



Rapport 2023:17

# Reglering av AI: För lite för sent eller för mycket för tidigt?

En rapport om generativ AI och AI-förordningen

Dnr: 2021/50

Myndigheten för tillväxtpolitiska utvärderingar och analyser

Studentplan 3, 831 40 Östersund

Telefon: 010 447 44 00

E-post: [info@tillvaxtanalys.se](mailto:info@tillvaxtanalys.se)

[www.tillvaxtanalys.se](http://www.tillvaxtanalys.se)

För ytterligare information kontakta: David Birksjö

Telefon: 010-447 44 78

E-post: [david.birksjo@tillvaxtanalys.se](mailto:david.birksjo@tillvaxtanalys.se)

Omslagsbild: AI-genererad. Prompt: The European Union is trying to regulate AI risks by arguing in large convention halls amongst lots of well dressed people from all over Europe, amongst robots, androids and humanoids using computers.

# Förord

Tillväxtanalys uppdrag är att utvärdera och analysera effekterna av statens insatser för en hållbar nationell och regional tillväxt. Vi ska också ge underlag och rekommendationer för utveckling, omprövning och effektivisering av politiken.

Syftet med den här rapporten är att analysera den framväxande AI-förordningen utifrån ett övergripande perspektiv som *belyser rättsliga utmaningar* vid reglering av (en allt snabbare) utveckling och spridning av ny teknik (det s.k. taktproblemet) och *sätt att hantera dessa utmaningar*. AI är idag en teknik som i olika utsträckning integrerats i olika tekniska lösningar och tjänster. Genomslaget förväntas bli än större de närmsta decennierna och föra med sig en rad komplexa risker och osäkerheter.

Studien är en fallstudie av ett stort regulatoriskt ingrepp i en "general-purpose-teknologi". Teknologierna och regleringen av dem berör flertalet branscher och näringslivet samt samhället i stort. Dessutom är det en reglering som spänner över ett spektrum av olika grader av risker och olika hårdhetsgrader.

Delstudien diskuterar lärdomar från AI-förordningen i relation till att hantera taktproblemet i det svenska näringslivet i stort idag. Rapporten är skriven av Stefan Larsson, jurist och docent i teknik och social förändring vid Lunds tekniska högskola, Lunds universitet.

David Birksjö har varit projektledare. I projektet har också Peter Svensson och Helena Löfgren medverkat.

Ett varmt tack till er som har bidragit med värdefulla inspel: Jockum Hildén och Joakim Wernberg, för läsning och nyttiga kommentarer; Joakim Olsson, för grafik, och till Abir Al-Sahlani med kollegor, Anna-Mari Rusanen, Fredrik Heintz, Sofia Ranchordás, Katja de Vries, Liane Colonna, Håkan Hydén, Theo Andersson, Masayuki Murayama, den svenska ambassaden i Tokyo, och Miriam Ben Hadj Ali för diskussioner och intervjuer.

Östersund, november 2023

Sverker Härd,  
Generaldirektör, Tillväxtanalys

# Innehållsförteckning

Förord .....	2
Innehållsförteckning .....	3
Sammanfattning .....	5
1. Introduktion: reglering av artificiell intelligens .....	11
1.1 Syfte och avgränsningar .....	13
1.2 Upplägg .....	14
1.3 Material och empiri .....	14
1.3.1 Den europeiska AI-förordningen .....	15
1.3.2 Expertintervjuer .....	15
1.4 Relevans: om formativitet och Brysseeffekten .....	16
2. Generativ AI som taktproblem .....	17
2.1 2023 som formativt år .....	17
2.1.1 2023: en ny våg av policyidéer för generativ AI .....	19
2.2 Omvärldsutblick av relevans för Europa och taktproblemet .....	20
2.2.1 G7 och generativ AI .....	20
2.2.2 USA och mjuk styrning .....	21
2.2.3 Storbritannien .....	22
2.2.4 Japan .....	23
2.3 Taktproblemet: mellan rättslig förutsägbarhet och teknologisk förändring .....	24
2.3.1 Fixering: samhällets behov av rättslig förutsägbarhet .....	24
2.3.2 Flexibilitet: adaptiv styrning som lösning på taktproblem .....	25
2.3.3 Om juridisk flexibilitet: experimentell reglering .....	26
3. EU och AI .....	28
3.1 Mjuk AI-styrning .....	28
3.1.1 2016–2020: medvetenhet om AI-risker .....	28
3.2 2021: förslag på AI-förordning .....	29
3.3 Regleringslandskapet – EU:s formativa teknikregleringsperiod .....	31
3.3.1 AI och digitalisering: ett regleringslandskap i förändring .....	32
3.3.2 Om ansvarsfördelning: direktiv om skadeståndsansvar .....	33
4. Analys: AI-förordningen som taktproblem .....	34
4.1 Vad är AI enligt EU? Definitionens konsekvenser .....	34
4.1.1 Terminologisk osäkerhet – vad är skillnaden mellan grundmodell och flerändamåls-AI? .....	36

4.1.2 AI-definitionen och högriskkategorisering .....	36
4.2 2023: generativ AI i rådets och parlamentets förslag .....	37
4.2.1 Flerändamåls-AI och grundmodeller .....	37
4.3 Standardisering .....	39
4.4 Delegerade akter och genomförandeakter .....	40
4.4.1 Delegerade akter .....	40
4.4.2 Genomförandeakter .....	41
4.5 Regulatoriska sandlådor: experimentell reglering .....	42
4.5.1 Tidsbestämddhet och flexibilitet .....	42
4.5.2 Medlemsstatskomplexitet .....	43
4.5.3 Möjlig motsättning mellan regulatoriska sandlådor och generativ AI- transparens .....	43
5. Summering och diskussion: AI-reglering och innovation .....	45
5.1 I vilken mån och på vilket sätt förekommer mjuk styrning på AI-området överlag? .....	45
5.2 Hur har den snabba utvecklingen på AI-området påverkat framtagandet av den europeiska AI-förordningen? .....	46
5.3 Vilka huvudelement av flexibilitet finns i förslaget/förslagen på AI-förordning? 46	
5.4 Hur föreslås s.k. generativ AI regleras i AI-förordningen och vilka huvudsakliga utmaningar är förenade med det i relation till pågående utveckling och innovation? .....	48
6. Slutsatser .....	50
6.1 Mellan fixering och flexibilitet .....	50
6.1.1 AI-förordningens flexibilitet .....	50
6.1.2 AI-förordningens konsekvenser för svensk tillsyn .....	51
6.1.3 Om rättslig komplexitet och behovet av kunskap .....	52
6.2 Avslutning .....	53
7. Referenser .....	55
7.1 Litteratur .....	55
7.2 Andra källor .....	58

# Sammanfattning

Syftet med den här rapporten är att analysera processen med att ta fram den europeiska AI-förordningen, dvs. placera in AI-regleringsaspekter i ett bredare perspektiv och samtidigt belysa rättsliga utmaningar kopplade till taktproblemet (se Birksjö m.fl., 2022). Taktproblemet, det vill säga när innovationstakten utmanar en trögrörlig reglering, skapar i sämsta fall rättslig osäkerhet som i sin tur leder till osäkerhet för företag att investera i teknikutveckling, men även till betydande utmaningar för myndigheters tillsynsarbete. AI är en flerändamålsteknik och kommer med största sannolikhet påverka utvecklingen inom en rad sektorer. De normativa val som görs nu, både i Europa och globalt, kommer att ha stor betydelse för AI-utvecklingen framöver.

Den här rapporten visar på sätt att hantera vissa av de utmaningar som har med taktproblem att göra. Trots den rättsliga *fixering* som AI-förordningen innebär vid ett ikraftträdande, ger den här analysen en del *flexibla element* som kan ge fingervisningar om hur rättsliga utmaningar kan hanteras i relation till innovation på AI-området.

Taktproblemet är särskilt relevant att analysera i relation till AI i skärningen av två centrala aspekter: 1.) teknikutvecklingen, framför allt på området för generativ AI gällande både kapacitet och antalet användare, har tagit ett språng under 2023, vilket också föranlett EU att revidera betydande delar av AI-förordningen under själva lagstiftningsprocessen, och 2.) EU tycks ha ambitioner att inte bara reglera den europeiska marknaden utan även sträva efter en global påverkan, vilket sker parallellt med tekniksprånget, och visar på en global dynamik i en formativ period i relation till initiativ i USA, Kina, OECD och G7, m.fl.

Att förstå dessa processer – tillblivelsen och framtagandet av AI-förordningen, dess inneboende dynamik med flexibla element, men även relationen till annan lagstiftning samt kommande detaljer i verkställighet och nationell tillsyn – är en förutsättning för att svenska intressenter ska kunna påverka, hantera och positionera sig väl. Att förstå de nya inslagen i AI-förordningen möjliggör för svenska intressenter att nyttja dessa.

Ett särskilt fokus har i den här rapporten lagts vid generativ AI eftersom det är ett särdeles intressant problemområde i relation till taktproblem. Regleringen har att fixera skyddsbehov och samtidigt erbjuda flexibilitet i förhållande till förändrade förutsättningar i en nära framtid, med delvis genuin osäkerhet i vilka AI-tekniker och tillämpningsområden som växer fram.

## Studiens frågeställningar

Mer specifikt besvarar rapporten följande fyra frågor som relaterar till taktproblem:

1. I vilken mån, och på vilket sätt, förekommer mjuk styrning på AI-området?
2. Hur har den snabba utvecklingen på AI-området påverkat framtagandet av den europeiska AI-förordningen?
3. Vilka huvudelement av flexibilitet finns i förslaget/förslagen på AI-förordning?
4. Hur föreslås s.k. generativ AI regleras i AI-förordningen och vilka huvudsakliga utmaningar är förenade med det i relation till pågående utveckling och innovation?<sup>[1][5][6]</sup>

Frågorna motiveras av att AI-förordningen kommer ha stor påverkan på både svenskt näringsliv, myndigheter och i förlängningen svenska medborgare. Även om

lagstiftningens upplägg inte slutgiltigt har förhandlats, så finns det en stor nytta att för svenskt vidkommande bättre förstå processen och dess utfall så att det går att 1) ex-ante analysera dess påverkan 2) förbereda och hantera dess utfall på bästa sätt, bland annat gällande frågor om verkställighet och tillsyn och de flexibla element som pekas ut nedan, samt 3) lära inför kommande revideringar och ny tekniklagstiftning på EU-nivå.

## Tillväxtanalys om AI och taktproblem

Den här rapporten bygger på insikter från en tidigare litteraturstudie från Tillväxtanalys (Birksjö m.fl., 2022) som redovisar kunskapsläget om skillnader i utvecklingstakt mellan reglering och teknisk innovation, det vill säga, taktproblemet. Den analys som görs här av den europeiska processen med att ta fram en AI-förordning relaterar även till tidigare analyser om AI-politik i relation till konkurrenskraft (Tillväxtanalys, 2022) och hur AI omformar näringslivet (Tillväxtanalys, 2023). Samtidigt som den här rapporten, publiceras också en fallstudie från Tillväxtanalys om regleringen av drönare och elsparkcyklar.

## AI-reglering: ett komplext men högst relevant taktproblem

Den europeiska kommissionen föreslog i april 2021 den förordning som fastställer harmoniserade regler om artificiell intelligens. AI-förordningen, ingår i vad som kallas "den nya lagstiftningsramen" (*new legislative framework*, NLF), som innebär en slags produktsäkerhetsreglering. Med det följer ett tydligt fokus på bedömningar av överensstämmelse (*conformity assessment*) som måste genomföras innan en AI-produkt inom högrisksegmentet placeras på marknaden. CE-märkning är därmed en central del.

Efter en stor mängd ändringsförslag samt rådets version, publicerad den 6 december 2022, och parlamentets bekräftade position från den 14 juni 2023, så pågår i skrivande stund förhandlingar för att nå en slutgiltig version.

Parallellt med lagstiftningsprocessen har AI-utvecklingen gått fort. Även om de skäl som anges i den här rapporten som stöd för en snabb AI-utveckling endast utgör ett smalt urval, finns det fog för argumentet att AI-landskapet har förändrats tämligen drastiskt bara under några månader från slutet av 2022 till och med 2023. Ett tecken på det är inte minst vad som ser ut som en ny våg av policydokument, där aktörer som OECD, G7, ACM, World Economic Forum och tillsynsmyndigheter i exempelvis Storbritannien utvecklar och publicerar riktlinjer för generativ AI.

Givet detta visar analysen hur AI-definitionen i AI-förordningen genomgår tämligen kraftiga förändringar från kommissionens förslag till rådets och parlamentets ändringsförslag). Det gäller den allmänna AI-definitionen, men lika intressant är det att koncept som *general purpose AI systems* och *generative AI systems* finns med först i rådets version och sedan i parlamentets version där framför allt s.k. *foundation models*, dvs. AI-grundmodeller, regleras särskilt. Med det sistnämnda eftersträvar parlamentet framför allt mer transparens kring träningsdata och dess hantering för AI-grundmodeller och generativ AI som ligger till grund för innovationer som ChatGPT, baserat på OpenAI:s GPT3 och 4, och språkmodeller som Metas LLaMa, Googles Bard, Anthropic chatbot Claude, Hugging Faces BLOOM eller kinesiska varianter som Huawei's PanGu- $\Sigma$  eller Baidus Ernie 3.0 Titan. Liknande underliggande modellträning ligger även till grund för

AI-system som kan generera bilder genom instruktioner, s.k. prompter, som DALL·E 3, Midjourney och Stable Diffusion. Parlamentet föreslår även transparens- och dokumentationskrav för energiförbrukning och i vilken mån träningsdata är upphovsrättsskyddade.

Även om det finns en rad säkerhetsutmaningar och problem med partiskhet i modeller, som kan leda till diskriminerande effekter, så är det svårt för lagstiftaren att fånga skyddsbehov och ställa välkalibrerade krav på rätt typ av AI-aktör när utvecklingen går så fort. Tumregeln tycks ofta bli att säkerställa transparens genom krav på dokumentation, för att säkerställa möjligheter till ansvarsutkrävande och överhuvudtaget var i den komplexa utvecklingsprocessen som något klandervärt har uppstått.

## AI-förordningens flexibilitet

Av de exempel på policyverktyg som litteraturen kring taktproblem tar upp fokuseras i denna analys framför allt rutiner för kontinuerlig teknikbevakning och riskbedömningar, och metoder för att testa teknik under kontrollerade former såsom regulatoriska sandlådor. Det gäller även annan mjuk reglering, i termer av frivilliga överenskommelser, standardiseringsprocesser eller – något som blivit vanligt inom AI-området – principiellt formulerade riktlinjer.

Man kan börja med att konstatera att mjuk styrning i termer av riktlinjer, ofta formulerat i termer av etik, har spelat en betydande roll för AI-utveckling och AI-användning de senaste sju åren. En ny policyvåg tycks ha kommit under 2023 med anledning av generativ AI. De principiella och etiska riktlinjer för AI som har tagits fram har ofta fokuserat behovet av bättre transparens och förklarbarhet, frågor om ansvarsfördelning, behovet av tillförlitliga tillämpningar och vikten av att tydligt hantera rättvisefrågor. Dessa tycks ha bidragit till en del av AI-förordningens fokus på transparens och tillit (parlamentet och rådet hänvisar explicit till den av kommissionen utsedda AI-expertgruppens riktlinjer från 2019). Styrkan i denna mjuka styrning ligger i både snabbhet och möjligheten att finna de principiellt viktigaste förhållningssätten. Svagheten ligger naturligtvis i bristen på sanktioner och avsaknad av process (Larsson, 2020).

Kombinationen av AI och regulatoriska sandlådor är en diskussion och utveckling som hänger starkt ihop med AI-förordningens process, med idén att deltagande företag får undantag från specifika lagar eller efterlevnadsprocesser för att förnya sig, och en spansk pilot startades redan 2022. Eftersom de olika förslagen på regulatoriska sandlådor i olika grad föreskrev inrättande av regulatoriska sandlådor i medlemsstaterna är det i skrivande stund inte klart hur tvingande AI-förordningen blir för att inrätta en sådan. Litteraturen om taktproblem lyfter dock fram regulatoriska sandlådor som en möjlig innovationsfrämjande väg, även om det finns kritik mot hur generellt lärandet egentligen är, utöver att främja den enskilde aktören, och att de ibland utvärderats otillfredsställande (jfr Ranchordas, 2021b).

- Det tycks därmed finnas behov av att utveckla mer transparent och objektiv design av experimentella regler och regulatoriska sandlådor.
- Det finns också en medlemsstatskomplexitet gällande AI som lyfts av analytiker (se Buocz m.fl., 2023), där relationen mellan nationell lag och EU-lagstiftning kan behöva förtydligas.



- De mildare formerna av regulatoriska sandlådor som innebär en närmre samverkan med tillsynsmyndigheter finns det positiva erfarenheter av, och starka skäl för att ta med sig för svenskt vidkommande i relation till när AI-förordningen ska verkställas. Svenska myndigheter bör därför planera för detta redan nu.

Även om det i den här kartläggningen inte funnits utrymme att fördjupa standardiseringsdelen av AI-förordningen kan det konstateras att det både kommer att vara ett viktigt instrument för efterlevnad, eftersom s.k. harmoniserade standarder förutsätts uppfylla krav som ställs i förordningen. Att bidra till standarderna kan därmed vara ett sätt att bidra till regleringens verkställighet och med minskade kostnader för efterlevnad. Med det sagt har standardiseringsproceduren kritiserats med tanke på resurssvaga intresseorganisationers möjligheter, samt ifrågasatts utifrån att tekniska standarder riskerar att inte vara lämpliga för att säkerställa grundläggande rättigheter (jfr BEUC, 2021; diskuteras av Micklitz, 2023).

En intressant aspekt rör de mandat som inom AI-förordningen kan ges till kommissionen att i efterhand ändra och förtydliga lagstiftningen genom s.k. *delegerade akter* och *genomförandeakter*. Detta gäller inte minst AI-definitionen i sig, där både kommissionens och rådets version föreslog möjliga tillägg till listan med AI-tekniker, och där rådets version gav även möjlighet till att ta bort tekniker från listan, under vissa specificerade omständigheter. Rådet introducerade *general purpose AI* dvs. flerändamåls-AI, och anger att kommissionen genom en genomförandeakt ska specificera hur högrisk-kraven ska tillämpas i förhållande till dessa mer generellt tillämpliga AI-system. Konsekvenserna av detta upplägg bör man för svenskt vidkommande titta närmre på, ur perspektivet innovation, med tanke på att förutsägbarheten minskar om delar kan ändras senare. Det är därmed också en fråga om rättssäkerhet.

Överhuvudtaget tycks fältet för dessa specifika möjligheter att ändra och förtydliga lagstiftning relativt obeforskat utifrån innovationspåverkan, rättssäkerhet och taktproblematik. Det finns en möjlig brist på förutsägbarhet här som behöver förstås bättre, speciellt för så centrala element som *definitionen av AI* i en AI-lagstiftning.

## Slutsatser

Det finns tre områden där studien har landat i slutsatser som berör svenskt politiskt beslutsfattande, och de områdena är AI-förordningens flexibilitet, AI-förordningens konsekvenser för svensk tillsyn, och dess del i en rättslig komplexitet.

### AI-förordningens flexibilitet

- Det finns en konceptuell osäkerhet i AI som styrningsbegrepp. Det gäller "AI" i sig, såväl som de nyare begreppen "grundmodell", "AI-system för allmänna ändamål" och "flerändamåls-AI". Att AI-förordningen knyter lika krav till de två begreppen kommer därför troligen leda till en osäkerhet hos praktiker om vilka krav deras AI-utveckling ska råda under. Betydelsen av detta behöver följas upp i mer detalj.
- Viss sorts flexibilitet gör att AI-regleringen förskjuter makt över bestämmandet i lagstiftningsprocess till kommissionen, en annan till standardiseringsorgan. Den europeiska lagstiftaren tycks både vilja adressera generativ AI och samtidigt skjuta det på framtiden. De verktyg som står till buds är s.k. delegerade akter och genomförandeakter, samt harmoniserade standarder.

- Vissa taktproblemlösningar tycks ha en betydande baksida även för marknader: oförutsägbarhet. Dvs., kunskapen kring vilken sorts taktproblemlösningar som bör användas, i relation till deras konsekvenser, behöver fördjupas.
- Den analys som gjorts här pekar mot att det tycks finnas vissa frågetecken kring hur användbara regulatoriska sandlådor är för AI-innovationer i Sverige. Det handlar mer generellt om huruvida det är möjligt med ett bredare lärande än för det enskilda testfallet, och samtidigt om förenligheten med svensk rätt. Detta beror bland annat på vad man avser med regulatorisk sandlåda, se mer om tillsynssidan nedan. Mer specifikt gällande grundmodeller finns det frågetecken kring om det är möjligt att förena ett befriande av dokumentations- och transparenskrav i en sandlådefas med ett tilläpande av samma krav vid ett senare inträde på marknaden.

### **AI-förordningens konsekvenser för svensk tillsyn**

- Mer flexibel praktik och vägledning hos tillsynsmyndigheter framstår som ett möjligt och önskvärt sätt att hantera taktproblem.
  - Det gäller särskilt *praktisk rättslig vägledning* från myndigheter i relation till AI-utveckling och reglering, vilket därmed kan uppmuntras.
  - Det bör ses som *ett lärande även för regleraren* i termer av att både förvaltningsmyndigheter och tillsynsmyndigheter troligen behöver kompetensutveckla vad AI-utvecklingen medför på deras verksamhets- och tillsynsområde. Flera goda exempel finns.
  - Avslutningsvis tycks *myndighetssamverkan* vara central, eftersom både AI-tillämpningar generellt och AI-förordningen specifikt träffar så många sakområden.
- Detta är särskilt relevant med tanke på AI-förordningens ämnesmässiga bredd, med alltifrån frågor om diskriminering och mänskliga rättigheter till tekniskt avancerade frågor, som kan relatera till finansiella marknader, konkurrens, konsumenter, medicinska applikationer, och dataskyddsfrågor, m.m.

### **Om rättslig komplexitet och behovet av kunskap**

- Det finns rättsliga överlapp i det europeiska regleringslandskapet som behöver förstås bättre, och för svenskt vidkommande troligen koordineras mer. I linje med litteratur om rättsliga överlapp (Crootof & Ard, 2021), finns det anledning att tro att det kommer uppstå en rad frågor i relationen mellan flertalet betydande rättsakter av betydelse för AI-området. Det gäller inte minst relationen mellan AI-förordningen, GDPR, rättsakten för digitala tjänster (DSA), rättsakten för digitala marknader (DMA) såväl som för upphovsrättsdirektivet.
- Även bortom rättsliga frågor behövs mer kunskapsutveckling om AI som teknik och dessa teknologiers samhällskonsekvenser. En starkare offentlig satsning vore eftersträvarsvärd. Tillväxtanalys (2022) har tidigare visat på hur det framför allt är privata satsningar på AI-forskning som utgör djup och tyngd. Någon motsvarande större offentlig satsning finns inte i Sverige.

### **Metodval**

Litteraturen som bistår analysen uttrycker i sig inte en systematisk genomgång, men bygger dels på tematiska sökningar kring AI och reglering men även på tematik runt till exempel experimentell reglering (jfr Ranchordás, 2021), och dels på en systematisk

genomgång som har gjorts av författarna till en tidigare rapport från Tillväxtanalys (Birksjö m.fl., 2022). Utöver genomgång av relevant litteratur analyseras ett betydande urval av dokument centrala för den europeiska AI-förordningsprocessen – inte minst kommissionens respektive rådets och parlamentets förslag – och som komplement till detta genomfördes 10 intervjuer med olika sorters expertis på området.

# 1. Introduktion: reglering av artificiell intelligens

Det tycks finnas en ständigt pågående dynamik mellan rättsliga regleringars anspråk och teknikinnovationers möjligheter och risker. Ofta uttrycks detta som ett taktproblem, det vill säga att det är en skillnad på utvecklingstakt mellan reglering och teknisk innovation, vilket också föranlett förslag på mer adaptiva sätt att styra (Birksjö m.fl., 2022). När det gäller artificiell intelligens (AI) befinner vi oss just nu mitt i ett skede där detta är extra påtagligt. Även om etiska frågor kring AI har diskuterats flitigt åtminstone sedan 2017 (jfr Larsson m.fl., 2019) och Europakommissionens förslag på en AI-förordning publicerades i april 2021, så kan 2023 ändå ses som ett år fyllt av särdeles livliga regleringsdiskussioner med fokus på AI. Orsaken till det är å ena sidan att kapaciteten hos de stora AI-modeller som kan generera texter, bilder och ljud blivit hög – ofta kallat generativ AI – att de därmed kan generera trovärdiga uttryck i mänsklig stil, å andra sidan att de blivit så vardagligt tillgängliga, vilket skapat en allmän medvetenhet. Tillgängligheten är kanske tydligast symboliserat av ChatGPT, som släpptes av Open AI den 30 november 2022, men det är bara toppen på ett isberg av en snabb utveckling med stora investeringar hos alla de större teknikbolagen. Två månader efter Chat GPT blev publik hade tjänsten 100 miljoner unika besökare, vilket pekar på en hastighet som samtidigt leder till frågor om hur styrning av AI-utveckling bör gå till, generellt, och hur den europeiska AI-förordningen ska kunna hantera den snabba innovationstakten på ett förutsägbart vis, specifikt (The Guardian, 2 december, 2023).

Våren 2023 har, i takt med att de generativa AI-modellerna har visat upp sin kapacitet till både fantastiska möjligheter och risker, lett till allt ifrån krav på moratorium på sex månader och efterlysning av bättre reglering från G7-länderna till en påbörjad förhandling mellan EU:s institutioner för att nå fram till en slutgiltig version på AI-förordningen. Hösten 2023 har fortsatt i samma anda med en presidentorder från Bidenadministrationen, riktlinjer och uppförandekod från G7, och ett gemensamt ställningstagande från ett brittiskt möte om AI och säkerhetsfrågor i början av november, som undertecknades av bl.a. USA och Kina.

Denna rapport analyserar det europeiska förslaget på AI-reglering utifrån detta taktproblem, inklusive en dynamik mellan fixering och flexibilitet, och med betoning på generativ AI, som gett särskilt avtryck i parlamentets version från den 14 juni 2023. I fokus ligger i den här analysen *inte* den slutgiltiga regleringen – den är heller inte klar i skrivande stund – utan processen i sig, indikativ som den är för en formativ period av betydande relevans för taktproblemet. Beslut som tas nu kommer att ha betydande konsekvenser för AI-utveckling och AI-tillämpning en lång tid framöver, i Sverige, Europa och troligen globalt. Förhoppningen med denna rapport, utöver att bidra till överblick över det komplexa fältet kring styrning (*governance*) av AI, är att bidra med rekommendationer till förhållningssätt utifrån svensk kontext.

Takten, kan konstateras, är i sig ett dilemma för både innovation och reglering. Att hitta rimliga balanser i reglering snabbt är svårt, varför också lagstiftare traditionellt sett har utvecklat en rad tekniker över tid, i allt från överprövning och ramlagstiftning till att lämna uttolkning av principiell reglering till domstolar att förtydliga i praxis. Denna flexibilitet står naturligtvis till buds även gällande AI-användning, oavsett en specifik AI-

förordnings vardande eller ej. Det faktum att den svenska avtalslagen (1915:218) från 1915 fortfarande fungerar för att reglera uttryck för avsikter och vad som ska anses rimligt förfarande i avtalslutande är ett bra exempel på det, i ljuset av den enorma skillnad i praktiker och teknologier som uppstått sedan innan första världskriget till nu. Och, givet en förordning, så står dels den Europeiska unionens domstol till buds för att reda ut oklarheter, dels kommer mycket av den praktiska flexibiliteten spelas ut i hur tillsynen designas och efterlevs, vilket ännu återstår att se gällande AI-förordningen.

Litteraturen i skärningen mellan innovation och reglering efterlyser dock ännu mer flexibilitet än vad rättsordningen traditionellt sett har erbjudit, i takt med att förhållandena och teknologierna blivit mer komplexa och snabbväxande. Och, mer specifikt gällande AI, är det rimligt att fråga i vilken mån man både bör och kan försöka reglera något som utvecklas så snabbt och har så stor potential till samhällsnytta, och vilka konsekvenser detta kan ha. Frågan ställdes i februari 2023 till en av teknikcheferna på Open AI, Mira Murati. Hon leder teamen bakom DALL·E, som använder AI för att skapa visuella uttryck, som konstverk och bilder, baserat på prompter. Murati konstaterade:

*Det är inte för tidigt. Det är väldigt viktigt för alla att börja engagera sig, med tanke på vilken effekt dessa tekniker kommer att få (Mira Murati, TIME, 5/2, 2023).<sup>1</sup>*

Liknande uttalanden gjorde även Sam Altman, vd för Open AI, när han kallades in till den amerikanska senaten den 16 maj, 2023. Han har i intervjuer uttryckt en oro för storskaliga desinformationskampanjer och möjligheterna till AI-programmerade verktyg för cyberattacker. Detta samtidigt som han ser enorma möjligheter till att var och en har en egen digital assistent, en "co-pilot", som kan effektivisera ens vardag och hjälpa till med allt möjligt (ABC News, 16 mars 2023). Han sätter därmed fingret på ett centralt AI-regleringsdilemma: hur säkerställer man nyttorna utan att dämpa innovationen med behovet av regleringskrav? AI-området, och särskilt för generativ AI, är illustrativt för taktproblemet, vilket gör att fältet är särdeles intressant för analys.

Först, ska sägas, är AI-användning på intet sätt oreglerad i svensk och europeisk rätt idag. Om man ser till dess påtalade risker inom olika användningsområden så kan man konstatera att det finns exempelvis diskrimineringslagstiftning, dataskydd och konkurrenslagstiftning, som kan vara relevant för både utveckling och tillämpning av AI. I samband med den konsekvensanalys som EU-kommissionen publicerade med förslaget på AI-förordning i april 2021 konstaterade kommissionen att lagstiftningsalternativet med ett horisontellt EU-lagstiftningsinstrument som fastställer obligatoriska krav för högrisk-AI-tillämpningar men samtidigt erbjuder frivilliga uppförandekoder för icke-högrisk-AI-tillämpningar är mest lämpligt. Detta eftersom det "erbjuder proportionerliga skydd mot riskerna med AI, samtidigt som administrations- och efterlevnadskostnaderna begränsas till ett minimum" (Kommissionen, 2021b, s. 2).

Det finns naturligtvis farhågor för negativa konsekvenser med den europeiska AI-förordningen, inte minst från ett industriperspektiv. Utöver det relativt omtalade moratorieförslag som publicerades av Future of Life Institute i mars 2023, så har en större grupp oberoende personer, chefer och forskare nyligen uttryckt mer konkurrens- och

---

<sup>1</sup> Murati utnämndes i oktober 2023 till en av 38 medlemmar i FN:s rådgivande organ för AI-frågor.

innovationsrelaterade farhågor i ett öppet brev adresserat till Europeiska unionen (Techzine, 30 juni 2023). Å andra sidan har över 100 organisationer som Algorithm Watch, BEUC och Privacy International, genom European Digital Rights (EDRI, september 2023), argumenterat för att rådets och parlamentets versioner innehåller ett väsentligt kryphål i och med att utvecklare av högrisksystem i dessa versioner mer bestämmer själva om de anser att systemet är högrisk (art. 6). Det skulle kunna undergräva syftet med regleringen i sig och som sämst även skapa mer av en politiserad ad hoc-tillämpning från kommissionens sida, vilket utvecklas mer nedan (se även Haataja & Bryson, 2023).

I litteraturen om taktproblem, som diskuteras i en tidigare rapport från Tillväxtanalys (Birksjö m.fl., 2022), diskuteras olika varianter av uttryck för dynamisk reglering och mjuk styrning som sätt att hantera taktproblemet. Åtgärder som pekas på är saker som kontinuerlig teknikbevakning och riskbedömning, regulatoriska sandlådor som tillåter olika grad av frihet i begränsade innovationsupplägg och varianter av mjuk reglering som branschcoder eller frivilliga standardiseringsförfaranden. Taktproblemet kan också ses i ljuset av en syn på sam- eller självreglering (*co-regulation*) som har utvecklats hos den europeiska lagstiftaren en tid, bland annat uttryckt i termer av rättsligt hårda krav på mjuk styrning genom uppförandekoder, till exempel för stora teknikplattformar (jfr Marsden, 2022).<sup>2</sup> Det mesta av detta tankegodts har sitt uttryck i den bredare diskussionen om styrning av teknologi – i termer av etisk styrning av AI i relation till tillit (Felländer m.fl., 2022; Larsson, 2020), standarder (Veale & Zuiderveen Borgesius, 2021) och regulatoriska sandlådor (Buocz m.fl., 2023; Ranchordás, 2021) – vilket har sina uttryck även i AI-förordningen. Det är i ljuset av denna litteratur som denna analys av den europeiska AI-förordningens process ska ses, som ett slags dialektik mellan juridikens fixering och innovationsstyrningens behov av samtidig flexibilitet.

## 1.1 Syfte och avgränsningar

Syftet med studien är att *belysa rättsliga utmaningar* vid reglering av utveckling och spridning av AI, det vill säga det så kallade taktproblemet, och *sätt att hantera dessa utmaningar*. Ett särskilt fokus läggs vid så kallad generativ AI eftersom det är ett särdeles relevant område i relation till snabb innovation och regelutveckling, som har att fixera skyddsbehov och samtidigt erbjuda flexibilitet i förhållande till förändrade förutsättningar i en nära framtid med delvis genuin osäkerhet i vilka AI-tekniker och tillämpningsområden som växer fram. Forskningsobjektet här är den europeiska *processen* att ta fram en övergripande AI-reglering. Processen uttrycker taktproblemet, i skärningen mellan behovet av att juridiskt *fixera* normativa ställningstaganden och behovet av att *flexibelt* förhålla sig till en rörlig teknikinnovation. Förhoppningen är att resultatet är relevant även för förståelsen av den slutgiltiga regleringen, särskilt när medlemsstaterna förbereder sig på sin roll i tillämpning och tillsyn.

Mer specifikt besvarar denna rapport följande fyra frågor som relaterar till taktproblem:

1. I vilken mån och på vilket sätt förekommer mjuk styrning på AI-området överlag?

---

<sup>2</sup> Se även Europeiska kommissionen (20 maj 2015) The Principles for better self- and co-regulation endorsed in the Better Regulation Package. <https://digital-strategy.ec.europa.eu/en/news/principles-better-self-and-co-regulation-endorsed-better-regulation-package> [senast besökt 5 oktober 2023].

2. Hur har den snabba utvecklingen på AI-området påverkat framtagandet av den europeiska AI-förordningen?
3. Vilka huvudelement av flexibilitet finns i förslaget/förslagen på AI-förordning?
4. Hur föreslås s.k. generativ AI regleras i AI-förordningen och vilka huvudsakliga utmaningar är förenade med det i relation till pågående utveckling och innovation?<sup>[11]</sup><sub>[SEP]</sub>

Den europeiska AI-regleringsprocessen tar sig tydligast uttryck i de olika förslagen på AI-förordningen från EU-kommissionen, Europeiska unionens råd ("rådet") och Europaparlamentet, men kan även studeras i de riktlinjer och principiella dokument som leder fram till förslagen på AI-förordning. Det gör att frågorna ovan besvaras med hjälp av litteraturstudier, analys av de olika lagstiftningsförslagen från kommissionen, rådet och parlamentet. Dessa källor kompletteras med intervjuer av experter på olika, relaterade områden, med tanke på både den tekniska, juridiska och marknadsmässiga komplexitet som AI-regleringen berör.

Eftersom var och en av dessa frågor bjuder på en oerhörd underliggande rikedom, kommer överblick eftersträvas snarare än detaljerat djup, hela tiden i nära anslutning till AI sett som ett taktproblem. Däri ligger också studiens tydligaste avgränsning: att AI-regleringen är omfattande och förslagen på förordning i sig komplexa, vilket gör att enbart ett mindre antal delar kan adresseras här. I strävan efter överblick placeras dock AI-förordningen in i en längre trend kring AI-styrningsfrågor, med både en förhistoria till AI-förordningen samt förordningens relation till andra betydande rättsakter. Dock med generativ AI i huvudsakligt fokus, eftersom det finns ett behov av mer analys på detta nya område.

## 1.2 Upplägg

*Fråga 1* besvaras framför allt i efterföljande kapitel 2, med global utblick och ett axplock av policyidéer från 2023, och kapitel 3, som redogör för den stora mängden principiella dokument och riktlinjer på AI-området som publicerats de senaste ca sju åren.

Perspektivet är globalt men med fokus på Europa, fram till och med kommissionens förslag på AI-förordning under 2021. *Fråga 2, 3* och *4* besvaras i kapitel 3.3., som ger en överblick över en formativ europeisk teknikregleringsperiod, och kapitel 4, som redogör för förändringar kring AI-begreppet över de olika förslagen till AI-förordning från 2021–2023, med fokus på flexibla element som regulatoriska sandlådor, standardisering och kommissionens mandat att förändra eller förtydliga delar av AI-förordningen genom s.k. delegerade akter eller genomförandeakter. Svaren summeras och diskuteras i kapitel 5, och slutsatser och rekommendationer presenteras i det avslutande kapitel 6.

## 1.3 Material och empiri

Den litteratur som bistår analysen uttrycker i sig inte en systematisk genomgång, men bygger på tematiska sökningar kring AI och reglering, på tematik runt till exempel experimentell reglering (jfr Ranchordás, 2021), samt på en systematisk genomgång som har gjorts av författarna till en tidigare rapport från Tillväxtanalys (Birksjö m.fl., 2022). Den rapporten tituleras *Utmaningar vid reglering av teknisk innovation – möjliga policyåtgärder* och redovisar kunskapsläget om skillnader i utvecklingstakt mellan reglering och teknisk innovation. Rapporten är en litteraturstudie som utgår från det engelska begreppet "pacing problem". Utöver genomgång av relevant litteratur analyseras ett betydande urval av dokument centrala för den europeiska AI-

förordningsprocessen, och som komplement till detta genomfördes 10 intervjuer med olika sorters expertis på området.

### 1.3.1 Den europeiska AI-förordningen

Den europeiska AI-förordningen är i skrivande stund inte färdigförhandlad och antagen, men processen och dess olika förslag är i sig talande för taktproblematiken. Denna process innehåller, i korthet, dels en rad dokument framtagna på kommissionens mandat – till exempel *Etiska riktlinjer för tillförlitlig AI* (AI HLEG, 2019), som en expertgrupp tillsatt av kommissionen tagit fram, och den vitbok om AI som kommissionen publicerade 2020 – och tre förslag till AI-förordning:

1. Kommissionen antog förslaget om en förordning som fastställer harmoniserade regler om artificiell intelligens (AI-förordningen) den 21 april 2021.
2. Rådet antog enhälligt sin allmänna inställning till förslaget den 6 december 2022.
3. Europaparlamentet bekräftade sin position i plenum den 14 juni 2023.

Dessa visar på intressanta förändringar i processen, som analyseras nedan, framför allt gällande generativ AI, men även regulatoriska sandlådor och så kallade delegerade akter, som anger en möjlighet för kommissionen att vid ett senare tillfälle föreslå ändringar som är tänkta att förtydliga eller specificera vissa oklarheter i lagstiftningen.

### 1.3.2 Expertintervjuer

Som komplement till litteratur och dokumentstudier har intervjuer genomförts med expertis på området. Följande personer har intervjuats:

- Abir Al-Sahlani, ledamot i Europaparlamentet sedan 2019, sitter bland annat i utskottet för medborgerliga fri- och rättigheter samt rättsliga och inrikes frågor (LIBE).
- Anna-Mari Rusanen, fil. dr., universitetslektor i kognitionsvetenskap, Institutionen för digitala humaniora, Helsingfors universitet och rådgivare för det finska finansministeriet. Deltog som representant för ministeriet i att formulera det finska förhållningssättet till AI-förordningen.
- Fredrik Heintz, professor i datavetenskap vid Linköpings universitet, AI-forskare och tidigare medlem i EU-kommissionens expertgrupp på AI-området (AI HLEG).
- Sofia Ranchordás, professor i förvaltningsrätt vid Tilburgs juridiska fakultet. Expert på regleringsutmaningar kopplade till ny teknik. Har bl.a. publicerat studier om smarta städer, AI och offentlighetsrätt (2022) och om experimentell reglering och regulatoriska sandlådor (2021).
- Katja de Vries, jur.dr och biträdande lektor i offentlig rätt vid Uppsala universitet. Forskar främst kring juridik och AI.
- Liane Colonna, jur.dr, biträdande lektor i juridik vid institutet för rättsinformatik, juridiska institutionen, Stockholms universitet. Har bland annat analyserat forskningsundantaget i förslaget på AI-förordning (Colonna, 2023).
- Håkan Hydén, emeritus professor i rättssociologi. Har under många år analyserat relationen mellan teknik och normer, vilket på senare år inkluderat olika aspekter av digitalisering och, mer specifikt, AI (jfr Hydén, 2022).
- Theo Andersson, omvärldsbevakare, enheten för analys, Digg – Myndigheten för digital förvaltning. Tidigare medlem i Kommittén för teknologisk innovation och etik (Komet).



- Miriam Ben Hadj Ali, departementssekreterare på finansdepartementet och samordnare för Sveriges deltagande i förhandlingarna om AI-förordningen.
- Intervju med innovationshandläggare vid svenska ambassaden i Tokyo, samt diskussioner med juridikprofessor Masayuki Murayama vid Meiji-universitetet i Tokyo, Japan.

Syftet med dessa intervjuer har framför allt varit att säkerställa att regleringsprocessen har uppfattats på ett korrekt och tillfredsställande sätt, på ett juridiskt, tekniskt eller teoretiskt plan.

## 1.4 Relevans: om formativitet och Brysseleffekten

Intrycket av att taktproblemet är särskilt intressant i relation till AI kan noteras i skärningen av två centrala aspekter: 1.) teknikutvecklingen tycks ta ett språng under 2023 och 2024, vilket vi återkommer till i nästa kapitel, 2.) EU tycks ha ambitioner att inte bara reglera den europeiska marknaden utan även sträva efter en global påverkan, vilket sker parallellt med tekniksprånget.

Utöver vad som brukar kallas för Brysseleffekten, dvs. att europeisk reglering har påverkan även utanför Europa (myntat av Bradford, 2012), uttalas just denna ambition gällande AI-förordningen på olika sätt. När Ursula von der Leyen som ordförande i Europeiska kommissionen höll tal den 13 september 2023 om tillståndet i unionen underströk hon både taktproblematiken och den roll EU spelar globalt.<sup>3</sup> Hon beskrev AI-förordningen som "världens första omfattande innovationsvänliga AI-lag" och som "en modell för hela världen". Hon efterlyste därmed en "ny global ram för AI, som bygger på tre pelare: skyddsåtgärder, styrning och vägledning för innovation". Hon pekade explicit på FN:s klimatpanel som modell för att hantera de effekter AI har i våra samhällen, dvs. en global panel som förser beslutsfattarna med den senaste forskningen. Även om von der Leyens globala ambitionsnivå är väldigt hög (med referens till G7:s Hiroshimaprocess vi återkommer till nedan) så pekar hon därmed på ett problem som taktproblemlitteraturen även identifierat kring informationsasymmetrier mellan teknikutvecklare och beslutsfattare, vilket vi återkommer till i relation till framför allt myndigheter nedan.

Kort sagt, AI-området sett som både som innovation och reglering, tycks ha något av ett formativt ögonblick både globalt och i Europa just nu, av betydande relevans även för svenskt vidkommande. Ju klokare vi kan bli kring den AI-taktproblematik där processen kring den europeiska AI-förordningen är den viktigaste komponenten, desto bättre kan vi hantera dess utmaningar.

---

<sup>3</sup> Se Europeiska kommissionen (13 september 2023) "Tal om tillståndet i unionen 2023 av kommissionens ordförande Ursula von der Leyen". [https://ec.europa.eu/commission/presscorner/detail/sv/speech\\_23\\_4426](https://ec.europa.eu/commission/presscorner/detail/sv/speech_23_4426) [senast besökt 6 november 2023].

## 2. Generativ AI som taktproblem

De stora språkmodellernas kapacitetsutveckling har pågått en tid. Ett större genombrott kom med den s.k. transformerarkitekturen som behövde mycket mindre träning än andra tekniker för att uppnå ööverträffade resultat i språköversättning.<sup>4</sup> Ökningen av kapacitet noterades mer publikt i Open AI:s GPT3, dvs. den tredje versionen av en *generative pre-trained transformer* som lanserades i juni 2020 och i Sverige bidrog till utvecklingen av en svensk version, GPT-SWE, som presenterades i december 2021 (AI Sweden, 7 december 2021). Det pågår en slags tävling mellan några av de största teknikaktörerna när det gäller stora språkmodellens kapacitet. Även om Chat GPT med sin publika lansering i november 2022 slog världen med häpnad i termer av hur snabbt den anammades av miljontals användare, så är det bara toppen på ett isberg av aktörer som utvecklar avancerade AI-modeller. Det inkluderar språkmodeller som Metas LLaMa, Googles Bard, Anthropic (som Amazon nyligen investerade 44 miljarder i)<sup>5</sup> chatbot Claude, Hugging Faces BLOOM eller kinesiska varianter som Huawei PanGu- $\Sigma$  eller Baidus Ernie 3.0 Titan.

Liknande underliggande modellträning ligger även till grund för AI-system som kan generera bilder genom prompter, som DALL-E 3, Midjourney och Stable Diffusion. Även ljud kan instrueras eller "promptas" i termer av att exempelvis VALL-E-modellen, som Microsoft publicerade i januari 2023, tycks kräva bara 3 sekunder från en ljudinspelning för att realistiskt kunna generera nya uttryck i samma stil. Forskarna bakom modellen ser också möjliga risker med att den typen av verktyg kan missbrukas för exempelvis bedrägeri (Wang m.fl., 2023, för diskussion, se Larsson, 2023).

### 2.1 2023 som formativt år

Våren 2023 är därmed särskilt intressant ur synvinkeln styrning av AI-utveckling. Här följer ett axplock av några betydande händelser:

- Moratoriebrev med över 31 000 signaturer från Future of Life Institute publiceras, som kräver en paus på åtminstone sex månader för träning av AI-system som är mer kraftfulla än GPT-4: **22 mars**.
- Forskare kopplade till forskningsinstitutet AI Now argumenterar för AI-förordningens behov av att adressera *general purpose AI* ur synvinkeln att risker för skadlig och diskriminerande bias förstärks av AI-modellerna, och riskerar att ducka ansvar: **14 april**.<sup>6</sup>
- Sam Altman, Open AI:s vd bjuds in till den amerikanska senaten för att vittna om AI och eventuella behov av reglering: **16 maj**.

---

<sup>4</sup> Den bygger på en djupinlärningsmodell som först introducerades 2017. Nätverksarkitekturen kallades för *Transformer*, och artikeln heter "Attention is all You Need", Vaswani et al., 2017.

<sup>5</sup> Enligt en artikel i The New York Times (25 september 2023) "Amazon Takes a Big Stake in the A.I. Start-Up Anthropic" av Adam Satariano and Cade Metz. <https://www.nytimes.com/2023/09/25/technology/amazon-anthropic-ai-deal.html> [senast besökt 6 november 2023].

<sup>6</sup> Se AI Now (14 april 2023) "Five considerations to guide the regulation of 'General Purpose AI' in the EU's AI Act".

- G7-ländernas publicerar en gemensam kommuniké efter ett möte i japanska Hiroshima. De påkallar behovet av en "Hiroshima-AI-process" för styrning av bland annat generativ AI, behovet av transparens, och uttrycker taktproblematiken mellan innovation och reglering. Här inkluderas OECD och GPAI: **21 maj**.
- Ytterligare ett namninsamlingsbrev, från Centre for AI Safety (CAIS), inklusive Sam Altman och Geoffrey Hinton, som kräver att minskandet av risken för utrotning från AI bör vara en global prioritet vid sidan av andra samhällsrisker som pandemier och kärnvapenkrig: **30 maj**.
- Margrethe Vestager, EU-kommissionär med ansvar för konkurrensfrågor och verkställande vice ordförande med ansvar för ett Europa rustat för den digitala tidsåldern, kallar generativ AI för en "game-changer", och tycker att regleringen går för långsamt och vill ha en code of conduct på plats "inom veckor": **31 maj**.
- EU och USA lanserade en gemensam AI-terminologi och taxonomi, som hänvisar både till EU-expertgruppens etiska riktlinjer för tillförlitlig AI, och visar på en synkronisering i policyutveckling: **31 maj**.<sup>7</sup>
- Europaparlamentets omröstning i plenum ("general purpose AI" och "foundation models" utgör utvecklade delar): **14 juni**.
- Den interinstitutionella triloggen mellan Europakommissionen, rådet och parlamentet för att enas om en slutgiltig AI-förordning pågår: **juni 2023**–.

Utvecklingen har också föranlett en konceptuell utveckling – det har inte upplevts som riktigt rättvisande att prata om stora *språkmodeller* enbart (från engelskans *large language models*, LLM) eftersom uttrycken som genereras sträcker sig långt bortom text och språk.

Stanforduniversitetets institut för *Human-Centered Artificial Intelligence* (HAI) publicerade i augusti 2021 (dvs. efter kommissionens första förslag på AI-förordning) en rapport om *foundation models*, dvs. vad vi kan kalla grundmodeller (Bommasani m.fl., 2022). Här definierar de grundmodeller som "vilken modell som helst som tränas på breda data (vanligtvis med självövervakning i stor skala) som kan anpassas (t.ex. finjusteras) till ett brett spektrum av nedströmsuppgifter."

Rapporten anger exempel som Googles BERT, GPT-3 och CLIP. Rapportförfattarna konstaterar att grundmodellerna ur teknisk synvinkel inte är nya – de är baserade på djupa neurala nätverk och självövervakat lärande, som använts i decennier. Men de senaste få årens exempel har visat på en skala och omfattning som gett en kapacitet som är förbluffande. Till exempel har GPT-3 175 miljarder parametrar och kan anpassas via naturliga språkuppmaningar (så kallade *prompter*) för att ge ett trovärdigt resultat på ett brett spektrum av uppgifter trots att den inte är explicit utbildad för att utföra många av dessa uppgifter. Samtidigt har befintliga grundmodeller potential att reproducera eller förstärka skador eller ojämlika och orättvisa samhällsstrukturer, vilket även konstateras i stanfordrapporten, och många av deras egenskaper är i allmänhet underbeforskade (för en teoretisk modell kring hur sociala orättvisor kan riskera att speglas i adaptiva eller lärande teknologier, se Larsson m.fl., 2023). Med tanke på AI-grundmodellernas

---

<sup>7</sup> Taxonomin finns på kommissionens hemsida, se Europakommissionen (31 maj 2023) EU-U.S. Terminology and Taxonomy for Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence> [senast besökt 6 november 2023].

omfattande spridning har de blivit ett ämne för intensiv granskning (jfr Bender m.fl., 2021).

### 2.1.1 2023: en ny våg av policyidéer för generativ AI

Utvecklingen inom generativ AI har också föranlett en rad aktörer att ta fram policyer och rekommendationer under 2023 för hur grundmodeller eller flerändamåls-AI bör hanteras. Orsakerna är troligen i vart fall delvis sammanvävda med AI-förordningen, där en del aktörer söker spela in rekommendationerna till den pågående förhandlingen om slutgiltig AI-förordning, men en rad betydande aktörer pekar på AI-utmaningar som de ser behöver hanteras, oaktat vad som sker inom EU.

Ett axplock från betydande aktörer:

- Japan: LDP Headquarters for the Promotion of Digital Society publicerade i april 2023 *The AI White Paper. Japan's National Strategy in the New Era of AI*.<sup>8</sup>
- World Economic Forum bjöd i april in en rad experter som bistod i att ta fram en policyrapport som publicerades i juni 2023 med titeln *The Presidio Recommendations on Responsible Generative AI*.<sup>9</sup>
- OECD publicerade *Initial policy considerations for generative artificial intelligence* i september 2023. Här pekar de bl.a. på hur generativ AI kan ge stora värden, med revolutionerande konsekvenser för industrier, men även ge risker. Dessa pekas på i termer av felinformation, reproduktion av bias, dvs. diskriminerande strukturer, utmaningar mot upphovsrättsligt ägande och några andra policyrelevanta frågor.<sup>10</sup>
- Teknikprofessionsorganisationen ACM:s teknologipolicyråd publicerade i juni 2023 *Principles for the development, deployment, and use of generative AI technologies*, där de bl.a. pekar på behovet av reglering och införandet av skyddsmekanismer, med human-in-the-loop, utmaningar med frågor kopplat till immaterialrättsligt ägande, dataskydd och behov av möjligheter och mekanismer för att korrigera felaktigheter.
- President Biden utfärdade en presidentorder om säker och tillförlitlig artificiell intelligens (31 oktober 2023).<sup>11</sup>
- G7 publicerade AI-riktlinjer för en "Hiroshimaprocess" för avancerade AI-system samt en uppförandekod för utvecklareorganisationer (30 okt, 2023).

<sup>8</sup> Policyn finns att hitta på LDP-politikern Masaaki Tairas officiella webbsida, se LDP Headquarters for the Promotion of Digital Society (april 2023) *The AI White Paper. Japan's National Strategy in the New Era of AI*. Project Team on the Evolution and Implementation of AIs. [https://www.taira-m.jp/ldp/s%20ai%20whitepaper\\_etrans\\_2304.pdf](https://www.taira-m.jp/ldp/s%20ai%20whitepaper_etrans_2304.pdf) [senast besökt 6 november 2023].

<sup>9</sup> World Economic Forum (juni 2023) *The Presidio Recommendations on Responsible Generative AI*. White Paper. <https://www.weforum.org/whitepapers/the-presidio-recommendations-on-responsible-generative-ai/> [senast besökt 6 november 2023].

<sup>10</sup> OECD (september 2023b) *Initial policy considerations for generative artificial intelligence*, OECD. [https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence\\_fae2d1e6-en](https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en) [senast besökt 6 november 2023].

<sup>11</sup> The White House (30 oktober 2023) FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> [senast besökt 6 november 2023]

- De länder som närvarade vid Storbritanniens AI Safety Summit skrev den 1 november under The Bletchley Declaration som bl.a. efterlyser internationellt samarbete för att hantera AI-risker, och pekar särskilt ut flerändamåls-AI och AI-grundmodeller.<sup>12</sup>

Utöver dessa kan även NIST nämnas, dvs. det amerikanska nationella institutet för teknologistandardisering (januari),<sup>13</sup> och Storbritanniens konkurrens- och marknadsmyndighet, CMA (september).<sup>14</sup> Sistnämnda, från CMA, sätter även fingret på hur specifika marknadsdomäner och deras tillsynsmyndigheter kan förstå betydelsen av generativ AI. Detta är även en poäng värd att beakta i relation till svenska förvaltningsmyndigheter, kanske särskilt de med tillsynsansvar.

Dessa policyer understryker att vi befinner oss i en formativ period gällande styrning och reglering av generativ AI. Det märks också i de svårigheter Europaparlamentet under våren 2023 tycks ha haft att kunna hantera utvecklingen med stora AI-grundmodeller och flerändamåls-AI som Chat GPT i relation till sitt förslag på AI-förordning, vilket jag återkommer till nedan.

För europeiskt vidkommande gällande tidsaspekter kan det konstateras att lagstiftningen i AI-förordningen blir tillämplig först 24 månader efter ikraftträdandet (enligt den troligen inte omförhandlade artikel 85). Det betyder att även om en slutgiltig lagstiftningstext är överenskommen i december 2023, så träder den i kraft 20 dagar efter offentliggörandet i EU:s officiella tidning, och blir tillämplig som helhet först under första halvåret 2026, även om delar kan träda ikraft tidigare.

AI är idag redan en teknik som i olika utsträckning integrerats i olika tekniska lösningar och tjänster. Genomslaget förväntas bli än större i närtid och föra med sin rad komplexa risker och osäkerheter – och möjligheter.

## 2.2 Omvärldsutblick av relevans för Europa och taktproblemet

Den globala policyrelaterade utvecklingen på AI-området är för omfattande för att kunna presenteras djuplodande här, men givet rapportens fokus på generativ AI så kan ändå några relevanta nedslag göras i ett litet urval av globalt intressanta aktörer.

### 2.2.1 G7 och generativ AI

Vid G7-toppmötet i Hiroshima i maj 2023 diskuterades bland annat AI-utvecklingen och dess behov av reglering i termer av ett taktproblem:

*Vi ser att även om snabba tekniska förändringar har stärkt samhällen och ekonomier, har den internationella styrningen av ny digital teknik inte nödvändigtvis hållit jämna steg. När takten i den tekniska utvecklingen accelererar, bekräftar vi vikten av att ta itu med*

---

<sup>12</sup> AI Safety Summit (1 november 2023) The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> [senast besökt 6 november 2023].

<sup>13</sup> NIST (26 januari 2023) The NIST AI Risk Management Framework. <https://www.nist.gov/itl/ai-risk-management-framework> [senast besökt 6 november 2023].

<sup>14</sup> Competition & Markets Authority (september 2023) *AI Foundation Models: Initial Report*.

*gemensamma styrningsutmaningar och att identifiera potentiella luckor och fragmentering i global teknikstyrning (G7, 2023, s. 27, författarens översättning).*

G7-ledarna enades om att inrätta en "Hiroshima-AI-process" för att reglera AI, och adresserar *generativ AI* specifikt. Regeringarna ska enligt kommunikén hålla diskussioner "på kabinetsnivå" om frågan och presentera resultaten i slutet av året. Oron enligt det gemensamma uttalandet är att framstegen inom generativ AI kan användas för desinformation och politisk störning. Men synen på åtgärder skiljer sig åt mellan länderna. Japan lutar åt att styra AI med mjukare riktlinjer, till skillnad från EU med förslaget på AI-förordning. G7-kommunikén från maj, följdes upp av en OECD-rapport i september 2023 där man hade intervjuat G7-länderna gällande deras syn på AI-frågorna och reglering (OECD, september 2023b), vilket också bidragit till både riktlinjer och en uppförandekod (G7, 30 oktober 2023).

## 2.2.2 USA och mjuk styrning

Diskussionerna kring AI-reglering har pågått sedan Obama-administrationen 2016, som började fokusera på riskerna och reglerna för artificiell intelligens.<sup>15</sup> USA har haft en National Security Commission on Artificial Intelligence fram till 2021, och har ett lagförslag om att upprätta ett federalt initiativ utformat för att påskynda forskning och utveckling om AI (The Artificial Intelligence Initiative Act). Den 4 oktober 2022 presenterade president Joe Biden en ny rättighetsförklaring för AI (*AI Bill of Rights*), som tecknar fem skydd som "amerikaner bör ha i AI-åldern": 1. säkra och effektiva system, 2. skydd mot algoritmisk diskriminering, 3. integritetsskydd, 4. notis och förklaring, och 5. mänskliga alternativ och hänsyn (MIT Technology Review, 2022).

Intressant ur ett styrningsperspektiv, och ett tecken på ett mer mjukt tillvägagångssätt än den europeiska, är att i juli 2023 säkrade Biden-Harris-administrationen frivilliga åtaganden från sju företag – Amazon, Anthropic, Google, Inflection, Meta, Microsoft och Open AI – för att hantera riskerna förknippade med AI.<sup>16</sup> Företagen åtog sig att säkerställa att AI-produkter genomgår både interna och externa säkerhetstestningar innan de släpps offentligt – vilket är likt AI-förordningens krav på högrisk-AI. Vidare tog de på sig en rad transparensåtaganden, som att dela information om hanteringen av AI-risker, att utveckla mekanismer för att informera användare när innehåll genereras av AI (exempelvis s.k. vattenmärkning), att offentligt rapportera om deras AI-systems kapacitet, begränsningar och användningsområden. I september 2023 tecknade ytterligare åtta företag dessa frivilliga åtaganden – Adobe, Cohere, IBM, Nvidia, Palantir, Salesforce, Scale AI och Stability.<sup>17</sup>

---

<sup>15</sup> Se här till exempel The White House (2016) The Administration's Report on the Future of Artificial Intelligence. <https://obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence> [senast besökt 5 oktober 2023]

<sup>16</sup> Se här The White House (21 juli 2023) FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> [senast besökt 5 oktober 2023].

<sup>17</sup> Se The White House (12 september 2023) FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Eight Additional Artificial Intelligence Companies to Manage the Risks Posed by AI. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris->

Frågan är om riktningen ändras med tanke på den presidentorder som Biden undertecknade den 30 oktober 2023. Det kan bero på hur kongressen hanterar AI-frågor framöver, och det kan bero på hur nästa val går, eftersom nya presidenter brukar se över tidigare presidenters presidentorder. Presidentordern är dock så nära en övergripande policy som har producerats i USA, med åtta olika sektorer som täcks: nationell säkerhet, individuell integritet, rättvisa och medborgerliga rättigheter, konsumentskydd, arbetsmarknadsfrågor, AI-innovation och USA:s konkurrenskraft, internationellt samarbete om AI-policy, och AI-kunskaper och expertis inom den federala regeringen (White House, 30 oktober 2023).

Gällande generativ AI kan noteras att Sam Altman, vd för OpenAI, bjöds in till den amerikanska senaten den 16 maj 2023. Altman förespråkade ett antal styrningsförslag – inklusive en ny statlig myndighet med uppgift att skapa standarder för området – för att ta itu med växande farhågor om att generativ AI skulle kunna förvränga verkligheten och skapa oöverträffade säkerhetsrisker. Vd:n räknade upp riskfulla beteenden som presenteras av teknik som Chat GPT, inklusive spridning av individualiserad "interaktiv desinformation" och känslomässig manipulation (Washington Post, 16 maj 2023). Man kan dock notera att Open AI också har lobbat för en mildare reglering, under 2022 (TIME, 20 juni 2023).

Utfrågningen i senaten inkluderade också Gary Marcus, en f.d. professor i psykologi och neurovetenskap vid New York University som också grundat AI-företag som sålts till Uber. Han ifrågasatte OpenAI:s brist på transparens om vilka data den använder för att utveckla systemen och konstaterade att

*Vi har oöverträffade möjligheter här men vi står också inför en perfekt storm av företagens ansvarslöshet, utbredd spridning, brist på adekvat reglering och inneboende opålitlighet (Senatsförhör, 16 maj, 2023, författarens översättning).<sup>18</sup>*

Snabbheten i AI-utvecklingen ställs därmed mot säkerhetsfrågor och riskerna med att inte ta dessa tillräckligt i beaktande. Dvs. att konkurrenssituationen i vad som beskrivs som en formativ period för stora språkmodeller inom AI-utveckling får de stora teknikjättarna att lansera produkter som kan få ödesdigra konsekvenser som de inte kan överblicka. Vi återkommer till transparens- och dokumentationsfrågor gällande grundmodeller nedan.

### 2.2.3 Storbritannien

Den brittiska regeringen har tagit en rad politiska initiativ under de senaste åren för att utveckla sin strategiska hållning för användningen av AI. Under 2023 har de bl.a.

- investerat 2 miljoner pund för en testversion av sandlådor för att hjälpa företag att testa AI-regler innan de kommer ut på marknaden

---

administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/ [senast besökt 5 oktober 2023].

<sup>18</sup> U.S. Senate Judiciary Committee Subcommittee on Privacy, Technology and the Law. Hearing: "Oversight of A.I.: Rules for Artificial Intelligence." Utskrift av senatsförhöret från den 16 maj, 2023 finns här: <https://techpolicy.press/transcript-senate-judiciary-subcommittee-hearing-on-oversight-of-ai/> [senast besökt 6 november 2023].

- investerat 100 miljoner pund för en "Foundation Model Taskforce"
- lanserat värdskapet för det globala säkerhetsmötet för AI under hösten 2023
- investerat 54 miljoner pund för att öka och utveckla säker och pålitlig AI-forskning.

I augusti 2023 publicerade de ett "pro-innovation"-ramverk som lyfter fram liknande principiella frågor som många andra:

- säkerhet (både *safety* och *security*), och robusthet
- lämplig transparens och förklarbarhet
- rättvisa (*fairness*)
- ansvarsskyldighet och styrning
- bestridbarhet och upprättelse.

Den myndighet som har ansvar för konkurrens- och konsumentfrågor, Competition and Markets Authority, CMA, publicerade i september 2023 en kartläggning av vilka konsekvenser de s.k. AI-grundmodellerna kan ha för deras verksamhetsområden gällande konkurrens och konsumentintressen.

Det globala säkerhetsmötet för AI, *AI Safety Summit*, hölls den 1–2 november 2023. Här lanserades nämnda The Bletchley Declaration. Mötet har dock kritiserats av en rad civilsamhällsorganisationer och forskare som i ett öppet brev till premiärminister Rishi Sunak kallar AI Safety Summit för "ett missat tillfälle". Kritikerna menar att AI-mötets fokus på "frontier AI" och risker med missbruk av AI och kontrollförluster är för ensidigt och okritiskt mot de teknikjättar som är närvarande vid mötet. Undertecknarna menar att bredare samhällsutmaningar med diskriminerande effekter, övervakning eller att mindre företag trycks ut tycks inte fokuseras, och de intressenter som drabbas av sånt är heller inte inbjudna (Open Letter, november 2023).

#### 2.2.4 Japan

I kontrast till hur den europeiska AI-förordningen är upplagd, tycks Japan mer inriktat mot vad Theo Andersson kallar en "självregleringsprincip" i en rapport från Digg (2022, s. 9), som dock också förutsätter transparens. I april 2023 publicerade regeringspartiet LDP:s projektgrupp (om utveckling och implementering av AI) en AI-vitbok med titeln "Japans nationella strategi i en ny tidsålder av AI". Här lämnas tre huvudsakliga rekommendationer:

1. behovet av att utveckla en ny nationell AI-strategi för den "nya tidsåldern av AI"
2. odla och stärka inhemsk kapacitet för AI-utveckling
3. främja användning av AI i myndigheter.

Man uttrycker i dokumentet, i kontrast till den europeiska lagstiftningsprocessen, att man sedan åtminstone 2019 i Japan tydligt har haft inriktningen att regleringen inte får "stå i vägen" för innovation, och att ramverk som föreslås inte är bindande (man lanserade 2019 bl.a. principer för ett människocentrerat AI-samhälle, se Larsson, 2020). Man uttrycker dock ett växande "regelgap" mellan Japan och övriga länder, som kan vara till nackdel för den japanska marknaden. Givet detta, tycks det finnas en motsättning mellan den inhemska politiken och hur man uttrycker sig inom G7, utvecklat ovan.

Den 24 april 2023 höll den japanska regeringen det första mötet i ett nytt AI-strategiteam. Den 11 maj lanserades ett AI-strategiråd i premiärministerns kansli som är tänkt att fungera samlande för AI-politik. Med utgångspunkt i att använda och sprida bl.a.



generativ AI, kommer regeringen att undersöka diskussioner om lämpliga bestämmelser som skydd av personlig information och fastställande av riktlinjer för dess användning.

## 2.3 Taktproblemet: mellan rättslig förutsägbarhet och teknologisk förändring

Taktproblemet är inte nytt, även om teknikinnovationen ofta känns accelererande. En grund för analysen i denna rapport ligger vid en tämligen enkelt uttryckt dialektik mellan *flexibilitet* och *fixering*. Eftersom det huvudsakliga undersökningsobjektet är de idéer som ingår i processen att ta fram den europeiska AI-förordningen, letar också analysen efter olika rättsliga verktyg däri som kan uttrycka flexibla inslag. Detta i ljuset av det uttryckta taktproblemet, inom en i övrigt relativt okritiserad idé om att juridikens uppdrag är att normativt fixera vad som ska gälla. Fixeringen spelar en roll inte minst i relation till marknadens behov av förutsägbarhet.

### 2.3.1 Fixering: samhällets behov av rättslig förutsägbarhet

Rättsutvecklingen har länge analyserats i termer av just förutsägbarhet, vilket ligger i en av taktproblematikens vågskålar. Förutsägbarhet beskrivs av rättsvetare Alexander Peczenik som en av de grundläggande demokratiska värderingarna i och en stat som styrs av lagstiftning. Många tongivande rättsvetare har genom åren pekat på just det principiellt viktiga i att offentliga myndigheters agerande i en demokratisk stat bör vara förutsägbart (Peczenik, 1995, s. 89–90). Den norske rättssociologen Vilhelm Aubert, exempelvis, ser förmågan att värna om förväntningar som en av rättens fem huvuduppgifter, och den tyske rättsvetaren Niklas Luhmann har å sin sida hävdad att det är dess viktigaste. Denna betoning på förutsägbarhet är särskilt intressant i ljuset av AI-förordningens försök att vara flexibel och "framtidssäker" med AI-definitionen, och hänskjutandet av vissa nyckelfrågor till ännu obefintliga tekniska standarder, och idéer om regulatoriska AI-sandlådor.

Samtidigt förändras samhällen, ofta genom ny teknik som erbjuder nya beteendemönster och sätt att genomföra transaktioner m.m. Det gör att även juridiken i sin fixering ständigt har behov av ett mått av samhällelig anpassning, dvs. flexibilitet. I introduktionen anfördes att även den traditionella rättsordningen tenderar att ha metoder för viss flexibilitet kopplad till hur lagstiftningens bokstav är fixerad. Jag visade exemplet med en mer än 100-årig avtalslag som är gällande och fungerande just för att domstolens uttolkning genom praxis har möjliggjort förtydliganden av hur de underliggande principerna ska förstås i ljuset av till exempel nya teknikmedierade sätt att sluta, eller bryta, avtal. Avtalsslutet regleras även för bussbiljetten som blippas med en smarttelefon eller läskburken som beställs i varuautomaten, oavsett att inget sådant existerade vid lagens tillblivelse. Även här bjuder den äldre rättsvetenskapen på exempel i Karl Renners analys av äganderättens oförändrade formulering genom det enormt förändrade samhälle som kom ur Västeuropas industrialisering (2010/1949, se kommentar i relation till AI i Larsson, 2020).

När det gäller AI-utvecklingens behov av styrning kan man fråga sig om vissa delar av mjuk styrning, som etiska ramverk, verkligen är reglering. Det riskerar dock att missa huvudsyftet här eftersom *styrningen* är det som fokuseras, inte nödvändigtvis *reglering*, i termer av lag som uppkommit enligt en viss typ av ordning. Även de allra mjukast formulerade ramverken gör trots allt någon typ av normativt anspråk på hur saker och

ting bör ske. Det är den normativiteten som är i fokus, men frågan måste ställas i relation till taktproblemet, dvs. hur kan något snabbt föränderligt styras bra? Vilka styrformer är rimliga? Och här erbjuder litteratur inom fältet för taktproblem en del idéer.

### 2.3.2 Flexibilitet: adaptiv styrning som lösning på taktproblem

En tidigare rapport från Tillväxtanalys (Birksjö m.fl., 2022) redovisar kunskapsläget om skillnader i utvecklingstakt mellan reglering och teknisk innovation. Exempel på policyverktyg som redovisas för en adaptiv styrmodell är:

- rutiner för kontinuerlig teknikbevakning och riskbedömningar
- metoder för att testa teknik under kontrollerade former såsom regulatoriska sandlådor
- bestämda tidsramar för regelförbättrande åtgärder, så kallade solnedgångsklausuler
- sam- eller medreglering<sup>19</sup>
- mjuk reglering, i termer av frivilliga överenskommelser, standardiseringsprocesser eller — något som blivit vanligt inom AI-området — etiska riktlinjer.

Regulatoriska sandlådor är avsedda att skapa en säker testbädd för innovation, antingen genom att tillåta tillfällig tillämpning av en annan reglering på en liten grupp av företag eller genom att erbjuda vägledning för efterlevnad (Ranchordas, 2021). Noterbart för svenskt vidkommande är att Integritetsskyddsmyndigheten (IMY) har haft pilotprojekt med regulatorisk testverksamhet, med fördjupad vägledning till ett projekt om hälsorelaterat federerat lärande, hur dataskyddsregelverket kan tolkas och tillämpas (IMY, 2023).<sup>20</sup>

Medan *adaptiv reglering* syftar till att främja regelförändringar och stödja innovation genom att anpassa befintliga regelverk, syftar *föregripande reglering (anticipatory regulation)* till att erbjuda en "iterativ utveckling av reglering och en bättre förståelse för teknologins påverkan på samhället" (Nesta 2017; se också Ranchordas, 2021b).

Som en del i samma dynamik mellan innovation och reglering, men för fenomenet när regler inte följs, används ibland begreppet *undvikande entreprenörskap (evasive entrepreneurship)*; se Birksjö m.fl. 2022, från bl.a. Hagemann m.fl., 2018). Skälen till undvikandet kan rimligen utgöras av alltifrån mer renodlade försök att skaffa sig konkurrensfördelar jämfört med de som följer reglerna, till uttryck för en mer utbredd brist på legitimitet i rådande regelverk, som därmed också kan komma att (behöva att) revideras. För det förre har ibland stora plattformaktörer hänvisats till, som i att de är "too big to ban" (jfr Birksjö m.fl., 2022, s. 25), och till det senare har analyser kring fildelning och upphovsrätt från tidigt 00-tal (jfr Palmås m.fl., 2014) visat att de första innovatörerna (som The Pirate Bay) ibland kan bana väg för efterföljande entreprenörskap (som Spotify).

<sup>19</sup> Vilket förvisso inte tycks förekomma så ofta i taktproblemslitteraturen, men som är ett passande komplement eftersom det också visar på element i till exempel Rättsakten om digitala tjänster, som vi återkommer till nedan, med ett slags tvingande framtagande av uppförandekoder.

<sup>20</sup> Vilket myndigheten har rapporterat om, se IMY (15 mars 2023) "Framgångsrik pilot med regulatorisk testverksamhet om AI", <https://www.imy.se/nyheter/framgangsrik-pilot-med-regulatorisk-testverksamhet-om-ai/> [senast besökt 5 oktober 2023].

Bar-Siman-Tov & Harari-Heit (2020, s. 245) ser till exempel fördelar med *tillfällig reglering* jämfört med permanent reglering för områden som är dynamiskt föränderliga. Det finns också exempel på rättsliga moratorier. Med den brittiska cloning law (bill 5758-1998), som i mycket togs fram i respons på de möjligheter till kloning som tydliggjordes med det klonade fåret Dolly i Skottland i slutet av 1990-talet, infördes ett femårigt moratorium gällande genetiska interventioner på människor. Orsaken var att man ansåg sig behöva mer tid för att undersöka moraliska, rättsliga, etiska, sociala och andra vetenskapliga följder av kloning (se Bar-Siman-Tov & Harari-Heit, 2020).

Som en del av Komets (Kommittén för teknologisk innovation och etik) utredning om regler och teknikutveckling (2021) togs en rad problembilder fram i skärningen mellan reglering och teknisk innovation. En av dessa träffar kärnan för analysen i denna studie, nämligen utmaningen med en målkonflikt mellan stabilitet och flexibilitet i regelverken vid snabb teknikutveckling (s. 19). Kometutredningen inkluderar här två särskilda kategorier:

- svårighet att få balans mellan kontinuitet och förutsägbarhet i förhållande till flexibilitet när en lag, förordning eller föreskrift tillkommer eller ändras
- teknikutvecklare upplever att beslutsfattarna inte ser till att förändring av regelverk interagerar med den snabba teknikutvecklingen.

Vilket ger att Kometutredningen i någon mening fångar upp en huvudsaklig tanke från rättsvetenskapliga sätt att beskriva rättens roll, och ställer den mot de behov som uppstår med snabb teknikutveckling.

### **2.3.3 Om juridisk flexibilitet: experimentell reglering**

Sofia Ranchordás, professor vid Tillburgs universitet och intervjuad för denna kartläggning, är kanske mest känd för sitt arbete inom experimentell reglering i pilotprojekt och regulatoriska sandlådor. Hon har konstaterat att det finns en lång tradition av experimentella policyer och reglering, åtminstone sedan franskt 1800-tal (Ranchordás, 2014). Precis som Bennet Moses (2010) har hon också konstaterat att innovation, i dess mer nutida mening, är ett regulatoriskt rörligt mål som inte passar bra med traditionella och i första hand reaktiva rättsliga ramverk. Hon har argumenterat för behovet av att få till en *god tajmning* på regleringsinitiativ i relation till innovation (2015), bland annat med tillfällig reglering.

Experimentella regleringar är potentiellt attraktiva för både tillsynsmyndigheter och innovatörer eftersom de är tänkta att skapa säkra förutsättningar för att främja innovation. Man vill minska vad som ibland beskrivs som en "regelbörda", förbättra kommunikationen mellan företag och tillsynsmyndigheter och samtidigt främja snabba regleringsbeslut utan att utsätta konsumenterna för risker (Buckley m.fl., 2019; Knight & Mitchell, 2020).

Experimentell lagstiftning har dock sina utmaningar. Ranchordas (2014) har visat att det finns brister i hur sådana experimentella tillvägagångssätt riskerar att bli politiserade, hur till exempel regulatoriska sandlådor ofta avslutas för tidigt samt innehåller otillfredsställande utvärderingar och en allmän frånvaro av metodologiska intressen som kan dra mer generella slutsatser. Hon argumenterar för att experimentella rättsliga system bör utformas på ett öppet, ansvarsfullt, proportionerligt och förutsägbart sätt. Trots den växande litteraturen om regulatoriska sandlådor och regleringen av AI så

kvarstår enligt Ranchordas (2021a) mycket att beforska och reda ut kring de juridiska, metodologiska och etiska utmaningarna med regulatoriska sandlådor.

Hon har i en analys av experimentell reglering och regulatoriska sandlådor sett viktiga metodologiska brister i experimentella bestämmelser och regulatoriska sandlådor (Ranchordás, 2021b). Hon ser tre negativa effekter: För det första är den interna giltigheten av experimentella rättsliga regimer begränsad eftersom det är oklart om de verifierade resultaten är det direkta resultatet av den experimentella interventionen eller andra omständigheter. Den begränsade externa giltigheten hos experimentella rättsliga regimer hindrar experimentets generaliserbarhet. För det andra ger experimentella rättsliga regimer som inte är vetenskapligt sunda ett begränsat bidrag till att främja evidensbaserad lagstiftning och rationalisering av regleringen. För det tredje kan metodologiska brister leda till brott mot rättsliga principer som kräver att experimentella regler följer objektiva, transparenta och förutsägbara standarder. Hon argumenterar därmed för behovet av mer transparent och objektiv design av experimentella regler och regulatoriska sandlådor.

Vi återkommer till flera aspekter av experimentell reglering och adaptiv styrning i analysen nedan. I relation till AI-förordningen ställs fixering mot element av flexibilitet mot varandra, och det sistnämnda syns framför allt i saker som standardisering kopplad till regleringen, möjligheten att föreskriva att kommissionen återkommer till förtydliganden av valda delar av förordningen genom så kallade delegerade akter och genomförandeakter samt utpekandet av regulatoriska sandlådor. AI-området i vidare mening har sedan några år kännetecknats av en stor rad riktlinjer av principiell eller etisk karaktär, vilket vi tittar närmre på i följande kapitel.

## 3. EU och AI

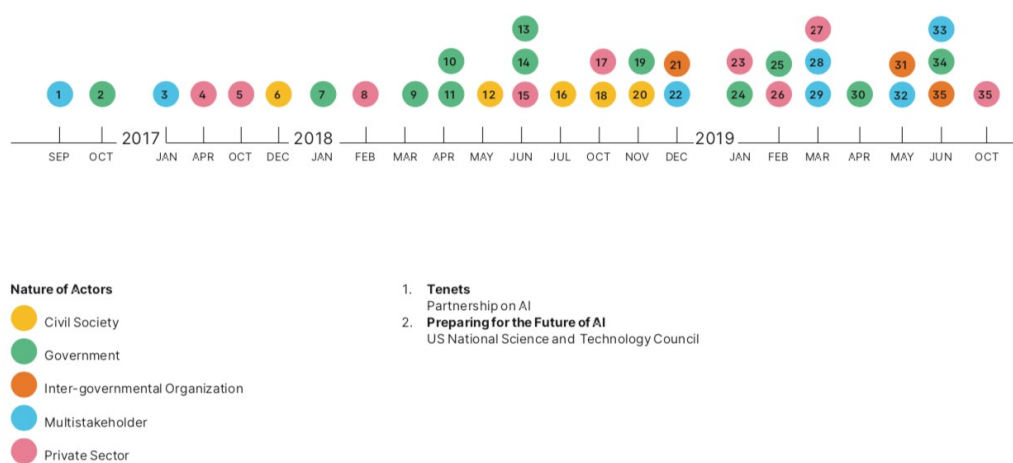
### 3.1 Mjuk AI-styrning

Det tycks finnas två vågor av mjuk AI-styrning: dels från de få åren innan kommissionens förslag på AI-förordning, dels under 2023, som en följd av utvecklingen inom generativ AI, vilket redovisats ovan. Här beskrivs den första vågen, som en upptakt till AI-förordningen.

#### 3.1.1 2016–2020: medvetenhet om AI-risker

Åren 2016 till omkring 2020 kan beskrivas som väldigt intensiva när det gäller framtagandet av etiska riktlinjer och principiella ställningstaganden gällande AI-utveckling och AI-tillämpning. En medvetenhet om inte bara möjligheter utan även risker för till exempel diskriminerande praktik i allt från AI-system för ansiktsgenkänning (Buolamwini & Gebru, 2018) till sökmotorer (Noble, 2018), myndighetsbeslut (Richardson, 2021) och HR-system (Ajunwa, 2023) spred sig. Kartläggningar av den här typen av principiella dokument och policyer visar på hur både företag, stater såväl som civilsamhället formerade AI-riktlinjer (Fjeld m.fl. 2020; Jobin m.fl. 2019), se Figur 1. Etik, tycks det, blev ett sätt att styra och utveckla styrning av AI (Larsson, 2020). Så även inom EU (Larsson, 2021), se Figur 2.

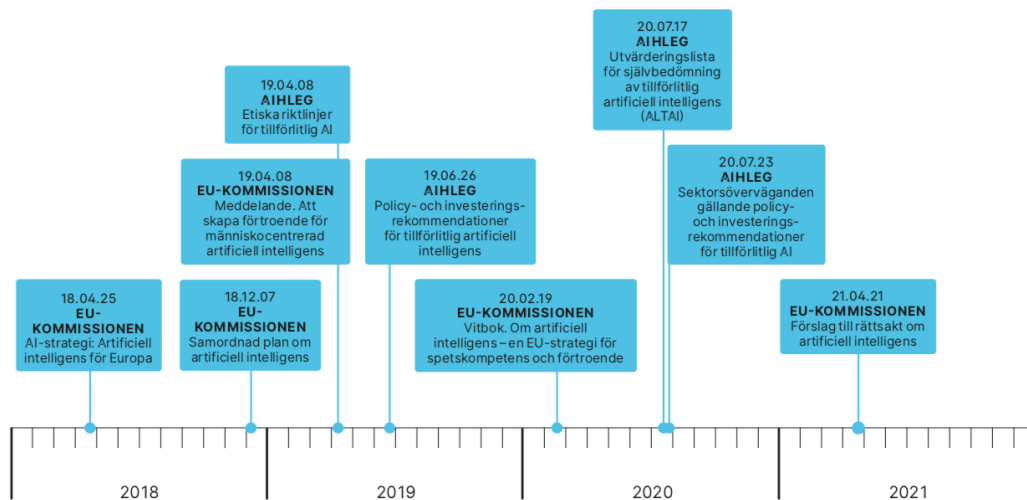
Figur 1 Tidslinje med AI-riktlinjer och principiella dokument. Adapterat från Fjeld m.fl., 2020.



En sådan kartläggning av riktlinjer (Jobin m.fl., 2019) visar att de vanligaste frågorna som tas upp i riktlinjerna handlar om (1) behovet av transparens (73 av de 84 källorna pekade på just detta), (2) rättvisa, eller det svåröversatta *fairness*, (3) icke-skadlighet, (4) ansvarsfördelningsfrågor och (5) integritet eller dataskydd. Det betyder dock inte att det fanns en enhetlighet i hur dessa begrepp används.

Alltsedan kommissionen publicerade sin AI-strategi under 2018 har arbetet med ett slags dubbel hantering av AI-frågorna pågått, av å ena sidan att försöka stimulera innovation, och å andra sidan begränsa risker, för en tidslinje, se Figur 2. Detta är tydligt även i den vitbok om AI som kommissionen publicerade 2020.

Figur 2 Tidslinje med EU-kommissionens AI-arbete fram till förslaget på en AI-förordning. Adapterad från Larsson & Ledendal, 2022.



Kommissionen inrättade en AI-expertgrupp (HLEG) som under 2019 och 2020 publicerade åtminstone fyra betydande rapporter, där *Etiska riktlinjer för tillförlitlig AI* tycks ha haft särskilt betydande roll för AI-styrningsområdet (för en genomgång av relevansen för svensk myndighetspraktik, se Larsson & Ledendal, 2022). Dessa riktlinjer nämns också explicit i skäl 13 i parlamentets förslag på AI-förordning.

AI-transparens är ett tämligen mångfacetterat område (Larsson & Heintz, 2020) som både handlar om AI-modellers förklarbarhet, hur förmedlingen går till i sig, såväl som mottagarens litteracitet eller informationsfärdigheter (Larsson m.fl., 2023). Samtidigt ligger mycket av vägen mot att bygga bra produkter och korrigera felaktigheter samt även behov av ansvarsutkrävande i just hur man säkerställer transparenta AI-modeller. Detta är tydligt inte minst i hur parlamentet i sin version av AI-förordning ser på generativ AI, och hur de s.k. grundmodellerna bör regleras genom omfattande transparenskrav, vilket vi återkommer till nedan.

### 3.2 2021: förslag på AI-förordning

Europeiska kommissionen föreslog i april 2021 den förordning som fastställer harmoniserade regler om artificiell intelligens (AI-förordningen). Förordningen, särskilt när det gäller AI-system med hög risk, ingår i vad som kallas "den nya lagstiftningsramen" (*New Legislative Framework, NLF*) som innebär ett slags produktsäkerhetsreglering. Med det följer till exempel ett tydligt fokus på bedömningar av överensstämmelse (*conformity assessment*) som måste genomföras innan en AI-produkt placeras på marknaden. CE-märkning är därmed en central del. Förordningen ska dock inte tillämpas på högrisksystem som regleras i andra riktade regelverk (till exempel civil luftfart och självkörande fordon), eller på AI-system som utvecklats eller används uteslutande för militära ändamål.

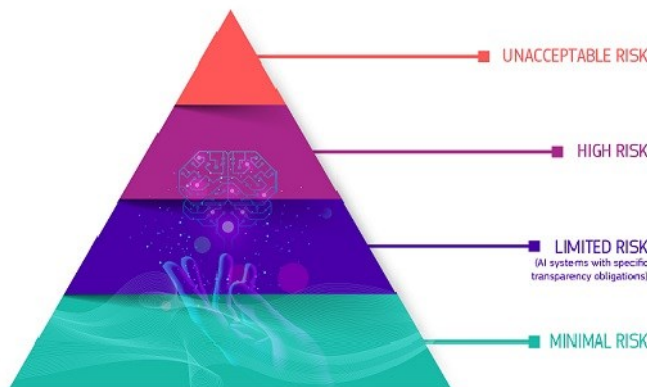
Efter en stor mängd ändringsförslag samt rådets version, dvs. allmänna inställning till förslaget, publicerat den 6 december 2022, och parlamentets bekräftade position från den 14 juni 2023 så pågår i skrivande stund förhandlingar för att nå en slutgiltig version. Förordningen är för omfattande och komplex för att någon djupare genomgång skulle

tillåtas här, men i korthet kan sägas något om de huvuddrag som kännetecknar förordningen.

De föreslagna reglerna fastställer skyldigheter för leverantörer och användare beroende på risknivån förknippad med olika sorters AI-system. Förordningen gäller leverantörer som släpper ut AI-system på marknaden och tar dem i bruk i unionen, oavsett om dessa leverantörer är etablerade inom unionen eller i ett tredjeland. Den består av 89 skäl, 85 artiklar och sju (kommissionens) eller sex (rådet och parlamentets) bilagor.

Indelningen av risknivåer är en huvudsaklig struktur för att sortera i och fördela de olika kravställningarna. Vitboken, nämnd ovan, hade en mer binär uppdelning, som bland annat kritiserades av den tyska regeringen. I Tyskland hade man publicerat en flernivåindelning i en nationell kommission, som tycks ha påverkat AI-förordningens upplägg (Larsson, 2021).

Figur 3 En riskbaserad ansats. Från kommissionens hemsida.



De olika risknivåerna beskrivs som oacceptabel risk, hög risk, begränsad risk, eller ingen risk, se Figur 3. *Oacceptabla risker* kännetecknar system som anses vara ett hot mot människor och därmed kommer att förbjudas. De inkluderar kognitiv beteendemanipulation av människor eller specifika utsatta grupper, social poängsättning, dvs. klassificering av människor utifrån beteende, socioekonomisk status eller personliga egenskaper och biometriska identifieringssystem i realtid och på distans, såsom ansiktsgenkänning. Vissa undantag kan tillåtas.

AI-system som negativt påverkar säkerhet eller grundläggande rättigheter kommer att betraktas som *högrisk*. Alla högrisk-AI-system kommer att utvärderas innan de släpps ut på marknaden och även under hela deras livscykel.

AI-system med begränsad risk bör uppfylla minimala krav på öppenhet som gör det möjligt för användare att fatta välgrundade beslut. Efter att ha interagerat med applikationerna kan användaren sedan bestämma om de vill fortsätta använda dem. Användare bör göras medvetna när de interagerar med AI. Detta inkluderar AI-system som genererar eller manipulerar bild-, ljud- eller videoinnehåll, till exempel så kallade *deepfakes*.

AI-förordningen föreslås ge möjligheter till höga sanktionsavgifter. Kommissionen föreslår upp till 30 miljoner euro eller upp till 6% av föregående års globala omsättning, som högst i en skala med tre nivåer som beror på vilka krav som inte har uppfyllts.

Parlamentets förslag introducerade fyra nivåer, istället för tre, med upp till 40 miljoner euro eller 7% som högst, om förbuden inte följs, se "oacceptabla risker" ovan (se artikel 71).

För svenskt vidkommande publicerades en faktapromemoria (2020/21:FPM109) den 26 maj 2021 om kommissionens förslag på AI-förordning. Ungefär samtidigt, den 21 maj 2021, remitterades förordningen och remissvaren skulle senast in till Infrastrukturdepartementet den 24 juni 2021. Givet förordningens komplexitet är det troligt att remissinstanserna har haft det svårt att sätta sig in i helheten på så kort tid. Flera reagerade dock på hur bred AI-definitionen var i förslaget, vilket vi återkommer till nedan.<sup>21</sup>

Definitionsfrågan är även ett tecken på taktproblemet, dvs. att innovationen går så snabbt under själva regleringsprocessen att de rättsliga begrepp som är tänkta att fånga det som utvecklas förnyas. Parlamentets version från juni 2023 inkluderar till exempel även "generativ AI", vilket inte finns i de tidigare förslagen från rådet och kommissionen, vilket jag återkommer till i mer detalj nedan. Parlamentet föreslår framför allt en rad transparenskrav som behöver uppfyllas för utvecklare av vad de kallar grundmodeller (*foundation models*), bl.a. att krav på att visa att innehållet har genererats av AI, och krav på sammanfattningar av vilka data som använts för modellträning, om de är upphovsrättsskyddade m.m.

### **3.3 Regleringslandskapet – EU:s formativa teknikregleringsperiod**

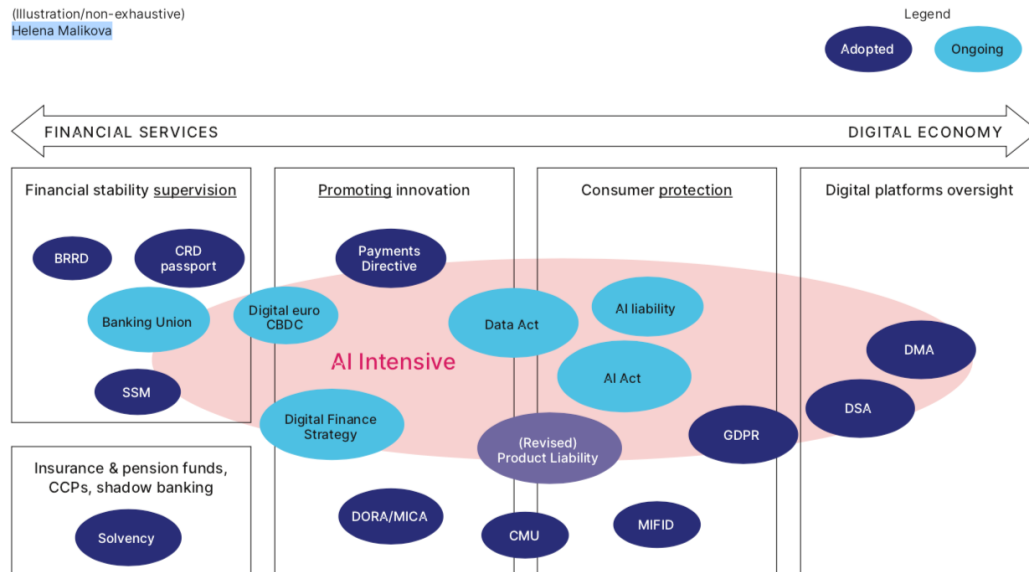
Även om AI-förordningen är komplex och mångbottnad i sig själv, kan betydelsen av rättsakten för innovation och samhälle inte helt förstås utan att man placerar in den i en större trend som har med EU:s digitala årtionde att göra. Utöver att det finns många överlappande redan befintliga regleringar att ta hänsyn till, pågår det också samtidigt, delvis parallellt, en rad lagstiftningsprocesser som har med överlappande frågeställningar att göra. Ramarna för denna rapport tillåter inte någon fördjupning, men det är av vikt att, om än bara översiktligt, placera in AI-förordningsprocessen i ett större europeiskt regleringslandskap, se fig. 4.

---

<sup>21</sup> Remissvaren finns samlade på regeringskansliets webb, se <https://www.regeringen.se/remisser/2021/05/remiss-av-europeiska-kommissionens-forslag-till-forordning-om-harmoniserade-regler-for-artificiell-intelligens/> [senast besökt 6 november 2023].



Figur 4 AI-förordningen som en del i en större europeisk lagstiftningstrend på det digitala området. Från Bryson (2023) :<sup>22</sup>



### 3.3.1 AI och digitalisering: ett regleringslandskap i förändring

En rapport med rekommendationer för AI-förordningen, som samlar från både aktivister och akademiker (bl.a. Natali Helberger), kan här få illustrera några av kopplingarna till andra rättsakter (Algorithm Watch, 2023). Rapportförfattarna konstaterar att AI-förordningen endast är en del av det regleringslandskap som styr över AI-system. De efterlyser att AI-förordningen bör ses och utformas i samband med andra relevanta rättsliga ramar och anger särskilt följande:

- Diskrimineringslagstiftning. Lagstiftaren, menar de, ska säkerställa att lagar om icke-diskriminering tillämpas fullt ut på utveckling och driftsättning av AI-system, inklusive vad de kallar "general purpose AI" (GPAI). För svenskt vidkommande kan Diskrimineringsombudsmannens arbete med myndigheters AI-användning i relation till diskriminering nämnas (DO, 2022) och att DO i november 2023 rapporterar ett regeringsuppdrag om AI och risker för diskriminering i arbetslivet (Arbetsmarknadsdepartementet, 2022).
- Rättsakten om digitala marknader (Digital Markets Act), som rapportförfattarna anser bör utforskas i relation till GPAI-leverantörer. Här efterlyser man även ett bredare säkerställande av att konkurrenslagstiftningen tar itu med inflytandet och marknadsmakten hos leverantörer av AI-grundmodeller och förhindrar konkurrensbegränsande metoder.
- Rättsakten om digitala tjänster (Digital Service Act), och då behovet av att klargöra i vilken utsträckning väldigt stora plattformar integrerar flerändamåls-AI, vilket

<sup>22</sup> Bryson, J. (maj 2023) Keynote: "No one should trust AI: How and why we govern production of intelligent artefacts", Rise of AI Conference, Berlin. <https://youtu.be/YgZD74jIMjY?si=Ekn12ycazNK1Ybvj> [senast besökt 12 november 2023].

förleder frågan huruvida skyldigheten att övervaka och mildra systemriskerna sträcker sig även till risker som är resultatet av utvecklingen och införandet av flerändamåls-AI på dessa plattformar;

- Behovet av revisioner i upphovsrätten (särskilt text- och datautvinningsbestämmelserna i artiklarna 3 och 4 i upphovsrättsdirektivet) kan behöva undersökas, vilket även lyfts av OECD (september 2023a) som en av de policyfrågor som väcks av generativ AI, och
- GDPR, och då med argumentet att utöka artikel 22 till att vara tillämplig på beslut som genereras av flerändamåls-AI; och åter betona GDPR-kraven i AI-förordningen, inklusive transparens om träningsdata.

Till listan kan även förordningen om medicintekniska produkter (MDR) läggas. Det är en EU-förordning (2017/745) som ersätter tidigare lagstiftning sedan den 26 maj 2021, och är tänkt att säkerställa säkerheten och prestandan för medicinteknisk utrustning. MDR refereras till i parlamentets version av AI-förordning, skäl 27. Det tycks finnas en del oklarheter i överlappen för AI-produkter som används inom hälsa och sjukvård (se Hauglid & Mahler, 2023).<sup>23</sup>

### 3.3.2 Om ansvarsfördelning: direktiv om skadeståndsansvar

Som utvecklingen med principiella eller etiska riktlinjer kring AI visar, inklusive HLEG:s variant, är ansvarsfördelningsfrågor en central utmaning. I september 2022 lade Europeiska kommissionen fram två förslag som har beskrivits som "den europeiska strategin för AI-ansvar" (jfr Hacker, 2023): ett nytt direktiv om skadeståndsansvar gällande artificiell intelligens och en översyn av produktansvarsdirektivet (PLD). De utgör därmed en viktig komponent i europeisk AI-reglering. Ansvarsförslagen och den föreslagna AI-förordningen är i sin natur sammanflätade: den senare innehåller inga individuella rättigheter för drabbade personer, och de förra saknar specifika, materiella regler om AI-utveckling och användning.

Direktivet om skadeståndsansvar gällande artificiell intelligens syftar till att främja utbyggnaden av pålitlig AI och att förbättra funktionen av den inre marknaden genom en harmonisering av medlemsstaternas lagstiftning om bl.a. bevisbörda vid vissa skadeståndsanspråk som rör AI (se t.ex. Regeringskansliets faktapromemoria 2022/23:FPM8). Direktivet är ett s.k. minimidirektiv och omfattar enbart skadeståndsskyldighet utanför avtalsförhållanden till följd av vårdslöshet hos skadevällaren och påverkar inte allmänna regler om skadestånd.

---

<sup>23</sup> AI-förordningens riskklassificeringssystem för medicintekniska produkter kan enligt Hauglid & Mahler (2023) potentiellt misslyckas med att fånga vissa AI-specifika risker såväl som vissa traditionella risker förknippade med medicintekniska produkter.

## 4. Analys: AI-förordningen som taktproblem

Det har konstaterats att AI-området bjuder på både oförutsägbarhet och snabbhet i innovationstakten, vilket därmed utgör en särdeles påtaglig utmaning för att skapa en förutsägbar och fixerad AI-reglering. Denna dynamik, här uttryckt som ett taktproblem, är särskilt intressant att notera i lagstiftningsprocessen för den europeiska AI-förordningen. Detta kapitel fördjupar i mer detalj ett urval av de särskilt intressanta aspekterna i denna lagstiftningsprocess, för att i efterföljande kapitel lyftas upp till att svara på de frågor som ställts upp i rapportens inledning, som i huvudsak handlar om att erbjuda ett lärande för hur taktproblem kan hanteras. De valda element av AI-förordningen som uttrycker eller medför flexibilitet och analyseras nedan är:

- AI-definitionen, inklusive "generativ AI", "flerändamåls-AI" och "grundmodeller"
- standardiseringens roll och förfarande
- delegerade akter och genomförandeakter
- regulatoriska sandlådor.

Dessa analyseras här i relation till taktproblemet och litteratur som presenterats ovan.

### 4.1 Vad är AI enligt EU? Definitionens konsekvenser

AI-definitionen har varit väldigt omdebatterad sedan den publicerades i april 2021 i kommissionens förslag. För svenskt vidkommande menade både remissinstanser som Linköpings universitet såväl som Lunds att den var så bred att den inkluderade långt mycket mer än vad de flesta ansåg vara AI. Det finns dels en diskussion kring AI-begreppet som sådant, som pågått länge och som är relevant för hela AI-förordningsprocessen, som jag inleder med här, dels en nyare diskussion kring "general purpose AI" och "foundation models", som jag återkommer till nedan.

Definitionen är i detta sammanhang viktig av två huvudsakliga skäl: 1. Den avgör omfattningen på regleringen. Om den är väldigt bred eller generell, framstår AI-förordningen som en närmast allmän teknikreglering riktad mot riskområden, vilket riskerar att bli en kostsam dokumentationsstruktur. Omvänt, om den är väldigt snäv och förenad med dyra krav, riskerar man vad vi ovan kallar ett undvikande entreprenörskap, där ett undvikande av regleringen kan bli en tydlig strategi hos företag. 2. Den utgör ett tydligt tecken på taktproblematik, och är därmed intressant för det fenomen som adresseras i den här rapporten. Om innovationen är så snabb att begreppen som är tänka att formaliseras i regleringen kraftigt behöver omformuleras under lagstiftningsprocessen i sig, visar detta på att regleringen behöver finna vägar för flexibilitet samtidigt som juridiken söker sin fixering. Som tankeexperiment kunde man reflektera över vad som hade hänt om AI-förordningen faktiskt hade följt sin första tidplan och blivit färdigförhandlad i slutet av 2022, samtidigt om grundmodellerna med Chat GPT som den publikt mest kända representanten tog fart. Risken för att regleringen inte hade fungerat tillfredsställande för så generella AI-system hade varit överhängande, med en rad potentiella problem som följde. Nu hann dock parlamentets version inkludera regleringsuttryck även för generativ AI, vilket gör att pågående förhandling kommer att försöka hitta en balans mellan fixering och flexibilitet även för dessa AI-modeller. En av

nycklarna, och utmaningarna, ligger därmed i hur man definierar AI och förstår hur de olika metoderna och teknikerna utvecklas och integreras i produkter på marknader.

Det finns en historisk dynamik i hur begreppsapparaten runt AI varit påtagligt och flexibel sedan dess introduktion under 1950-talet. Och, som jag konstaterar på annan plats (jfr Larsson, 2021; Larsson och Ledendal, 2022), finns det en betydande skillnad i att å ena sidan ha en flexibel begreppsapparat inom dynamiska forskningsdiscipliner, och å den andra syfta till att reglera innovativa industriella marknadspraktiker som idag omsätter miljardbelopp. Historiskt har intelligensbegreppet dels använts i åtminstone 70 olika varianter i olika sorters forskning (Legg & Hutter, 2007), dels haft en tendens att reserveras för *ännu olösta* teknikproblem, som är tänkta att lösas i en nära framtid (jfr Stone m.fl., 2016).

En relativt vanlig approach de senaste åren har varit att närma sig AI-begreppet så som EU-kommissionens strategi från april 2018 gjorde. Här betonades hur AI "avser system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder – med viss grad av självständighet – för att uppnå särskilda mål". Denna definition syns också i flertalet svenska myndighetsrapporter, från bl.a. Digg (2020) och SCB (2020; för vidare analys, se Larsson & Ledendal, 2022).

I kommissionens förslag på AI-förordning definieras AI i artikel 3(1) som "system med artificiell intelligens (AI-system)" med referens till tillhörande bilaga I. Här nämns under rubriken "Tekniker och metoder som omfattar artificiell intelligens" bland annat "logik- och kunskapsbaserade metoder" och "statistiska metoder, bayesisk beräkning, sök- och optimeringsmetoder". Definitionen av AI-system är därmed väldigt bred.

I kommissionens förslag anges dessutom en viktig flexibilitet, där kommissionen i artikel 4 ges rätt att genom s.k. *delegerade akter* vid behov lägga till tekniker i bilaga I (i enlighet med villkoren i artikel 73). Dvs. AI-definitionen kan här komma att utökas genom delegerade akter, vilket är ett lagstiftningsverktyg vi därför återkommer till nedan.

Parlamentets version från juni 2023 tycks lite grann i polemik med kommissionens kritiserade definition poängtera att den måste uppfattas som legitim av de som finns i fältet:

*Begreppet AI-system i denna förordning bör vara tydligt definierat och nära anpassat till det arbete som utförs av internationella organisationer som arbetar med artificiell intelligens för att säkerställa rättssäkerhet, harmonisering och bred acceptans, samtidigt som den ger flexibilitet för att hantera den snabba tekniska utvecklingen på detta område (skäl 6 i parlamentets förslag på AI-förordning).*

Enligt parlamentets definition av AI-system (artikel 3[1]) är det ett "ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som, för uttryckliga eller underförstådda mål, kan generera utdata såsom förutsägelser, rekommendationer eller beslut som påverkar fysiska eller virtuella miljöer". Det liknar därmed – ånyo – den AI-definition som lanserades i kommissionens AI-strategi från 2018, och den definition som tagits fram av OECD.<sup>24</sup> Parlamentet föreslår att kommissionens

---

<sup>24</sup> Ett AI-system, enligt OECD-definitionen "uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models through analysis in an automated

mandat att ändra i definitionen inte ska ske genom delegerade akter utan genom s.k. *genomförandeakter* (artikel 4). Skillnaden i det återkommer jag därför till nedan.

#### **4.1.1 Terminologisk osäkerhet – vad är skillnaden mellan grundmodell och flerändamåls-AI?**

Som nämnt har ett forskningsinstitut vid Stanforduniversitetet föreslagit "foundation models" som begrepp (Bommasani m.fl., 2022). Detta är inte minst intressant eftersom det bara är parlamentets version av förslag på AI-förordningen som använder sig av detta begrepp, och att det inte tycks finnas någon egentlig konsensus i forskarsamhället i hur begreppsapparaten bäst förhåller sig till exempelvis "multi-purpose AI", här kallat flerändamåls-AI, vilket jag återkommer till i analysen av AI-förordningen nedan. Även policyrekommendationerna om generativ AI från World Economic Forum (2023) och Algorithm Watch (2023) efterlyser en enhetligare terminologi. Det kan medföra utmaningar för regleringens efterlevnad, och osäkerhet för marknadsaktörer.

#### **4.1.2 AI-definitionen och högriskklassificering**

Utöver gränsdragningen till förbjudna praktiker så är högriskklassificeringen en väldigt central del av AI-förordningen. Och här finns det en skillnad som tycks betydande mellan kommissionens och parlamentets versioner. Europaparlamentets förslag inför ett extra lager av hänsyn när det gäller att klassificera ett system som högrisk. De system som används i högriskdomäner (som i bilaga III) skulle bara klassificeras som högrisk om de utgör en "betydande risk för skada" för hälsa och säkerhet eller grundläggande rättigheter för personer (eller, i vissa fall, miljön).

Rent praktiskt skulle detta innebära att en leverantör utifrån sin egen bedömning av riskens svårighetsgrad, intensitet, sannolikhet att inträffa och varaktighet kan klassificera sin tjänst som högrisk – och därmed kan meddela en tillsynsmyndighet och vara fri att placera sitt system på marknaden. Som svar kan tillsynsmyndigheten vid behov invända mot ett sådant meddelande inom tre månader från underrättelsen. Detta tycks vara ett element av mjuk styrning inbakat i det annars lite hårdare ramverket.

Kritiker tycker att denna föreslagna process är problematisk. Det har bl.a. pekats på av en sammanslutning av intresseföreningar (EDRI, september 2023). Haataja & Bryson (2023) ser förslaget som ett undergrävande av lagstiftarens förmåga att skapa trovärdiga kategorier av högrisksystem. De menar också att det öppnar ett stort rörelseutrymme för högst ifrågasatta tolkningar av risknivåer av leverantörer som vill undvika att följa lagstiftningen. Vidare menar de att en sådan ändring skulle ställa högre krav på myndigheternas tillsynskapacitet för att ha resurser att bedöma dessa anmälningar.

Överhuvudtaget kommer mycket av regleringens huvudsakliga syften att spela ut genom hur tillsynen arrangeras i praktiken, som därför behöver analyseras mycket mer än vad hittills har gjorts. Den föreslagna självvärderingen kan tyckas attraktivt mjuk och potentiellt kostnadsbesparande för ett företag som lanserar en AI-produkt, men det öppnar för en risk för slumpmässighet i tillämpningen, kanske rentav ett slags politisering av juridiken, där makten ligger hos kommissionen eller tillsynsmyndigheten

---

manner (e.g., with machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy" (OECD AI Principles).

att efter eget omdöme använda sig av sanktioner mot vissa aktörer. På medlemsstatsnivå kunde detta leda till tämligen obalanserad tillämpning baserad på tillgängliga resurser snarare än en normativ bedömning.

Haataja & Bryson (2023) argumenterar rentav för att det tycks som att Europaparlamentet försöker göra lagen till "ett vapen" som endast ska användas under extrema omständigheter, kanske alltid i efterhand, där en mycket hög regelbörda kan hävdas mot endast ett litet antal handplockade företag. Detta skulle i så fall gå stick i stäv med visionen om en europeisk digital ekonomi bestående av säkra, rättvisa, tillförlitliga AI-produkter. Haataja och Bryson argumenterar därför för att AI-förordningen bör hålla sig till en tillräckligt snäv men bestämd lista över högrisk kategorier för att undvika de skadliga effekterna av att införa ytterligare lager i högrisk-klassificeringen av ett AI-system.

## 4.2 2023: generativ AI i rådets och parlamentets förslag

Från kommissionens förslag på AI-förordning i april 2021, via rådets i december 2022 och parlamentets i juni 2023 har mycket hänt inom AI-utvecklingen. Framför allt under den period som Europaparlamentet har förhandlat fram sin version, med två parallella utskott, dels utskottet för den inre marknaden och konsumentskydd (IMCO), med Brando Benifei som rapportör, dels utskottet för medborgerliga fri- och rättigheter samt rättsliga och inrikes frågor (LIBE), med Dragoş Tudorache som rapportör. Chat GPT släpptes sent i november 2022 och blev väldigt populär redan i januari 2023, Microsoft bekräftade i mars att deras sökmotor Bing körs på en anpassad version av GPT-4, och Googles Bard, en chattbot byggd på grundmodellen LaMDA, lanserades också i mars. Man kan se i en kort skrift från parlamentets Tambiama André Madiega från den 30 mars 2023 att förståelsen för flerändamåls-AI utvecklats i parlamentet under våren 2023.<sup>25</sup>

### 4.2.1 Flerändamåls-AI och grundmodeller

Noterbart är att "generativ AI" inte har reglerats separat i kommissionens förslag till AI-förordning från april 2021. Rådets version, som börjar med versionen från november 2021 som publicerades under det slovenska ordförandeskapet, introducerade begreppen "General Purpose AI Systems" (GPAIS) och *generativa* AI-system (artikel 3, om definitioner). Detta utvecklades i en ändring som föreslogs under det franska rådets ordförandeskap i maj 2022, och antogs som rådets allmänna strategi i december 2022.

*General purpose AI systems* eller *AI-system för allmänna ändamål*, som den officiella översättningen säger (Europaparlamentet, 2023), är även med i parlamentets version, som ett "AI-system som kan användas i och anpassas till ett brett spektrum av tillämpningar för vilka det inte avsiktligt och specifikt utformats" (artikel 3[1d]). Utöver formuleringarna i rådets version gällande generativa AI-system ställde parlamentsversionen flera krav på leverantörer av "foundation models", dvs. *grundmodeller* (se definition i artikel 3), vilket var en ny terminologi.

---

<sup>25</sup> Se Europaparlamentet (30 mars 2023) "General-purpose artificial intelligence. At a Glance" av Tambiama André Madiega. [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_ATA\(2023\)745708](https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2023)745708) [senast besökt 5 oktober 2023].

Rådet tycks vilja reglera den typen av grundmodeller (genom GPAI-terminologin) på samma sätt som högrisksystem. Parlamentet utvecklar och ger mer nyanserade förslag på hur de bör regleras, genom de olika transparens- och dokumentationskrav de anger (se även Haataja & Bryson, 2023).

Leverantörer av grundmodeller ska enligt parlamentets förslag vara skyldiga att följa specifika krav relaterade till datahantering, riskhantering, modellutvärdering, energieffektivitet och kvalitetsstyrning (artikel 28b).

Mycket handlar om dokumentation och "data governance", dvs. dataförvaltningsåtgärder. De ska bland annat

- utarbeta omfattande teknisk dokumentation och begripliga bruksanvisningar för att göra det möjligt för leverantörer i senare led att uppfylla sina skyldigheter (artikel 28b[2e]), och detta ska ställas till förfogande för de nationella behöriga myndigheterna (artikel 28b[3])
- inrätta ett kvalitetsstyrningssystem (28b[2f]) och registrera denna grundmodell i en EU-databas (28b[2g])
- behandla och införliva endast datauppsättningar som omfattas av lämpliga dataförvaltningsåtgärder för grundmodeller (artikel 28b[2b])
- utforma och utveckla grundmodellen så att den under hela dess livscykel kan "uppnå lämpliga nivåer av prestanda, förutsägbarhet, tolkningsbarhet, korrigerbarhet, säkerhet och cybersäkerhet" (artikel 28b[2c])
- se till att "minska energianvändningen, resursanvändningen och avfallet" och öka energieffektiviteten och "utformas med kapacitet som gör det möjligt att mäta och registrera energi- och resursförbrukning" (artikel 28b[2d]).

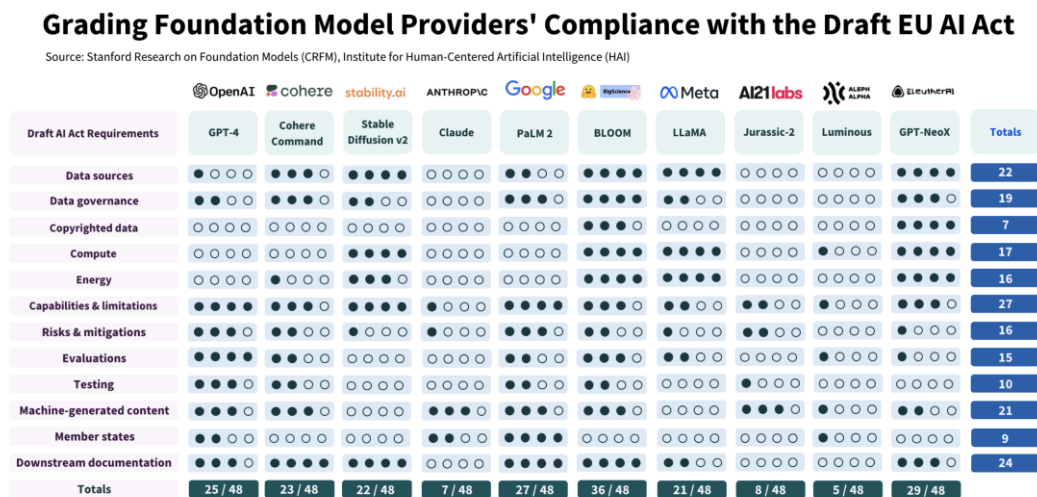
Slutligen, gällande upphovsrättsliga anspråk anges särskilda skyldighet för leverantörer av generativa AI-grundmodeller, dvs. sådana som är specifikt avsedda att med varierande grad av autonomi generera innehåll såsom komplex text, bilder, ljud eller video ("generativ AI"). För dessa ska dessa leverantörer "dokumentera och offentliggöra en tillräckligt detaljerad sammanfattning av användningen av träningsdata som skyddas enligt upphovsrättslagstiftningen" (artikel 28b[4d]).

Detta är en tämligen transparensfokuserad approach till regleringen av grundmodeller.

#### Grundmodeller och transparenskrav

Hur skulle då dagens leverantörer stå sig i relation till de krav som parlamentets version ställer upp? Intressant nog har forskare vid HAI-institutet vid Stanforduniversitetet kontrollerat detta i relation till lagförslaget (Bommasani m.fl. 2023). Forskarna gjorde först ett urval om 12 krav i lagförslaget som specifikt gäller grundmodeller, se Figur 5. De granskade därefter tio stora leverantörer och deras flaggskeppsmodeller i relation till varje krav. De aggregerade avslutningsvis leverantörers prestanda för varje krav i en enda poäng som en proxy för efterlevnad och jämförde resultat mellan leverantörer för att mäta trender inom branschen.

Figur 5 Analys av grundmodeller i relation till AI-förordningens transparenskrav, i parlamentets version. Från Bommasani m.fl., 2023



Det huvudsakliga resultatet visade följande:

- Öppna versioner av modeller ger generellt höga poäng på resursupplysningskrav relaterade till data och beräkning (jfr BLOOM från Hugging Face). Med sådana öppna grundmodeller kan det dock vara mer utmanande att övervaka eller kontrollera implementeringen än hos mer begränsade/stängda modeller som leder till bättre poäng på implementeringsrelaterade krav.
- Även om många leverantörer av grundmodeller räknar upp risker, är det relativt få som avslöjar de begränsningar de implementerar och effekten av dessa begränsningar. Lagen kräver också att leverantörer beskriver "icke-reducerade risker med en förklaring till varför de inte kan mildras" (min översättning), vilket ingen av leverantörerna bedöms göra.
- Få leverantörer avslöjar någon information om upphovsrättsstatus för träningsdata (inkl. Open AI:s GPT-4, Googles PaLM 2, och Metas LLaMA), kanske delvis för att försvåra krav från rättighetshavare. De rapporterar också inkonsekvent kring koldioxid-utsläpp från träning eller deras strategier för att mäta utsläpp.

### 4.3 Standardisering

Även om det inte finns plats för någon djupare analys av standardiseringens roll i relation till AI-förordningen i denna kartläggning, kan en viktig observation göras här. Det kan konstateras att ett sätt att hantera de behov av att säkerställa efterlevnad som följer med de krav som ställs i AI-förordningen är genom så kallade *harmoniserade standarder*. Dvs. kommissionen har bett standardiseringsorganisationerna European Committee for Standardization (CEN) och the European Committee for Electrotechnical Standardization (CENELEC) att utveckla standarder som kan erhålla en status där det antas att om man följer dessa standarder så tillgodoser man också AI-förordningens krav



på utvärdering och överensstämmelse (*presumption of conformity*).<sup>26</sup> Standardiseringsförordningen har dock väckt frågor om hur resurssvaga intresseorganisationer ska kunna bidra i och påverka frågor i svårtillgängliga standardiseringsprocesser (Veale och Zuiderveen Borgesius, 2021).<sup>27</sup>

En av de utpekade utmaningarna ligger även i att tekniska standarder förväntas definiera när och hur AI-system kommer att respektera (bland annat) grundläggande rättigheter, vilket beskrivits som en atypisk användning av standarder.<sup>28</sup> Civilsamhällets organisationer har varit kritiska till att utvidga tekniska standarder till att behandla grundläggande rättigheter och hävdar att beslutet om i vilken grad tekniska standarder överensstämmer med grundläggande rättigheter måste överlåtas till de demokratiska organen i EU eller medlemsstaterna (jfr BEUC, 2021; diskuteras av Micklitz, 2023).

## 4.4 Delegerade akter och genomförandeakter

En genomlysning av de olika förslagen på AI-förordning ger vid handen att så kallade delegerade akter (*delegated acts*) och genomförandeakter (*implementation acts*) tycks spela en särskilt viktig roll i hanteringen av AI-definitionen och för regleringen av flerändamåls-AI och grundmodeller.

### 4.4.1 Delegerade akter

Genom de olika förslagen från kommissionen, rådet och parlamentet tycks delegerade akter ha varit ett sätt att hantera utmaningar med rörliga begrepp och genuin osäkerhet. I kommissionens ursprungsförslags artikel 4 kan kommissionen ges rätt att genom delegerade akter vid behov "ändra förteckningen över de tekniker och metoder som förtecknas i bilaga I i syfte att uppdatera förteckningen så att den speglar marknadsutvecklingen och den tekniska utvecklingen" (i enlighet med villkoren i artikel 73). Dvs. den kan ges rätt att ändra i definitionen av AI. Enligt medierapportering har ledande medlemmar av parlamentet också uttryckt (i oktober 2022) att de önskar utöka det område för vilket kommissionen skulle kunna använda dessa befogenheter till att revidera i efterhand (Euractive, 2 november 2022).

Det är intressant ur denna kartläggnings grundtes, i dragkampen mellan fixering och flexibilitet. Det är också intressant ur ett slags rättssäkerhetsaspekt. Hur öppen mot extralegala framtida förändringar kan man lämna själva regleringsobjektets definition innan rättsakten blir alltför rättsosäker? Minns här de ovan nämnda rättsfilosofernas poäng om rättssystemets förutsägbarhet. Hur kan det påverka teknikutvecklare att det som vid lagens ikraftträdande inte ses som AI vid ett senare tillfälle kan komma att göra det?

---

<sup>26</sup> För en överblick över standardiseringsförfarande så som det är kopplat till AI-förordningen, se <https://artificialintelligenceact.eu/standard-setting/> [senast besökt 13 september 2023].

<sup>27</sup> Även konstitutionella frågor har väckts i relation till hur privata standardiseringsorgan med mandat under den nya lagstiftningsramen (NLF) inte bör befrias från rättslig granskningen av (de facto-) lagstiftare (Veale och Zuiderveen Borgesius, 2021, s. 105).

<sup>28</sup> Se CEPS (25 april 2023) "With the AI Act, we need to mind the standards gap", av Clément Perarnaud. <https://www.ceps.eu/with-the-ai-act-we-need-to-mind-the-standards-gap/> [senast besökt 6 oktober].

För att förtydliga: EU-lagstiftaren kan ge kommissionen befogenheter genom särskilda regler som är inskrivna i den befintliga rättsakten, här AI-förordningen, att anta så kallade *delegerade akter* (enligt artikel 290 i fördraget om Europeiska unionens funktionssätt). Delegerade akter beskrivs som "icke-lagstiftningsakter" som antas av kommissionen och som syftar till att ändra eller komplettera de icke väsentliga delarna av lagstiftningen. Det är dock inte självklart att se hur AI-definitionen i en reglering om AI kan ses som *icke-väsentlig*.

Kommissionen samråder med experter från EU:s medlemsstater innan den antar delegerade akter. En delegerad akt som antagits av kommissionen kan endast träda i kraft om inga invändningar görs av rådet eller parlamentet, inom den tidsfrist som anges i den grundläggande akten, dvs. i det här fallet AI-förordningen, som vanligtvis är två månader. Delegerade akter används vanligtvis när lagstiftningsakter – inklusive deras bilagor – måste anpassas (regelbundet) för att ta hänsyn till tekniska och vetenskapliga framsteg.

Rådets version från nov 2022 införde mer möjligheter till delegerade akter i relation till AI-definitionen, tycks det. Artikel 7(1) har ändrats för att ge en möjlighet att inte bara lägga till användningsfall med hög risk till listan genom delegerade akter, utan också att stryka dem. För att säkerställa att de grundläggande rättigheterna skyddas på ett adekvat sätt vid sådana strykningar har ytterligare bestämmelser lagts till i artikel 7(3) som specificerar de villkor som skulle behöva uppfyllas innan en delegerad akt kan antas. Det är också här som "general purpose AI" introducerades, fast med ett slags krav på vidare utredning kopplad till en *genomförandeakt* (artikel 4b[1]).

#### 4.4.2 Genomförandeakter

*General purpose AI* – här kallat flerändamåls-AI – tillkom i rådets version på AI-förordning. Här föreslås att vissa krav för högrisk-AI-system också skulle gälla för flerändamåls-AI (artikel 4b[1]). I stället för direkt tillämpning av dessa krav skulle dock en så kallad *genomförandeakt* specificera hur de ska tillämpas i förhållande till flerändamåls-AI, baserat på samråd med medlemsstaterna<sup>29</sup> och en detaljerad konsekvensbedömning och med hänsyn till specifika egenskaper hos dessa system och tillhörande värdekedja, teknisk genomförbarhet samt marknadsutveckling och teknisk utveckling.<sup>30</sup> Kommissionen kan använda sig av expertis för att förbereda ett utkast på en genomförandeakt, till exempel genom att konsultera expertgrupper, beställa en utredning eller hålla allmän utfrågning (Spasova, 2021).

Genomförandeakter är ett verktyg i EU-lagstiftningens verktygslåda som syftar till att skapa enhetliga förutsättningar för genomförandet av den aktuella rättsakten, genom att ge befogenhet till kommissionen (eller i undantagsfall Europeiska unionens råd) att utfärda genomförandeakter (enligt artikel 291 i fördraget om Europeiska unionens funktionssätt). Procedurmässigt antas genomförandeakter av kommissionen efter att ha hört kommittéer som består av tekniska experter från EU:s medlemsstater, i enlighet med

---

<sup>29</sup> Genom hänvisning till artikel 74(2) i AI-förordningen, som i sin tur hänvisar till artikel 5 i Europaparlamentets och rådets förordning (EU) nr 182/2011 av den 16 februari 2011 om fastställande av allmänna regler och principer för medlemsstaternas kontroll av kommissionens utövande av sina genomförandebefogenheter.

<sup>30</sup> Se EUR-Lex "Delegated acts" <https://eur-lex.europa.eu/EN/legal-content/glossary/delegated-acts.html>.

de regler och förfaranden som fastställs i förordning (EU) nr 182/2011. Europaparlamentet och rådet hålls informerade under utarbetandet av genomförandeakter och har rätt till granskning, men kan inte blockera antagandet av genomförandeakter.

Detta tycks också vara en flexibilitet i AI-förordningen – om dessa delar av förslaget antas – som skjuter väl valda delar av regleringen av flerändamåls-AI på framtiden. Det medför också ett slags förskjutning mot tekniska kommittéer, vilket väcker liknande frågor kring representation och deltagande som med standardiseringsförfarandet ovan.

## 4.5 Regulatoriska sandlådor: experimentell reglering

Kombinationen av AI och regulatoriska sandlådor är en diskussion och utveckling som hänger starkt ihop med AI-förordningens process (jfr EU-parlamentet, 2019; Madiega och Van De Pol, 2022). OECD publicerade en rapport om regulatoriska sandlådor för AI i juli 2023, helt i linje med AI-förordningens idéer. Rapporten fokuserar på sandlådor där myndigheter stimulerar företag att testa innovativa produkter eller tjänster som utmanar befintliga rättsliga ramar. Deltagande företag får undantag från specifika lagar eller efterlevnadsprocesser för att förnya sig. Den belyser positiva effekter och pekar ut utmaningar, risker och policyöverväganden för AI-sandlådor, med betoning på tvärvetenskapligt samarbete, uppbyggnad av AI-expertis, regleringskompatibilitet och handelspolitik.

En pilot kopplad till EU med ”den första regulatoriska sandlådan för AI” har också redan presenterats, i juni 2022, av Spaniens regering och kommissionen. Det gjordes vid ett evenemang som hölls i Bryssel i närvaro av spanska och europeiska myndigheter såväl som experter på området.<sup>31</sup>

### 4.5.1 Tidsbestämmdhet och flexibilitet

Artikel 53 i kommissionens förslag innehåller en bestämmelse om sandlådor för att tillhandahålla en kontrollerad miljö som underlättar utveckling, testning och validering av innovativa AI-system, men tycks inte tillåta någon avvikelse från reglering (jfr Colonna, 2023). Rådets version behåller bestämmelsen om sandlådor men ändrar den väsentligt. Den klargör att myndigheter har möjlighet att ”utöva sina tillsynsbefogenheter på ett flexibelt sätt” (författarens översättning). Parlamentets version av sandlådor tycks tydliggöra mer av den typen av tidsbestämda förslag på experimentell reglering som en del av litteraturen på taktproblem-området diskuterar. Dvs. att de ska

*tillhandahålla en kontrollerad miljö som främjar innovation och underlättar utveckling, testning och validering av innovativa AI-system under en begränsad tid innan de släpps ut på marknaden eller tas i bruk i enlighet med en särskild plan som de presumtiva leverantörerna och den utfärdande myndigheten kommer överens om (artikel 53[1][d]).*

---

<sup>31</sup> Vilket rapporteras om på kommissionens hemsida, se Europeiska kommissionen (27 juni 2022). ”First regulatory sandbox on Artificial Intelligence presented” <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented#:~:text=The%20regulatory%20sandbox%20is%20a%20way%20to%20connect,compliance%20with%20the%20requirements%20of%20the%20AI%20Regulation> [senast besökt 6 november 2023].

Märkligt nog lämnar rådets utkast, även om AI-lagstiftningen kommer att vara en förordning snarare än ett direktiv, utrymme för medlemsländerna att besluta om de ska inrätta AI-reglerande sandlådor eller inte (Buocz m.fl., 2023). Parlamentet, å andra sidan, kräver att medlemsstaterna upprättar minst en AI-regulatorisk sandlåda (art 53[1], skäl 71). Var detta landar vet vi inte än, men litteraturen om regulatoriska sandlådor lyfter både förtjänster och utmaningar. Dessa har delvis med vilken sorts regulatoriska lättnader som avses i testbädden, eller om man mest avser ett slags myndighetsinstitutionellt lärande där alla parter är med och lär sig. Detta talar för att flera myndigheter bör se över sina fält i ljuset av AI-utvecklingen. När det gäller sandlådor pekar en nyligen publicerad forskningsöversikt utförd åt Konkurrensverket på risker när dominerande företag ingår i en regulatorisk sandlåda, och ser det som motiverat att konkurrensmyndigheterna är vaksamma på frågan, vilket också sätter ljuset på hur urvalsprocessen till regulatoriska sandlådor går till (Norbäck & Persson, 2023). En annan utmaning när det gäller EU-reglering har med relationen till nationell reglering att göra, som kan erbjuda olika förutsättningar.

#### **4.5.2 Medlemsstatskomplexitet**

Buocz m.fl. (2023) använder sig av ett exempel med en tänkt sandlåda i Österrike kring ett medicinskt cancerupptäckningssystem. De konstaterar att varhelst regleringen av ett visst AI-system berör lagstiftning på både EU- och medlemsstatsnivå, kan implementeringen av AI-regulatoriska sandlådor bli en komplex uppgift. I Buocz m.fl.s österrikiska exempel anförs till exempel att harmoniseringskraven i Europarätten inte tillåter AI-förordningen att undanta de administrativa böter som föreskrivs i nationell lagstiftning med mindre att en ny nationell lagstiftning tas fram som möjliggör det. Här finns det rimligen behov av mer nationell analys kring rättsläget.

#### **4.5.3 Möjlig motsättning mellan regulatoriska sandlådor och generativ AI-transparens**

Även om det nog kan ses som en generell motsättning mellan regulatoriska sandlådor och regelverk som ställer höga krav på dokumentation och styrning, så kan man spekulativt undersöka huruvida det finns en motsättning i (parlamentsversionen av) AI-förordningen mellan transparenskraven för grundmodeller och möjligheterna till sandlådor. Kan grundmodeller både utvecklas i en regulatorisk sandlåda och samtidigt uppfylla dokumentationskraven som förordningen föreskriver? Dvs. ställs i samma krav på dokumentation om datastyrning, var data kommer ifrån, om de är upphovsrättsligt skyddade m.m. i den regulatoriska sandlådan? Om inte, vilket vore i linje med regulatoriska sandlådors innovationsfrämjande syfte, är det möjligt att tillgodose den typen av krav i efterhand, dvs. när träningsfasen är avslutad? Forskningen kring förklarbarhet i träningsfasen för AI-modeller har tydligt visat på utmaningar med att påvisa hur träningen går till, och vilka aspekter kring underliggande data som i detalj leder till ett visst utfall. Att i efterhand utröna huruvida stora mängder träningsdata var upphovsrättsligt skyddade riskerar att bli oöverstigligt. Hur förhåller sig kraven på rättslig grund enligt dataskyddslagstiftning till framtagandet av AI-modeller som bygger på personuppgifter? Om man å ena sidan kan träna modellerna utan rättslig grund, hur förhåller sig då det till den fas där AI-modellen ska släppas på marknaden, om dataskyddslagstiftningen inte uppfylls? Och, hur förhåller sig ad hoc-lättnader för enskilda aktörer till konkurrensrättsliga frågor?

Om inget undantag medges från tidskrävande och ekonomiskt kostsamma åtgärder, är det då verkligen en fungerande och innovationsfrämjande regulatorisk sandlåda?

## 5. Summering och diskussion: AI-reglering och innovation

Givet hur mycket som har hänt på AI-området under 2023 är det förståeligt om de europaparlamentariker som sattes att förhandla fram en position på AI-förordningen hade svårigheter med att fånga komplexiteten av det fenomen de sökte reglera. AI-förordningen har dels en förhistoria, dels samspelar den med en rad mjukare styrformer som standardisering och regulatoriska sandlådor. I det här kapitlet summerar jag först svaren på de fyra frågor som ställdes i rapportens inledning, och diskuterar ett urval av de mest relevanta resultaten.

### 5.1 I vilken mån och på vilket sätt förekommer mjuk styrning på AI-området överlag?<sup>32</sup>

Den etiskt inriktade principiella styrningen av AI har varit tongivande de senaste åren, både globalt (Larsson, 2020) såväl som inom EU (Larsson, 2021). Parallellt har även en rad standarder utvecklats, eller är under utveckling, för AI. För europeisk AI-styrning kommer troligen de s.k. harmoniserade standarder, som är under utveckling för att komplettera AI-förordningen, att bli av stor betydelse, vilket jag återkommer till under 5.3 nedan. AI-landskapet har förändrats tämligen drastiskt under 2023. Ett tecken på det är inte minst en ny våg av policydokument, där aktörer som OECD, G7, Bidenadministrationen, World Economic Forum m.fl. publicerat riktlinjer för generativ AI. Styrkan i denna mjuka styrning ligger i både snabbhet och möjligheten att finna de principiellt viktigaste förhållningssätten, där även privata aktörer kan tillföra information om den senaste teknikutvecklingen. Svagheter ligger naturligtvis i bristen på sanktioner och en process som kan hantera aktörer och normativa frågor på ett ordnat och förutsägbart vis.

De principiella och etiska riktlinjer för AI som har tagits fram har ofta fokuserat behovet av bättre transparens och förklarbarhet, frågor om ansvarsfördelning, behovet av tillförlitliga tillämpningar och rättvisefrågor. Det gäller en rad NGO:er, stora teknikföretag som Microsoft och Alphabet men även EU-kommissionen genom den expertgrupp de tillsatte (AI HLEG) som tog fram etiska riktlinjer för tillförlitlig AI under 2019. Dessa tycks ha bidragit till en del av AI-förordningens fokus på transparens och tillit, och parlamentet och rådet hänvisar explicit till expertgruppens riktlinjer. Liknande idéer kan noteras i svensk offentlig förvaltning (DO, 2022; Blom Lussi m.fl., kommande), och det svenska regeringsuppdrag som koordinerades av Digg (2020; jfr Larsson & Ledendal, 2022), och som bland annat tog fram en förtroendemodell för AI i offentlig förvaltning.

---

<sup>32</sup> Denna fråga besvaras framför allt i kapitel 2, som innehåller en global utblick och ett axplock av policyidéer från 2023, och kapitel 3, som redogör för den stora mängden principiella dokument och riktlinjer på AI-området som publicerats de senaste ca sju åren.

## 5.2 Hur har den snabba utvecklingen på AI-området påverkat framtagandet av den europeiska AI-förordningen?<sup>33</sup>

I denna rapport har den taktproblematik som framför allt lyfts fram haft att göra med 2023 års utveckling på området för stora språkmodeller och AI-grundmodeller, ofta refererat till som en del i eller ett uttryck för generativ AI. Diskussionen får ofta symboliseras av ChatGPT, som publicerades i november 2022 och som enligt vissa bedömare hade över 100 miljoner användare bara två månader senare.

Just den applikationen är dock bara toppen på ett isberg av en snabb utveckling av en rad AI-system och AI-grundmodeller för alltifrån text, programmeringskod, bild och ljud till kommande varianter av rörliga media, översättningar och lärande förmågor, med stora investeringar hos alla de större teknikbolagen.

Givet detta visar jag på hur AI-definitionen i AI-förordningen genomgår tämligen kraftiga förändringar från kommissionens förslag (april 2021) till rådets (december 2022) och parlamentets ändringsförslag (juni 2023). Detta gäller den allmänna definitionen, men lika intressant är det att koncept som *general purpose AI systems* och *generative AI systems* finns med i rådets version (december 2022) och i parlamentets version (juni 2023), där även *foundation models*, dvs. AI-grundmodeller, regleras särskilt.

Även om de skäl som anges i den här rapporten som stöd för en snabb AI-utveckling endast utgör ett smalt urval, finns det fog för argumentet att AI-landskapet har förändrats tämligen drastiskt bara under några månader från slutet av 2022 till och med 2023. Ett tecken på det är inte minst vad som ser ut som en ny våg av policydokument, där aktörer som OECD, G7, ACM, World Economic Forum, Bidenadministrationen, och tillsynsmyndigheter i exempelvis Storbritannien utvecklar och publicerar riktlinjer för generativ AI. Denna bild stöds även av respondentintervjuer som gjorts för denna rapport. Den snabba AI-utvecklingen får ses som en bidragande orsak till att AI-förordningen tagit mer tid än planerat att nå enighet kring förslag från rådet och parlamentet. Man kan också spekulera i om AI-förordningens flexibla element kommit att få en mer betydande roll i ljuset av osäkerheten i hur tekniken och dess marknader utvecklas. Man kan tänka sig att lagstiftaren skjuter fler detaljer på framtiden ju mer svårbedömd teknikutvecklingen är.

## 5.3 Vilka huvudelement av flexibilitet finns i förslaget/förslagen på AI-förordning?

Fokus här är några av AI-förordningens mer betydande flexibla element, som i) regulatoriska sandlådor, ii) standardiseringsprocesser och iii) kommissionens mandat att förändra eller förtydliga delar av AI-förordningen genom s.k. delegerade akter och genomförandeakter, särskilt gällande iv) AI-definitionen.

Kombinationen av AI och regulatoriska sandlådor är en diskussion och utveckling som hänger starkt ihop med AI-förordningens process med idén att deltagande företag får

---

<sup>33</sup> Denna fråga (2), samt efterföljande fråga 3 och 4, besvaras i kapitel 3.3., som ger en överblick över en formativ europeisk teknikregleringsperiod, och i kapitel 4, som redogör för förändringar kring AI-begreppet över de olika förslagen till AI-förordning 2021–2023.

undantag från specifika lagar eller efterlevnadsprocesser för att förnya sig. En spansk pilot startades redan 2022. Eftersom de olika förordningsförslagen på regulatoriska sandlådor i olika grad föreskriver inrättande av regulatoriska sandlådor i medlemsstaterna är det i skrivande stund inte klart hur tvingande AI-förordningen blir för att inrätta en sådan.

- Taktproblemslitteraturen lyfter fram regulatoriska sandlådor som en möjlig innovationsfrämjande väg, även om det finns kritik mot hur generellt lärandet egentligen är, utöver att främja den enskilde aktören, och att de ibland utvärderats otillfredsställande (jfr Ranchordas, 2021b).
- Det finns behov av att utveckla mer transparent och objektiv design av experimentella regler och regulatoriska sandlådor för AI.
- Det finns också en medlemsstatskomplexitet gällande AI som lyfts av analytiker (se Buocz m.fl., 2023), där relationen mellan nationell lag och EU-lagstiftning kan behöva utredas närmre.
- Det är i skrivande stund också oklart hur eventuella dokumentationskrav på leverantörer av AI-grundmodeller ska förhålla sig till regulatoriska sandlådor, men en potentiell motsättning har här identifierats.
- De mildare formerna av regulatoriska sandlådor som innebär en närmre samverkan med tillsynsmyndigheter finns det flertalet positiva erfarenheter av, och därmed fog för att ta med sig för svenskt vidkommande i relation till när AI-förordningen ska verkställas, vilket utvecklas mer nedan.

Även om det i den här kartläggningen inte funnits utrymme att fördjupa standardiseringsdelen av AI-förordningen kan konstateras att det kommer att vara ett viktigt instrument för efterlevnad, eftersom s.k. harmoniserade standarder förutsätts uppfylla krav som ställs i förordningen. Att bidra till standarderna kan därmed vara ett sätt att bidra till regleringens verkställighet och med minskade kostnader för efterlevnad. Med det sagt har standardiseringsproceduren kritiserats med tanke på resurssvaga intresseorganisationers möjligheter, samt ifrågasatts utifrån att tekniska standarder riskerar att inte vara lämpliga för att säkerställa grundläggande rättigheter (jfr BEUC, 2021; diskuteras av Micklitz, 2023).

En ur ett flexibilitetsperspektiv mycket intressant aspekt rör de mandat som inom AI-förordningen kan ges till kommissionen att i efterhand ändra och förtydliga lagstiftningen genom s.k. delegerade akter och genomförandeakter. Detta gäller inte minst AI-definitionen i sig, där både kommissionens och rådets version föreslog möjliga tillägg till listan med AI-tekniker, och där rådets version även gav möjlighet till att ta bort tekniker från listan, under vissa specificerade omständigheter. Rådet introducerade *general purpose AI*, dvs. flerändamåls-AI, och anger att kommissionen genom en genomförandeakt ska specificera hur högrisk-kraven ska tillämpas i förhållande till dessa mer generellt tillämpliga AI-system. Medlemsstaterna tycks inte föredra delegerade akter, möjligen för att dessa ger mindre mandat att påverka utförandet än genomförandeakter.

Överhuvudtaget tycks fältet för dessa specifika möjligheter att ändra och förtydliga lagstiftning relativt obeforskat utifrån innovationspåverkan och taktproblematik. Det finns en möjlig brist på förutsägbarhet här som behöver förstås bättre, speciellt för så centrala element som definitionen av AI i en AI-lagstiftning.



## 5.4 Hur föreslås s.k. generativ AI regleras i AI-förordningen och vilka huvudsakliga utmaningar är förenade med det i relation till pågående utveckling och innovation?

Som nämnt dyker *general purpose AI* upp först i rådets förslag på AI-förordning, nära kopplat till kraven för högrisk-AI. Mest relevant i relation till taktproblemet är att parlamentets förslag från juni 2023 – dvs. efter en vår fylld med debatt kring ChatGPT och andra generativa AI-applikationer – inkluderar förslag på hur s.k. grundmodeller bör regleras. Med detta avser parlamentet en "AI-systemmodell som tränats på en stor datamängd, utformats för att ge generella resultat och kan anpassas till ett brett spektrum av olika uppgifter". Detta skiljer de från *general purpose*-system som rådet först formulerade, och som parlamentet definierar som "AI-system för allmänna ändamål: ett AI-system som kan användas i och anpassas till ett brett spektrum av tillämpningar för vilka det inte avsiktligt och specifikt utformats."

Detta fält, som kommit att refereras till som *generativ AI*, har som nämnts föranlett framtagandet av ny våg av policydokument bara under 2023. Innovationshastigheten i relation till en redan pågående process att ta fram en AI-reglering för hela Europa, med uttalade ambitioner från kommissionens ordförande Ursula von der Leyen att bli "en modell för hela världen", har skapat en betydande utmaning för pågående förhandlingar om AI-förordning. Det har å ena sidan visat sig vara svårt att nå konsensus kring AI-begreppet som sådant, trots att det som forskningsfält inrättades redan på 1950-talet. Å andra sidan tycks det vara förenat med svårigheter att skapa en legitimt uppfattad reglering kring begrepp (*foundation models*) som inom forskningen blev etablerat så sent som omkring 2021 (jfr Bommasani m.fl., 2022), dvs. först efter kommissionen lagt sitt förslag på AI-förordning.

En sådan utmaning kan finnas i AI-förordningens fokus på risknivåer, och parlamentets försök att passa in generativ AI i de risknivåer som AI-förordningen delats upp i. Helberger & Diakopoulos (2023), exempelvis, argumenterar istället för att betrakta generativ AI och flerändamåls-AI mer allmänt som en generell riskkategori i sig. De tycker att det nuvarande riskbaserade tillvägagångssättet måste anpassas bättre till avsaknaden av ett fördefinierat syfte – som ju generativ AI i termer av grundmodeller ofta inte har – och risker relaterade till omfattningen av användning och utvinning av träningsdata.

Litteraturen kring taktproblem lyfter även utmaningen med undvikande entreprenörskap (se Birksjö m.fl., 2022), där vissa aktörer strävar efter att undvika de delar av regleringen som är mer betungande och kostsamma. Ett skäl, vilket diskuterats i kapitel 2 ovan, kunde vara att vissa aktörer försöker nå en konkurrensfördel i relation till de aktörer som följer regleringen. Ett annat skäl, kunde vara att lagstiftningen mer utbrett inte upplevs som legitim och rimlig, och därmed ger ett entreprenörskap som försöker undvika att kalla sina innovationer för de begrepp som triggar krav enligt förordningen.

Vidare, om regleringen inte upplevs som legitim, samtidigt som det finns element av självvärdering (som parlamentets förslag kring högrisk, uttryckt ovan), så kan man misstänka att utvecklare använder de frihetsgraderna för att undvika regleringskrav, oavsett om en mer objektiv bedömning skulle kategorisera ett befintligt AI-system som

högrisk. Här skulle regleringen i värsta fall riskera att hamna på ett gungfly som dels kan undergräva syftet att säkerställa en enhetlig rättslig ram med hög skyddsnivå för hälsa, säkerhet och grundläggande rättigheter, dels leda till en ad hoc-tillämpning med skönsmässiga inslag i tillsynen (se Haataja & Bryson, 2023, och 4.1.2 ovan).

Det finns även en konceptuell problembild för regleringsförslag i spännvidden mellan i) generativ AI, ii) AI-grundmodeller och iii) AI-system för generella ändamål/flerändamåls-AI. Hur leverantörer och utvecklare av AI-system av mer generell karaktär kommer att agera i relation till dessa uppdelningar är svårbedömt.

Gällande parlamentets förslag på krav på leverantörer av AI-grundmodeller kan man konstatera att de i mycket handlar om transparens- och dokumentationskrav (se även Haataja & Bryson, 2023). Det handlar om specifika krav relaterade till datahantering, riskhantering, modellutvärdering, energieffektivitet och kvalitetsstyrning, inklusive i vilken grad träningsdata skyddas av upphovsrätt. Intressant nog visade en utvärdering av ett antal av de mest kända AI-grundmodellerna att de är långt ifrån att uppfylla de krav som parlamentets förslag ställer upp (Bommasani m.fl., 2023). Få leverantörer avslöjar de begränsningar de implementerar och effekten av dessa begränsningar, och få avslöjar någon information om upphovsrättsstatus för träningsdata, och de rapporterar inkonsekvent kring koldioxid-utsläpp från träning eller deras strategier för att mäta utsläpp. Man kan tänka sig att de pågående lagstiftningsförhandlingarna laborerar med olika nivåer av grundmodeller, i syfte att inte ställa för omfattande dokumentationskrav på de mindre betydande. Huruvida så blir fallet återstår att se.

Ett försök till konceptuell uppdelning av juridisk relevans vore att se generativ AI som mer av en mjukvaruintegrerad tjänst, och AI-grundmodellerna som hänförliga till utvecklingssteget innan eller bakom detta. En sådan uppdelning tycks inte stödjas helt av AI-utvecklare och AI-forskare, men skulle kunna ha en poäng i relation till vilka risker som är förenade med de två delarna. Riskerna associerade med generativ AI behöver därmed inte helt överlappa fullständigt med de som förknippas med grundmodeller. Grundmodellernas problematik kan därmed kopplas mer till *data governance*, dvs. dataförvaltningsåtgärder: varifrån data kommer, hur modellen har tränats, hur beslut som grundmodellen når kan förklaras; är det grundmodellen eller applikationsnivån som är bakom problemen? De generativa systemen kan man kvalitetsgranska på ett annat sätt för att deras output är fullt synlig. Det kan visserligen uppstå frågor om vilka data systemet tränats på från ett upphovsrättsligt perspektiv, men black box-problematiken, dvs. problem med brist på transparens – handlar mer om output än om träningsdata.

## 6. Slutsatser

Syftet med den här rapporten har varit att analysera processen med framtagandet av den europeiska AI-förordningen, dvs. placera in AI-regleringsaspekter i ett brett perspektiv, och samtidigt därigenom belysa rättsliga utmaningar kopplade till taktproblemet. I förlängning av detta visar rapporten också på sätt att hantera dessa utmaningar. Trots den rättsliga fixering som AI-förordningen innebär vid ett ikraftträdande, ger den här analysen vid handen en del rörliga element som kan ge fingervisningar om hur vissa rättsliga utmaningar kan hanteras i relation till innovation på AI-området.

Vi befinner oss i en formativ period gällande AI-reglering där de beslut som tas kommer att ha stor påverkan för en väsentlig tid framåt, globalt, i Europa såväl som i Sverige. EU har som lagstiftare beslutat sig för att inte vara passiv när det gäller digitala frågor, inklusive gällande de stora aktörer som kännetecknar dominansen över det moderna internet, konkurrensfrågor, datahanteringsfrågor – och AI. Betydelsen av denna rättsutveckling och hur överlapp och glapp mellan rättsakter ser ut kommer att sysselsätta många analytiker framöver. AI-förordningen kan potentiellt komma att bli en av de viktigaste regleringarna i det globala landskapet.

Att förstå dessa processer – tillblivelsen och framtagandet av AI-förordningen, dess inneboende dynamik med flexibla element men även relationen till annan lagstiftning samt kommande detaljer i verkställighet och nationell tillsyn – är en förutsättning för att svenska intressenter ska kunna påverka, hantera eller positionera sig väl. Att förstå de nya inslagen i AI-förordningen möjliggör för svenska intressenter att nyttja dessa.

Ett särskilt fokus har i den här rapporten lagts vid generativ AI eftersom det är ett särdeles intressant problemområde i relation till taktproblem. Regleringen har att fixera skyddsbehov och samtidigt erbjuda flexibilitet i förhållande till förändrade förutsättningar i en nära framtid, med delvis genuin osäkerhet i vilka AI-tekniker och tillämpningsområden som växer fram.

### 6.1 Mellan fixering och flexibilitet

Man kan först konstatera att mjuk styrning i termer av riktlinjer, ofta formulerad i termer av etik, har spelat en betydande roll för AI-utveckling och AI-användning de senaste sju åren, med en ny våg för generativ AI under 2023. Styrkan ligger i idéutvecklingen med formulerandet av principiella förhållningssätt, specifikt ofta fokus på rättvisefrågor, transparensbehov och ansvarsfördelning. Svagheten ligger i brist på sanktioner för den aktör som avviker.

Det finns tre områden där studien har landat i slutsatser som berör svenskt politiskt beslutsfattande, och de områdena är AI-förordningens flexibilitet, AI-förordningens konsekvenser för svensk tillsyn och dess del i en rättslig komplexitet.

#### 6.1.1 AI-förordningens flexibilitet

Det finns en konceptuell osäkerhet i AI som styrningsbegrepp som är helt central för AI-förordningen. Det gäller "AI" i sig, såväl som för de nyare begreppen "grundmodell", "AI-system för allmänna ändamål" och "flerändamåls-AI" som tillförts regleringsförslagen i takt med att AI-innovationerna fått ett starkt och snabbt genomslag. Att AI-förordningen knyter lika krav till två relativt nya begrepp är förenat med vissa

utmaningar i relation till hur praktiker, innovatörer och forskare förstår begreppen. Jag har ovan argumenterat för att snabbheten i AI-utvecklingen hänger samman med flexibla inslag i AI-förordningen, som sätt att hantera osäkerhet. De flexibla inslagen har följder som bör uppmärksammas:

- Vissa flexibiliteter i AI-regleringen förskjuter makten över bestämmandet från lagstiftningsprocess till kommissionen, andra till standardiseringsorgan. Den europeiska lagstiftaren tycks både vilja adressera generativ AI och samtidigt skjuta det på framtiden. Några av de verktyg som står till buds är s.k. delegerade akter och genomförandekter, men även så kallade harmoniserade standarder.
- Vissa taktproblemlösningar tycks ha en betydande baksida även för marknader: oförutsägbarhet. Tillväxtanalys har i en tidigare studie (2023) konstaterat att företagen upplever en oro för att nya regleringar kan komma att skapa en osäkerhet vad avser spelreglerna på marknaden och efterfrågar tydlig reglering. Det finns en risk för att centrala delar av AI-förordningen blir principiellt inriktade men osäkert reglerade just för att precisionen inte låter sig nås än. De förtydliganden som senare ska göras av kommissionen, tillsynsmyndigheter eller rättspraxis betyder rimligen att oförutsägbarheten råder fram till dess.
- Det tycks finnas vissa frågetecken kring hur användbara regulatoriska sandlådor är för AI-innovationer i Sverige: Det handlar mer generellt om huruvida det är möjligt med ett bredare lärande än för det enskilda testfallet, och samtidigt om förenligheten med svensk rätt. Detta beror också på vad man avser med regulatorisk sandlåda, se mer om tillsynssidan nedan. Mer specifikt gällande grundmodeller finns det frågetecken kring om det är möjligt att förena ett befriande av dokumentations- och transparenskrav i en sandlådefas med ett tillämpande av samma krav vid ett senare inträde på marknaden.

### 6.1.2 AI-förordningens konsekvenser för svensk tillsyn

Utöver det mer generella lärandet om hanteringen av taktproblem kan man mer specifikt fråga sig vad det svenska utrymmet är för att påverka den europeiska AI-regleringen. I takt med att förordningen går över i skedet att blir tillämplig kommer tillsynsmyndigheternas förhållningssätt att bli en avgörande fråga. Det finns några taktproblemsrelevanta observationer att göra här. Mer flexibel praktik och vägledning hos tillsynsmyndigheter framstår som ett möjligt och önskvärt sätt att hantera taktproblem, vilket särskilt gäller följande:

- *Praktisk rättslig vägledning* från myndigheter i relation till AI-utveckling och reglering. Det tycks till exempel oklart om det är möjligt att i Sverige göra undantag för relevant lagstiftning för regulatoriska AI-sandlådor, och i så fall i vilken grad. Mer direkt vägledning som praktik hos tillsynsmyndigheter är högst möjligt, vilket pilotprojekt hos IMY visat (IMY, 2023), och behövt.
- *Även regleraren behöver förstå tekniken*, i termer av att både förvaltningsmyndigheter och tillsynsmyndigheter troligen behöver kompetensutveckla vad AI-utvecklingen medför på deras verksamhets- och tillsynsområde. Ett vanligt utpekad taktproblem har att göra med informationsasymmetrier mellan lagstiftare och teknikutvecklare (jfr Taihagh m.fl., 2021). I ljuset av att AI-utvecklingen så mångbottnat kan påverka en

- rad olika intressen och områden, bör därmed även svenska tillsynsmyndigheter kontinuerligt utveckla kunskap om hur utvecklingen påverkar just deras område.<sup>34</sup>
- *Myndighetssamverkan* tycks central. Även om regeringen såklart noggsamt behöver utreda hur tillsyn enligt förordningen bäst utses eller fördelas, är det inte otänkbart att AI-förordningens tämligen breda ämnesområden – med alltifrån frågor om diskriminering och mänskliga rättigheter till väldigt tekniskt avancerade frågor, som kan relatera till finansiella marknader, konkurrens, konsumenter, medicinska applikationer, dataskyddsfrågor – kommer kräva en utökad samverkan mellan flera tillsynsmyndigheter.

Dessa slutsatser kan också ses i ljuset av att GDPR utvärderades relativt nyligen av kommissionen, vilket diskuterats i parlamentet som tagit fram en resolution (Europaparlamentet, 25 mars 2021). Även om det finns betydande skillnader mellan en dataskyddsförordning med individuella rättigheter och en produktsäkerhetsinriktad AI-förordning, så kan man dra vissa lärdomar från problem med verkställighet och tillsyn som GDPR visade på. Till exempel uttrycktes i utvärderingen problemet med ojämn och ibland obefintlig verkställighet hos medlemsstaternas tillsynsmyndigheter; att olika tolkningar görs på medlemsstatsnivå, vilket skapar rättsosäkerhet; och att endast ett fåtal fall följs upp av den stora mängd klagomål som sker, vilket hänger ihop med en brist på resurser och teknisk kompetens.

Detta är rimligen lärdomar att ta fasta på för svenskt vidkommande, som kommer att avgöra hur väl myndigheter kan ge vägledning som kan minska den tid och de resurser som företag och andra myndigheter behöver lägga på att utreda huruvida deras AI-utveckling och AI-användning överensstämmer med förordningens krav.<sup>35</sup>

### 6.1.3 Om rättslig komplexitet och behovet av kunskap

Man kan ta med sig att den europeiska lagstiftaren är väldigt aktiv på områden som har med digitalisering, data och AI att göra. Det finns därmed rättsliga överlapp i det europeiska regleringslandskapet som behöver förstås bättre framöver:

- Det finns rättsliga överlapp i det europeiska regleringslandskapet som behöver förstås bättre, och för svenskt vidkommande troligen koordineras mer. I linje med litteratur om rättsliga överlapp (*legal overlaps*, Crootof & Ard, 2021), finns det anledning att tro att det kommer uppstå en rad frågor i relationen mellan flertalet

<sup>34</sup> Brittiska *Competition and Markets Authority*, CMA, publicerade i september 2023 en kartläggning av vilka konsekvenser de s.k. AI-grundmodellerna kan ha för deras verksamhetsområden gällande konkurrens och konsumentintressen. Liknande kunde man föreställa sig även vore välkommet för motsvarigheter hos IMY, Konsumentverket, Konkurrensverket, m.fl. Här kan Diskrimineringsombudsmannen lyftas fram som ett exempel, med flertalet publikationer på AI-området, sett utifrån en högst relevant diskrimineringshorisont (exv. DO, 2022). DO rapporterade också nyligen ett regeringsuppdrag om risker för diskriminering i användning av AI och annat automatiserat beslutsfattande i arbetslivet (DO, 2023), vilket dels pekar på att myndigheterna kan ges i uppdrag att fördjupa både sin egen och den allmänna kunskapen på relevanta områden.

<sup>35</sup> Vidare finns det ett bredare rättsligt landskap att analysera och förbereda sig på. För rättsakten för digitala tjänster (DSA) föreslås till exempel tre myndigheter att bli tillsynsmyndigheter (SOU 2023:2): Post- och telestyrelsen (också som samordnare för digitala tjänster), Konsumentverket och Myndigheten för press, radio och tv. Kanske vore en uppdelad eller samverkande tillsyn också vägen för AI-förordningen, givet dess mångfacetterade ämnes- och ansvarsområden.

betydande rättsakter av betydelse för AI-området; i relationen mellan AI-förordningen, GDPR, rättsakten för digitala tjänster (DSA), rättsakten för digitala marknader (DMA) såväl som för upphovsrättsdirektivet. Forskningsundantaget i AI-förordningen är en sådan fråga som adderar komplexitet till ett redan komplext regleringslandskap (Colonna, 2023).

- Mer spekulativt kan man konstatera att kompetens, forskning och utveckling kostar. En djupare och starkare offentlig satsning på kompetensutveckling och forskning på AI, både kring tekniken och dess samhällskonsekvenser, vore eftersträvansvärd.

Tillväxtanalys (2022) har tidigare visat på hur det framför allt är privata satsningar på AI-forskning som utgör djup och tyngd. Någon motsvarande större offentlig satsning finns inte i Sverige. Som jämförelse kan Norge nämnas, som nyligen beslutade att satsa 1 miljard norska kronor över fem år på AI-forskning, och har skapat ett nytt departement och utsett en ny minister för digitalisering och styrning, med ansvar för AI. Norge har även en AI-strategi från 2020, tydligt inspirerad av HLEG:s etiska riktlinjer (af Malmborg, 2020), på ett sätt Sverige inte riktigt har.

## 6.2 Avslutning

Avslutningsvis vill jag återkomma specifikt till generativ AI och de mer allmänt användbara AI-grundmodellerna. När jag i rapportens inledning hänvisar till uttalanden från Mira Murati, teknikchefen på OpenAI som utvecklar bildgenererande AI-grundmodeller, så var just argumentet att det inte är för tidigt att reglera något som trots allt rör sig så snabbt. Detta för att det finns nytta att införliva, som är beroende av spelregler och sätt att på samhällsnivå hantera en rad utpekade AI-relaterade risker. En reglering fokuserad på artificiell intelligens, dess utveckling och, framför allt, implementering, är ett sätt för den europeiska lagstiftaren att på ett oerhört ambitiöst sätt med världens blickar på sig satsa på att det är den mest framkomliga vägen.

Till hjälp finns ett par taktproblemsinriktade insikter som lagstiftaren införlivar i termer av flexibiliteter inbakade i en i övrigt strikt risknivåuppdelning. Med flexibilitet följer dock oförutsägbarhet, vilket kan vara ett tveeggat svärd för marknadens aktörer. Å ena sidan kan regulatoriska sandlådor och nära myndighetssamverkan fungera väl, å andra sidan kan regleringens centrala begrepp ändras på kort tid genom de mandat som förordningen anger, och därmed regleringens träffyta. Någonstans finns en gräns för när flexibilitet i reglering medför för mycket oförutsägbarhet som ytterst kan betyda rättsosäkerhet. Risken är att kostnader förenade med rättsliga krav på utvärderingar och dokumentation – och oförutsägbarheten om kraven träffar verksamheten ovanpå det – är innovationshämmande. Å andra sidan är riskerna för oönskade effekter som diskriminerande AI-system, skalbara desinformationskampanjer, felande eller överutnyttjade övervakningsteknologier omvittnade, och kan i sin tur hämma tilliten till AI-tillämpningar som sådana. Vilket kan medföra djupt oroande samhällskostnader.

Det kanske inte heller är rimligt att förvänta sig ett koncist ställningstagande hos den europeiska regleraren, givet hur svårbedömt det är att sja om vilken sorts teknisk utveckling vi står inför, vilken typ av marknadsdynamik den medför och därmed hur riskerna fördelas och bäst motverkas. Att förstå den här dynamiken och dess inneboende motsättningar utgör dock grunden för alla aktörers förmåga att positionera sig väl, såväl för svenska reglerare som för marknadsaktörer och tillsynsmyndigheter. Vi tycks därtill befinna oss i en formativ period gällande utveckling och reglering av AI. De val som görs

nu är betydande och kommer att få konsekvenser för en lång tid framöver. Desto viktigare är det då att göra dessa val så väl som möjligt.

## 7. Referenser

### 7.1 Litteratur

- Aubert, V. (1989). *Continuity and development in law and society*. Oxford University Press, USA.
- Ajunwa, I. (2023). *The Quantified Worker: Law and Technology in the Modern Workplace*. Cambridge University Press.
- Bar-Siman-Tov, I., & Harari-Heit, G. (2020). The legisprudential and political functions of temporary legislation. In: Ranchordás, S., & Roznai, Y. (Eds.). *Time, Law, and Change: an interdisciplinary study*. Bloomsbury Publishing.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? 🦜. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623).
- Bennett Moses, L. (2013). How to think about law, regulation and technology – problems with “technology” as a regulatory target. *Law, Innovation & Technology*, vol. 5 (1), 1–20.
- Birksjö, D., Alfredsson, E. och Svensson, P. (2022) *Utmaningar vid reglering av teknisk innovation – möjliga policyåtgärder*. Rapport 2022:04, Myndigheten för tillväxtpolitiska utvärderingar och analyser.
- Blom Lussi, E., Larsson, S., Högberg, C., Kaun, A. (kommande) Governing the automated welfare state: Translations between AI ethics and anti-discrimination regulation, *Nordic Welfare Research*.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2022). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*. Authored by the Center for Research on Foundation Models (CRFM) at the Stanford Institute for Human-Centered Artificial Intelligence (HAI).
- Bradford, A. (2012). The brussels effect. *Nw. UL Rev.*, 107, 1.
- Buckley, R. P., Arner, D. W., Veidt, R. & Zetsche, D. A. (2019). Building FinTech ecosystems: Regulatory sandboxes, innovation hubs and beyond. *Washington University Journal of Law and Policy*, 61, 55–56. doi:10.2139/ssrn.3455872.
- Buocz, T., Pfothenauer, S., & Eisenberger, I. (2023). Regulatory sandboxes in the AI Act: reconciling innovation and safety? *Law, Innovation and Technology*, 1–33.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91), PMLR.
- Colonna, L. (2023) “The AI Act’s Research Exemption: A Mechanism for Regulatory Arbitrage?” In *YSEC Yearbook of Socio-Economic Constitutions*. Springer, Cham.
- Crootof, R. och Ard (2021) Structuring Techlaw. *Harvard Journal of Law & Technology*, 34(2): 347–417.



- Diskrimineringsombudsmannen. (2022). *Transparens, träning och data. Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering*. Rapport 2022:1, Stockholm: Diskrimineringsombudsmannen.
- Edwards, L. (april 2022) *The EU AI Act: a summary of its significance and scope*. Ada Lovelace Institute.
- Felländer, A., Rebane, J., Larsson, S., Wiggberg, M., & Heintz, F. (2022). Achieving a Data-driven Risk Assessment Methodology for Ethical AI. *Digital Society* 1(2): 1–27.
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication*, (2020-1).
- Hacker, P. (2023). The European AI liability directives—Critique of a half-hearted approach and lessons for the future. *Computer Law & Security Review*, 51, 105871.
- Hagemann, R., Huddleston Skees, J., & Thierer, A. (2018). Soft law for hard problems: The governance of emerging technologies in an uncertain future. *Colo. Tech. LJ*, 17, 37.
- Helberger, N., & Diakopoulos, N. (2023). ChatGPT and the AI Act. *Internet Policy Review*, 12(1).
- Haataja, M., & Bryson, J. J. (2023). The European Parliament's AI Regulation: Should We Call it Progress? *Amicus Curiae*, 2(4): 707–718.
- Hauglid, M. K., & Mahler, T. (2023). Doctor Chatbot: The EU's Regulatory Prescription for Generative Medical AI. *Oslo Law Review*, (1), 1–23.
- Högberg, C. & Larsson, S. (2022) "AI and Patients' Rights: Transparency and information flows as situated principles in public health care", in de Vries & Dahlberg (eds.) *Law, AI & Digitalization. De Lege – Yearbook Uppsala Faculty of Law 2021*. Uppsala, Sweden: Iustus Förlag.
- Hydén, H. (2022). "Regulation of AI: Problems and Options." In: Liane Colonna, Stanley Greenstein (eds.) *Nordic Yearbook of Law and Informatics 2020–2021: Law in the Era of Artificial Intelligence*. The Swedish Law and Informatics Research Institute.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, 1(9), 389–399.
- Knight, B. R. & Mitchell, T. E. (2020). The sandbox paradox: Balancing the need to facilitate innovation with the risk of regulatory privilege. *South Carolina Law Review*, 72(2), 445–476.
- Larsson, S. (2018). Algorithmic governance and the need for consumer empowerment in data-driven markets. *Internet Policy Review*, 7(2).
- Larsson, S. (2019) The Socio-Legal Relevance of Artificial Intelligence, "Law in an Algorithmic World", Special Issue of *Droit et Société*.
- Larsson, S. (2020) On the Governance of Artificial Intelligence through Ethics Guidelines, *Asian Journal of Law and Society*, 7(1): 1–23.

- Larsson, S. (2021) "AI in the EU: Ethical Guidelines as a Governance Tool", in Bakardjieva Engelbrekt, Leijon, Michalski & Oxelheim (eds.) *The European Union and the Technology Shift*. Cham, Switzerland: Palgrave Macmillan.
- Larsson, S., Anneroth, M., Felländer, A., Felländer-Tsai, L., Heintz, F., Cedering Ångström, R. (2019) *HÅLLBAR AI. Inventering av kunskapsläget för etiska, sociala och rättsliga utmaningar med artificiell intelligens*, Stockholm: AI Sustainability Center.
- Larsson, S. & Heintz, F. (2020) Transparency in artificial intelligence, *Internet Policy Review* 9(2): 1–16.
- Larsson, S., Liinason, M., Tanqueray, L., & Castellano, G. (2023). Towards a Socio-Legal Robotics: A Theoretical Framework on Norms and Adaptive Technologies. *International Journal of Social Robotics*.
- Larsson, S & Ledendahl, J (2017). *Personuppgifter som betalningsmedel*. Konsumentverket rapport 2017:4.
- Larsson, S. & Ledendahl, J. (2022) "AI i offentlig sektor: Från etiska riktlinjer till lagstiftning", in de Vries & Dahlberg (eds.) *Law, AI & Digitalization. De Lege – Yearbook Uppsala Faculty of Law 2021*. Uppsala: Iustus Förlag.
- Larsson, S., Haresamudram, K., Högberg, C., Lao, Y., Nyström, A., Söderlund, K., & Heintz, F. (2023) Four Facets of AI Transparency, in Lindgren, S., (ed.) *Handbook of Critical Studies in Artificial Intelligence*, Edward Elgar Publishing.
- Legg, S., & Hutter, M. (2007). "A Collection of Definitions of Intelligence". In B. Goertzel & P. Wang (Eds.), *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms* (pp. 17–24), Proceedings of the AGI Workshop 2006 (Vol. 157), IOS Press.
- Luhmann, N. (2008). *Rechtssoziologie*. Springer-Verlag.
- af Malmberg, F. (2020) "AI policy in Norway – looking to the future and harmonised with the EU", i Larsson, S., Ingram Bogusz, C., & Andersson Schwarz, J. (red.) *Human-Centred AI in the EU. Trustworthiness as a strategic priority in the European Member States*. Brussels: European Liberal Forum.
- Marsden, C. (2022). Platform Law in Europe-Combating Digital Harms Through Co-Regulation. *경제규제와 법*, 15(1), 34–66.
- Micklitz, H.W. (2023) *The Role of Standards in Future EU Digital Policy Legislation. A Consumer Perspective*. ANEC & BEUC.
- Noble, S. U. (2018). *Algorithms of Oppression*. New York university press.
- Palmås, K., Andersson Schwarz, J., & Larsson, S. (2014). The liability of politicalness: Legitimacy and legality in piracy-proximate entrepreneurship. *International Journal of Entrepreneurship and Small Business*, 22(4), 408–425.
- Peczenik, A. (1995). Vad är rätt? Om demokrati, rättssäkerhet, etik och juridisk argumentation. Norstedts Juridik AB.
- Ranchordás, S. (2014). *Constitutional sunsets and experimental legislation: A comparative perspective*. Cheltenham: Edward Elgar.

- Ranchordas, S. (2021a). Experimental Regulations and Regulatory Sandboxes: Law without Order? *Law and Method* 1–35.
- Ranchordas, S. (2021b). Experimental Regulations for AI: Sandboxes for Morals and Mores. *MORALS & MACHINES*, (1), 86–100.
- Ranchordas, S. (2015). Innovation experimentalism in the age of the sharing economy. *Lewis & Clark L. Rev.*, 19, 871.
- Renner, Karl (2010/1949) *The Institutions of Private Law and Their Social Functions*, New Brunswick & London: Transaction Publishers.
- Richardson, R. (2021). Defining and demystifying automated decision systems. *Md. L. Rev.*, 81, 785.
- Spasova, R. (2021). Powers of the European Commission – delegated and implementing acts in practical terms. In *ERA Forum* (Vol. 22, No. 3, pp. 507-521). Berlin/Heidelberg: Springer Berlin Heidelberg.
- Stone, P., et al. (2016). *Artificial Intelligence and Life in 2030, Report of the* (pp. 2015–2016). Stanford University: Study Panel. Stanford.
- Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*, 15(4), 1009–1019.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Veale, M. & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, 22(4), 97–112.
- Wang, C., Chen, S., Wu, Y., Zhang, Z., Zhou, L., Liu, S., ... & Wei, F. (2023). Neural codec language models are zero-shot text to speech synthesizers. *arXiv preprint arXiv:2301.02111*.

## 7.2 Andra källor

- ABC News (16 mars 2023) "OpenAI CEO Sam Altman says AI will reshape society, acknowledges risks: 'A little bit scared of this'", av Victor Ordonez, Taylor Dunn och Eric Noll. <https://abcnews.go.com/Technology/openai-ceo-sam-altman-ai-reshape-society-acknowledges/story?id=97897122> [senast besökt 5 oktober 2023].
- ACM Technology Policy Council (juni 2023) Principles for the development, deployment, and use of generative AI technologies.
- AI HLEG (2019a) Etiska riktlinjer för tillförlitlig AI.
- AI Safety Summit (1 november 2023) The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> [senast besökt 6 november 2023].

AI Sweden (7 december 2021) "What's the latest with the GPT-SWE?"

<https://www.ai.se/en/news/whats-latest-gpt-swe> [senast besökt 6 november 2023].

Algorithm Watch (2023) The AI Act and General Purpose AI: Charting a path forward.

Algorithm Watch & University of Amsterdam – AI, Media & Democracy Lab.

[https://algorithmwatch.org/en/wp-content/uploads/2023/09/PolicyBrief\\_GPAI\\_AW-3.pdf](https://algorithmwatch.org/en/wp-content/uploads/2023/09/PolicyBrief_GPAI_AW-3.pdf)

[senast besökt 12 november 2023]

Arbetsmarknadsdepartementet (2022), Regeringsbeslut 2022-06-07, A2022/00887.

BEUC (2021) Regulating AI to Protect the Consumer, Position Paper on the AI Act, 7.10.2021 in particular under 9 at 25, co-authored with ANEC

[https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-](https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf)

[088\\_regulating\\_ai\\_to\\_protect\\_the\\_consumer.pdf](https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf) [senast besökt 6 oktober 2023].

Bommasani, R, Klyman, K, Zhang, D & Liang, P. (juni 2023) "Do Foundation Model Providers Comply with the Draft EU AI Act?" <https://crfm.stanford.edu/2023/06/15/eu-ai-act.html> [senast besökt 18 september 2023].

Bryson, J. (maj 2023) Keynote "No one should trust AI: How and why we govern production of intelligent artefacts", Rise of AI Conference, Berlin.

<https://youtu.be/YgZD74jIMjY?si=Ekn12ycazNK1Ybvj> [senast besökt 20 september 2023].

CEPS (25 april 2023) "With the AI Act, we need to mind the standards gap", av Clément

Perarnaud. <https://www.ceps.eu/with-the-ai-act-we-need-to-mind-the-standards-gap/>

[senast besökt 6 oktober].

Competition & Markets Authority (september 2023) AI Foundation Models: Initial Report.

COM(2022) 496 final. Förslag till EUROPAPARLAMENTETS OCH RÅDETS DIREKTIV Om anpassning av reglerna om utomobligatoriskt skadeståndsansvar vad gäller artificiell intelligens (direktivet om skadeståndsansvar gällande AI).

Digg (2022) Självsmål i målstyrningen? Digg reflekterar kring digitaliseringens komplexitet. 2022-3147.

Diskrimineringsombudsmannen. (2022). *Transparens, träning och data. Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering*. Rapport 2022:1, Stockholm: Diskrimineringsombudsmannen.

Diskrimineringsombudsmannen, DO (2023) *AI och risker för diskriminering i arbetslivet*.

Slutredovisning av regeringsuppdraget om risker för diskriminering vid användandet av artificiell intelligens och annat automatiserat beslutsfattande inom arbetslivet.

EDRi (September 2023) "EU legislators must close dangerous loophole in AI Act"

[https://edri.org/wp-content/uploads/2023/09/AI-Act\\_Article-6-NGO-statement-draft-FINAL.pdf](https://edri.org/wp-content/uploads/2023/09/AI-Act_Article-6-NGO-statement-draft-FINAL.pdf) [senast besökt 13 september, 2023].

EUR-Lex "Delegated acts" <https://eur-lex.europa.eu/EN/legal-content/glossary/delegated-acts.html>

Europeiska kommissionen, 25.4.2018, COM(2018) 237 final.

Europeiska kommissionen. Förslag till regelverk om artificiell intelligens <https://digital-strategy.ec.europa.eu/sv/policies/regulatory-framework-ai> [senast besökt 6 november, 2023].

Europeiska kommissionen (2018) Meddelande från kommissionen till Europaparlamentet, Europeiska rådet, rådet, Ekonomiska och sociala kommittén och Regionkommittén om artificiell intelligens för Europa, 25.4.2018, COM(2018) 237 final.

Europeiska kommissionen (31 maj 2023) EU-U.S. Terminology and Taxonomy for Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence> [senast besökt 5 oktober 2023].

Europeiska kommissionen (2020) Vitbok – Om artificiell intelligens – en EU-strategi för spetskompetens och förtroende, 19.2.2020, COM(2020) 65 final.

Europeiska kommissionen (13 september 2023) "Tal om tillståndet i unionen 2023 av kommissionens ordförande Ursula von der Leyen". [https://ec.europa.eu/commission/presscorner/detail/sv/speech\\_23\\_4426](https://ec.europa.eu/commission/presscorner/detail/sv/speech_23_4426) [senast besökt 5 oktober 2023].

Europeiska kommissionen (20 maj 2015) The Principles for better self- and co-regulation endorsed in the Better Regulation Package. <https://digital-strategy.ec.europa.eu/en/news/principles-better-self-and-co-regulation-endorsed-better-regulation-package> [senast besökt 5 oktober 2023].

Europeiska kommissionen (27 juni 2022) "First regulatory sandbox on Artificial Intelligence presented" <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented#:~:text=The%20regulatory%20sandbox%20is%20a%20way%20to%20connect,compliance%20with%20the%20requirements%20of%20the%20AI%20Regulation.>

Europeiska kommissionen (2021b) COMMISSION STAFF WORKING DOCUMENT EXECUTIVE SUMMARY OF THE IMPACT ASSESSMENT REPORT Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS {COM(2021) 206 final} - {SEC(2021) 167 final} - {SWD(2021) 84 final}.

Europeiska kommissionen (21 april 2023) Impact Assessment of the Regulation on Artificial intelligence <https://digital-strategy.ec.europa.eu/en/library/impact-assessment-regulation-artificial-intelligence> [senast besökt 6 november 2023]

Europaparlamentet (2019). Resolution of 12 February 2019 on a Comprehensive Industrial Policy on Artificial Intelligence, 2018/2088 (INI). Främja den offentliga förvaltningens förmåga att använda AI. Delrapport i regeringsuppdraget I2019/01416/DF || I2019/01020/DF (delvis).

Europaparlamentet (30 mars 2023) "General-purpose artificial intelligence. At a Glance" av Tambiama André Madiega. [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_ATA\(2023\)745708](https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2023)745708) [senast besökt 5 oktober 2023].

Europaparlamentet (14 juni 2023) "MEPs ready to negotiate first-ever rules for safe and transparent AI", <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai> [senast besökt 5 oktober 2023].

Europaparlamentet (2023) *Rättsakten om artificiell intelligens*. Europaparlamentets ändringar antagna den 14 juni 2023 av förslaget till Europaparlamentets och rådets förordning om harmoniserade regler för artificiell intelligens (rättsakt om artificiell intelligens) och om ändring av vissa unionslagstiftningsakter (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)). [https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236\\_SV.pdf](https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_SV.pdf) [senast besökt 18 oktober 2023].

Europaparlamentets och rådets förordning (EU) nr 182/2011 av den 16 februari 2011 om fastställande av allmänna regler och principer för medlemsstaternas kontroll av kommissionens utövande av sina genomförandebefogenheter.

Europaparlamentets resolution av den 25 mars 2021 om kommissionens utvärderingsrapport om genomförandet av den allmänna dataskyddsförordningen två år efter dess tillämpning ([2020/2717\[RSP\]](https://www.europarl.europa.eu/doceo/document/TA-9-2021-0236_SV.pdf)).

Faktapromemoria från regeringen till riksdagen (2020/21:FPM109). [https://www.riksdagen.se/sv/dokument-och-lagar/dokument/fakta-pm-om-eu-forslag/forordning-om-artificiell-intelligens\\_h806fpm109/](https://www.riksdagen.se/sv/dokument-och-lagar/dokument/fakta-pm-om-eu-forslag/forordning-om-artificiell-intelligens_h806fpm109/) [senast besökt 4 oktober 2023].

Future of Life Institute (22 mar 2023) "Pause Giant AI Experiments: An Open Letter", <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> [senast besökt 5 oktober 2023].

G7 (20 maj 2023) G7 Hiroshima Leaders' Communiqué, <https://www.consilium.europa.eu/en/press/press-releases/2023/05/20/g7-hiroshima-leaders-communique/> [senast besökt 5 oktober 2023].

G7 (30 oktober 2023). Hiroshima Process International Guiding Principles for Advanced AI system. <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system> [senast besökt 6 november 2023].

G7 (30 oktober 2023) Hiroshima Process International Code of Conduct for Advanced AI Systems. <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems> [senast besökt 6 november 2023].

HLEG, High-Level Expert Group on AI (2019) *Etiska riktlinjer för tillförlitlig AI*. EU-kommissionen.

IMY (2023) Federerad maskininlärning mellan två vårdgivare. Slutrapport om Integritetsskyddsmyndighetens pilotprojekt med regulatorisk testverksamhet om dataskydd. IMY-2023-2602 <https://www.imy.se/publikationer/slutrapport-om-imys-pilotprojekt-med-regulatorisk-testverksamhet-om-dataskydd/> [senast besökt 5 oktober 2023].

IMY (15 mars 2023) "Framgångsrik pilot med regulatorisk testverksamhet om AI", <https://www.imy.se/nyheter/framgangsrik-pilot-med-regulatorisk-testverksamhet-om-ai/> [senast besökt 5 oktober 2023].

LDP Headquarters for the Promotion of Digital Society (april 2023) The AI White Paper. Japan's National Strategy in the New Era of AI. Project Team on the Evolution and Implementation of AIs. [https://www.taira-m.jp/ldp/s%20ai%20whitepaper\\_etrans\\_2304.pdf](https://www.taira-m.jp/ldp/s%20ai%20whitepaper_etrans_2304.pdf).

Madiega, T. och Van De Pol, A.L. (2022) Artificial Intelligence Act and Regulatory Sandboxes. European Parliamentary Research Service. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS\\_BRI\(2022\)733544\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf).

MIT Technology Review (4 Oktober 2022) "The White House just unveiled a new AI Bill of Rights" av Melissa Heikkilä. <https://www.technologyreview.com/2022/10/04/1060600/white-house-ai-bill-of-rights> [senast besökt 5 oktober 2023].

The New York Times (25 september 2023) "Amazon Takes a Big Stake in the A.I. Start-Up Anthropic" av Adam Satariano and Cade Metz. <https://www.nytimes.com/2023/09/25/technology/amazon-anthropic-ai-deal.html> [senast besökt 5 oktober 2023].

Nesta (2017). *A Working Model for Anticipatory Regulation*. London: Nesta. [https://media.nesta.org.uk/documents/working\\_model\\_for\\_anticipatory\\_regulation\\_0.pdf](https://media.nesta.org.uk/documents/working_model_for_anticipatory_regulation_0.pdf).

NIST (26 januari 2023) The NIST AI Risk Management Framework. <https://www.nist.gov/itl/ai-risk-management-framework>.

OECD (7 september 2023a) *G7 Hiroshima Process on Generative Artificial Intelligence (AI). Towards a G/ Common Understanding on Generative AI*. Report prepared for the 2023 Japanese G7 Presidency and the G7 Digital and Tech Working Group.

OECD (september 2023b) Initial policy considerations for generative artificial intelligence, OECD. [https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence\\_fae2d1e6-en](https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en).

OECD (juli 2023c) Regulatory sandboxes in artificial intelligence. <https://www.oecd.org/sti/regulatory-sandboxes-in-artificial-intelligence-8f80a0e6-en.htm>.

Office for Artificial Intelligence (3 augusti 2023) *A pro-innovation approach to AI regulation*. Parliament by the Secretary of State for Science, Innovation and Technology by Command of His Majesty. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.

An Open Letter on the UK's Global Summit on AI Safety (November 2023). The letter has been convened by Connected by Data, The Trades Union Congress and Open Rights Group in response to the UK Government's summit being held on the 1st and 2nd November. 2023 <https://ai-summit-open-letter.info> [senast besökt 6 november 2023].

Regeringskansliet (2021) Remiss av Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens. <https://www.regeringen.se/remisser/2021/05/remiss-av-europeiska-kommissionens-forslag-till-forordning-om-harmoniserade-regler-for-artificiell-intelligens/> [senast besökt 6 november 2023].

Regeringskansliet Faktapromemoria (2022) *Direktiv om skadeståndsansvar gällande artificiell intelligens*.

2022/23:FPM8. <https://www.regeringen.se/faktapromemoria/2022/11/202223fpm8/>.

SCB, Statistiska centralbyrån (2020) *Artificiell intelligens i Sverige*.

SOU 2023:2 (31 januari 2021) En inre marknad för digitala tjänster – ansvarsfördelningen mellan myndigheter

Techzine (30 juni 2023) "150 European companies write open letter on dangers of AI Act", av Laura Herijgers. <https://www.techzine.eu/blogs/analytics/108372/150-european-companies-write-open-letter-on-dangers-of-ai-act/> [senast besökt 5 oktober 2023].

Tillväxtanalys (2022) *AI-politik för konkurrenskraft*, Rapport AU 2022:02:01.

Tillväxtanalys (2023) *Hur omformar AI näringslivet och hur kan politiken utvecklas?*

Serienummer: Rapport 2023:04. <https://www.tillvaxtanalys.se/studieomraden/projekt-klara/genomforda-ramprojekt/2023-04-27-hur-omformar-ai-naringslivet-och-hur-kan-politiken-utvecklas.html>.

TIME (5 februari, 2023) "The Creator of ChatGPT Thinks AI Should Be Regulated" av John Simons. <https://time.com/6252404/mira-murati-chatgpt-openai-interview/> [senast besökt 14 september, 2023].

TIME (20 juni 2023) "Exclusive: OpenAI Lobbied the E.U. to Water Down AI Regulation" av Billy Perrigo <https://time.com/6288245/openai-eu-lobbying-ai-act/>.

Washington Post (16 maj 2023) "CEO behind ChatGPT warns Congress AI could cause 'harm to the world'" av Cat Zakrzewski, Cristiano Lima and Will Oremus.

<https://www.washingtonpost.com/technology/2023/05/16/sam-altman-open-ai-congress-hearing/> [senast besökt 5 oktober 2023].

The White House (2016) The Administration's Report on the Future of Artificial Intelligence. <https://obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence> [senast besökt 5 oktober 2023].

The White House (4 oktober) "Blueprint for an AI Bill of Rights".

<https://www.whitehouse.gov/ostp/ai-bill-of-rights/> [senast besökt 5 oktober 2023].

The White House (21 juli 2023) FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> [senast besökt 5 oktober 2023].

The White House (30 oktober 2023) FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.

<https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> [senast besökt 6 november 2023].

The White House (12 september 2023) FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Eight Additional Artificial Intelligence Companies



to Manage the Risks **Posed** by AI. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> [senast besökt 5 oktober 2023].

World Economic Forum (juni 2023) *The Presidio Recommendations on Responsible Generative AI*. White Paper. <https://www.weforum.org/whitepapers/the-presidio-recommendations-on-responsible-generative-ai/>.

U.S. Senate Judiciary Committee Subcommittee on Privacy, Technology and the Law. Hearing: "Oversight of A.I.: Rules for Artificial Intelligence." Utskrift av senatsförhöret med Sam Altman den 16 maj, 2023: <https://techpolicy.press/transcript-senate-judiciary-subcommittee-hearing-on-oversight-of-ai/>.

På vilket sätt statens insatser bidrar till svensk tillväxt och näringslivsutveckling står i fokus för våra rapporter.

Läs mer om vilka vi är och vad nyttan med det vi gör är på [www.tillvaxtanalys.se](http://www.tillvaxtanalys.se). Du kan även följa oss på LinkedIn och YouTube.

Anmäl dig gärna till vårt [nyhetsbrev](#) för att hålla dig uppdaterad om pågående och planerade analys- och utvärderingsprojekt.

Varmt välkommen att kontakta oss!



**Tillväxtanalys**  
Studentplan 3, 831 40 Östersund  
Telefon: 010-447 44 00  
E-post: [info@tillvaxtanalys.se](mailto:info@tillvaxtanalys.se)  
Webb: [www.tillvaxtanalys.se](http://www.tillvaxtanalys.se)