificering som juridiken finner sitt enda hållbara fäste.

Bibliography of References / Källförteckning

Bathae, Yavar, 'The Artificial Intelligence Black Box and the Failure of Intent and Causation' (2018) 31 Harvard Journal of Law & Technology 889.

Bergström, Svante, Pantavtal eller köp? (Norstedts 1969).

Braun, Daniel, 'Why "Artificial Intelligence" Should Not Be Regulated' (2025) SSRN Working Paper.

Brożek, Bartosz, Furman, Michał, Jakubiec, Marek och Kucharzyk, Bartłomiej, 'The Black Box Problem Revisited: Real and Imaginary Challenges for Automated Legal Decision Making' (2023) 32 Artificial Intelligence and Law 427.

Doshi-Velez, Finale och Kim, Been, 'Towards a Rigorous Science of Interpretable Machine Learning' (2017) arXiv:1702.08608.

Ebers, Martin, 'Truly Risk-Based Regulation of Artificial Intelligence' (2024) European Journal of Risk Regulation 686.

Faber, Wolfgang och Martinson, Claes, 'Can Ownership Limit the Effectiveness of EU Consumer Contract Law Directives?' (2019) 15 European Review of Contract Law 91.

Floridi, Luciano, 'Turing's Three Philosophical Lessons and the Philosophy of Information' (2012) 14 Information 21.

Gam, Henrik, Timeshare (Juristförlaget 1998).

Larsson, Stefan, 'Mellan juridisk fixering och flexibilitet i den europeiska AI-förordningen' (2024) Juridisk Tidskrift 287.

Llewellyn, Karl Nickerson, The Common Law Tradition (Little, Brown & Co 1960).

Martinson, Claes, 'The Scandinavian Approach to Property Law: A Functionalist View' (2014) 3 Scandinavian Studies in Law 17.

Mik, Eliza, 'Smart Contracts: Terminology, Technical Limitations and Real World Complexity' (2017) 9 Law, Innovation and Technology 269.

Mossberg, Oskar, 'Den konstruktiva riktningen och "begreppsjurisprudensen"' (2022) 135 Tidsskrift for Rettsvitenskap 289.

Ross, Alf, 'Tû-Tû' (1957) 70 Harvard Law Review 812.

Rudin, Cynthia, 'Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead' (2019) 1 Nature Machine Intelligence 206.

Schytzer, Jonatan, Fordrans uppkomst inom insolvensrätten (Jure 2020).

Searle, John Rogers, 'Minds, Brains, and Programs' (1980) 3 Behavioral and Brain Sciences 417.

Söderlund, Kasia, *"AI Transparency in Trustworthy AI: From Metaphor to Governance Tool in EU Technology Regulation"* (Lund University 2025).

Turing, Alan Mathison, 'Computing Machinery and Intelligence' (1950) 59 Mind 433.

Wachter, Sandra, Mittelstadt, Brent och Floridi, Luciano, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 76.

Föreläsningar och opublicerat material

Bruncevic M, föreläsningar, fördjupningskursen *"AI och rätten"*, Handelshögskolan vid Göteborgs universitet (VT 2026).

Heidbrink, Jakob, föreläsningar, fördjupningskursen "*Avancerad kontraktsrätt*",
Handelshögskolan vid Göteborgs universitet (VT 2026).