

Generativ AI, etik och immaterialrätt

Det förekommer en hel del samhällseliga frågor och etiska frågeställningar som pockar på en lösning i samband med generativ artificiell intelligens. Det immaterialrättsliga läget är också i många fall oklart och under utveckling då det gäller användningen av generativ AI.

Särskilt under det senaste året har artificiell intelligens, AI, varit på allas läppar. Det betyder inte nödvändigtvis att vi talar om samma sak. Till AI kopplas därtill både förhoppningar och rädslor. Då ett företag tar i bruk AI hoppas man få en konkurrensfördel, medan en arbetstagare kan vara rädd för att förlora sin arbetsplats.

Det förekommer följaktligen olika uppfattningar om vad AI är och vilken verkan den har varför det är utmanande att ge en uttömmande definition.¹ Ett förslag är att tala om att med hjälp av en dator automatisera verksamhet som framstår som intelligent. Ett annat att "artificiell intelligens är en metafor med hjälp av vilken ett datorprogram jämförs med mänsklig intelligent verksamhet. Det är egentligen inte frågan om intelligens utan någon slags jämförelse med intelligens." Experter tenderar att avdramatisera AI, såsom i exemplen ovan, medan medier gärna underblåser överdrifter.²

Jämförelsen med mänsklig intelligens känns utan tvivel kittlande. Användningen av benämningen AI utan en närmare precisering har bidragit till "hopen" kring företeelsen.

Min tanke är likväl inte att diskutera

Max Oker-Blom

är ekon.dr och jur.lic.
samt associerad
forskare vid Hanken.



”I dagsläget är det oklart om domstolar kan stoppa ett generativt program, som utan upphovsrättsinnehavarnas samtycke tränats med immaterialrättsligt skyddat material.”

och analysera själva fenomenet. Avsikten här är att betrakta de immaterialrättsliga utmaningar ett delområde eller en kategori inom AI, nämligen generativ AI, ställs inför. Såsom övriga kategorier kan nämnas expertsystem, maskininlärning och djup inlärning.³

Viktiga frågor såsom hur generativ AI inverkar på våra arbetsplatser och på miljön – träning av generativ AI är energislukande

– förbigås här.⁴ Den senare frågan har för övrigt berörts överraskande litet i offentligheten.

Generativ AI och IPR

Generativ AI är en form av AI som kan skapa nytt innehåll som svar på en användares uppmaning eller instruktion ("prompt"), till exempel en beskrivning av ett önskat utfall. Det kan röra sig om text, datakod, bild, ljud (musik) eller video. Som exempel på generativa AI-verktyg kan nämnas Chat GPT, Midjourney, Copilot och Firefly.

Litet mer precist sagt är generativ AI ett sannolikhetsystem som letar efter samband. Till exempel om man modifierar ett fotografi analyserar systemet ett massivt bibliotek med tidigare fotografier och bestämmer om det noterar en vit pixel vilken sannolikheten är att följande pixel närmast den första pixeln är vit, svart, grå eller någon annan färg. Med hjälp av detta koncept kan man konstruera helt nya fotografier, noveller, datakod, olika slag av ansikts- och röstförfalskningar ("deep fakes"), etcetera.⁵

Generativa AI-verktyg tränas med enorma mängder data, vilka kan innehålla miljontals sidor text eller bilder. Beroende på den som utvecklar verktyget består träningsmaterialet av fritt tillgänglig information eller data, så kallad ren data ("unencumbered pure data"), skyddade data, såsom upphovsrättsskyddade verk eller en blandning av dessa. Det är här immaterialrätten kommer in i bilden.

Enligt en vedertagen ekonomisk uppfattning är avsikten med att skydda immateriella rättigheter ("Intellectual property rights, IPR") att skapa incitament för innovativ eller skapande verksamhet och innovativt arbete i samhället.⁶ Genom att ge uppfinnaren eller konstnären en ensamrätt, som grundar sig på lag, ger samhället också rätts-

medel för att skydda dessa från andra, det vill säga närmast konkurrenter. Härigenom kan rättighetens innehavare förhindra andra att utnyttja uppfinningen eller det upphovsrättsskyddade verket. Ur ekonomisk synvinkel får alltså innehavaren ett temporärt 'monopol', som har ett visst skyddsomfång och är i kraft en viss tid, för sin uppfinning eller sitt konstverk.

Då de generativa AI-verktygen tränas med material eller data, som är immaterialrättsligt skyddat, kan det uppstå immaterialrättsintrång. Dessa fall handlar för det mesta om upphovsrätt och varumärken, men också andra immaterialrätter, såsom mönsterskydd⁷, databas- och patenträttigheter kan beröras.

I det följande diskuterar jag en del av dessa fenomen litet närmare.⁸

Immaterialrättsliga utmaningar 1. Träningsmaterialets status

För att kunna träna generativa AI-verktyg behöver således utvecklaren av verktyget samtycke av den, som har upphovsrätt till träningsmaterialet. Producenter av generativ AI har därför slutit avtal med till exempel mediehus för att ha tillgång till ett så stort material som möjligt. Om en producent trots allt låtit bli att inhämta samtycke av upphovsrättsinnehavaren eller -innehavarna finns det självfallet en risk för rättstvister. Många rättstvister har också de facto ägt rum. Under det kommande året förväntas domstolsavgöranden i USA om bland annat i vilken utsträckning skyddat material kan användas för att falla inom ramen för det amerikanska undantaget "fair use", med andra ord vad som kan betraktas som skäligt.⁹

Ett tilläggsproblem är att upphovsrätt till ett verk inte kräver registrering för att uppstå. Det uppkommer automatiskt i och med



Nyligen hävdade skådespelaren Scarlett Johansson att Open AI utnyttjat hennes röst utan hennes samtycke, vilket Open AI:s vd Sam Altman dementerat. Foto: Gage Skidmore

att verket, till exempel texten eller konstverket, bedöms vara originellt och individuellt. I USA krävs visserligen registrering om innehavaren avser tvista om sin rättighet. Det är alltså inte alltid självklart att man har att göra med ett skyddat verk.

Som ett motdrag till Open AI:s relativt färskaste påståenden om att det är omöjligt att träna en utvecklad AI-modell utan upphovsrättskyddat material har man nyligen publicerat ett upphovsrättsvänligt alternativ, nämligen Common Corpus, som utmanar denna uppfattning.¹⁰

Common Corpus är resultatet av internationellt samarbete och publicerades den 20 mars 2024. Det är världens största öppna datasamling för att kunna träna stora språkmodeller (Large language models, LLM). Det består enligt uppgift av 500 miljarder ord och innehåller miljoner böcker och

övriga publikationer. Det är för närvarande världens största öppna engelskspråkiga data set med 180 miljarder ord. Som drivkraft för projektet fungerar det franska start-up företaget Pleias. Grundarnas tanke är att skapa konkurrens med företag som Open AI.

Det är inte enbart utvecklaren av AI-verktyget, som bör beakta innebörden av immaterialrätt, närmast upphovsrätt, utan också det enskilda företaget, det vill säga användaren. Här handlar det om att värna om de uppmaningar företaget ger applikationen. Det finns nämligen en risk att immaterialrättsligt skyddad information på det sättet integreras i den modell producenten ställer till företagets förfogande och senare hamnar i händerna på en konkurrent. Risken att detta skall ske med konfidentiell information är likaså stor om företaget inte har introducerat vederbörliga kontroller.

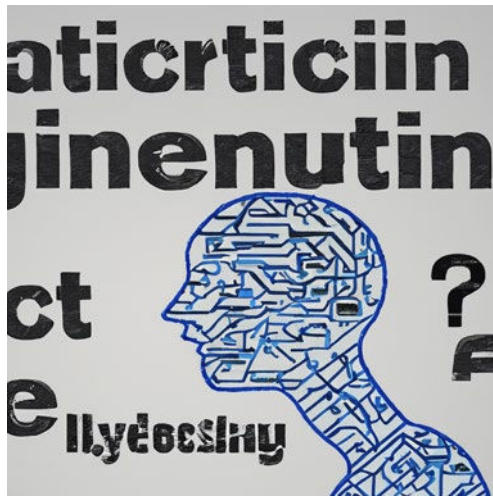
En ytterligare utmaning handlar om att användarföretaget möjligen kan bli ansvarigt för immaterialrättskränkningar, som är ett resultat av att den generativa AI-modellen tränats på skyddat material, trots eventuella utfästelser om motsatsen, som i och för sig inte härstammar från det aktuella företaget.

I dagsläget är det också oklart om domstolar kan utfärda en interimistisk dom för att stoppa ett visst generativt program, som utan upphovsrättsinnehavarnas samtycke tränats med immaterialrättsligt skyddat material.

Hur domstolarna förhåller sig till olika undantag, såsom det ovan nämnda "fair use", tillfällig kopiering, etcetera, är också oklart. Det är dessutom skäl att notera att domstolsutslag kan variera beroende på jurisdiktion och de faktiska omständigheterna i det aktuella fallet.

2. Public domain

Trots att det inte i ett visst fall föreligger immaterialrättsliga hinder att träna genera-



vändaren erhåller upphovsrätt till resultatet. Det krävs enligt den amerikanska upphovsrättsmyndigheten, the US Copyright Office, en högre grad av kreativitet för att erhålla en dylik rätt. I Kina har man vid internetdomstolen i Bejing kommit till motsatt slutsats då domstolen ansett att användarens estetiska uppfattning reflekterats i de instruktioner som getts till verktyget.¹³

Etiska utmaningar

Det immaterialrättsliga läget är sålunda ännu i många fall oklart och under utveckling då det gäller användningen av generativ AI. Men det förekommer självfallet också en hel del samhällliga frågor som pockar på en lösning.

Hur kan vi vara säkra på att det inte förekommer etiska problem eller snedvridningar i samband med generativ AI? Träningsmaterialet kan, också utan avsikt, premiera majoritetsbefolkningen framom minoriteter. Ett typiskt exempel är beviljande av lån. Ett av kriterierna kan vara boningsort. Fastän låntagaren skulle ha betalningsförmåga kan boningsorten, till exempel ett låglöneområde, bidra till att diskvalificera låntagaren.

Ett annat ofta använt exempel är be-

Då man utvecklar generativa AI-verktyg är det viktigt att inhämta samtycke av den som har upphovsrätt till träningsmaterialet. Här några förslag av Firefly i Photoshop på prompten "Artificiell intelligens väcker frågor om upphovsrätt och etik".

handlingen av arbetsökande. Om systemet tränats enbart med representanter för majoritetsbefolkningen kan en person som hör till minoriteten diskrimineras. Resultatet av generativ AI bör med andra ord vara under ständig och kritisk uppsikt. Sunt förnuft bör inte undervärderas i begeistringen över den generativa AI:ns snabbhet och förmåga att skanna stora mängder material.

Ett annat drag som kännetecknar också den generativa AI:n är att om vi enbart betonar negativa konsekvenser och risker leder tänkandet till teknosolutionism eller -regulationism, det vill säga till att de tekniska lösningarna betonas på bekostnad av sociala och samhällliga frågeställningar.¹⁴ Komplicerade samhällliga problem, såsom diskriminering och orättvisor, ses enbart som ett tekniskt problem i systemet. Diskriminering kan med andra ord bli ett programmeringsproblem för att uttrycka det litet provokativt.

I dagsläget finns det därtill redan en hel del

forskning som visar att digitaliseringen och med den diverse AI-system medför utvecklings- och inlärningsproblem för ungdomar.¹⁵

Den generativa AI:n innebär alltså inte enbart immaterialrättsliga utmaningar, utan också etiska och samhälleliga frågeställningar bör vara i fokus. Det här talar enligt mig för att diverse AI-verktyg uttryckligen bör ses som sådana, alltså verktyg. Kombinationen människa och AI bör betonas. Det är vi som i sista hand skall styra AI:n och inte tvärtom.

FOTNOTER

¹ Se t.ex. Hannu Toivonen, Mitä tekoöly on ?, MeediaZone, Viro 2023, Petri Myllymäki, Tekoälyn älykkydestä, i verket Älykäs huominen, Gaudeamus 2021, s. 9 under rubriken "Tekoälyn määrittelemisen vaikeus" och David L. Shrier, Basic AI, A Human Guide to Artificial Intelligence, Robinson 2024.

² Se Shrier s. 19 där han uttrycker sig så här "It's important to understand artificial intelligence, because it is more complicated and less predictable than how its is depicted in the media."

³ Se Shrier s. 21 och Max Oker-Blom, Artificiell intelligens och immaterialrätten som incitament för utveckling och kontroll, Ekonomiska samfundets Tidskrift 2/2019 s. 22 ff. om bl.a. definitioner och AI:s historia. Se också Mika Viljanen, Menikö juna jo? Tekoälyn sääntelemisen mahdollisuuksista, Lakimies 7-8/23 s. 1207 där han indelar AI enligt använd teknologi.

⁴ Shrier 45 ff. och s. 36 f. Se också Martin Ford, The Rise of the Robots, 2015. Miljöpåverkan berörs bl.a. av Torsten Fagerholm, Varje AI-körning har en prislapp som betalas av planeten, HBL 6 juni 2024 och i notisen "AI kan äta upp Googles klimatmål", HBL 3.7.2024.

⁵ Shrier s. 32 f.

⁶ Till exempel Max Oker-Blom, "Incentive vs access" – lyhyt katsaus immateriaalioikeuden ikuisuusksymykseen, Lakimies 7-8/2013 s. 1356-1363 och hänvisningarna där.

⁷ Se t.ex. Hanna Pynnönen, Mallioikeuksilla suojatut tuotteet tekoälyohjelmistojen koulutusdatassa – onko kyse oikeusloukkauksista?, IPRinfo 4/2024.

⁸ Min utgångspunkt är WIPO:s färsk rapport "Generative AI, Navigating intellectual property, 2024". WIPO, dvs. World Intellectual Property Organization, Världsorganisationen för den intellektuella äganderätten, är underställd Förenta nationerna, FN. Se också Christian Nilsson Zamel, Artificiell intelligens och immaterialrätt – några reflektioner från det svenska Patent- och registreringsverket, NIR 4 2021 s. 488 ff. Artikeln innehåller allmänna synpunkter på förhållandet mellan AI och immaterialrätt. När det gäller reglering av AI bör naturligtvis EU:s Artificial Intelligence Act av den 13.3.2024 noteras.

⁹ Ett exempel är fallet The New York Times v. Open AI. Se Roosa Karus, Generatiivisesta tekoälystä ja tekijänoikeudesta vuonna 2024, IPR info 1/2024.

¹⁰ Roosa Karus, Tekoälyn kouluttaminen mahdollista ilman tekijänoikeudella suojattua materiaalia, IPR info 2/2024.

¹¹ Nyligen hävdade skådespelaren Scarlett Johansson att Open AI utnyttjat hennes röst utan hennes samtycke, vilket Open AI:s VD Sam Altman dementerat. Se HBL:s nätupplaga 21.5.2024 och Roosa Karus, Scarlett Johansson esittää Open AI:n kopioineen hänen äänensä, IPRinfo 24.5.2024.

¹² Se WIPO:s rapport i fotnot 8 ovan. En annan sak är naturligtvis om AI utgör grunden för ett företags affärsmodell. Se Thomas H. Davenport, Nitin Mittal, All in on AI, How Smart Companies Win Big with Artificial Intelligence, Harvard Business Review Press 2023.

¹³ Se not 11 ovan.

¹⁴ Anna-Mari Rusanen, Algoritmien aakkoset, i verket Älykäs huominen, Gaudeamus 2021, s. 43.

¹⁵ Men också minnesutmaningar för äldre. Uttrycket "google-hjärna" har använts i detta sammanhang.