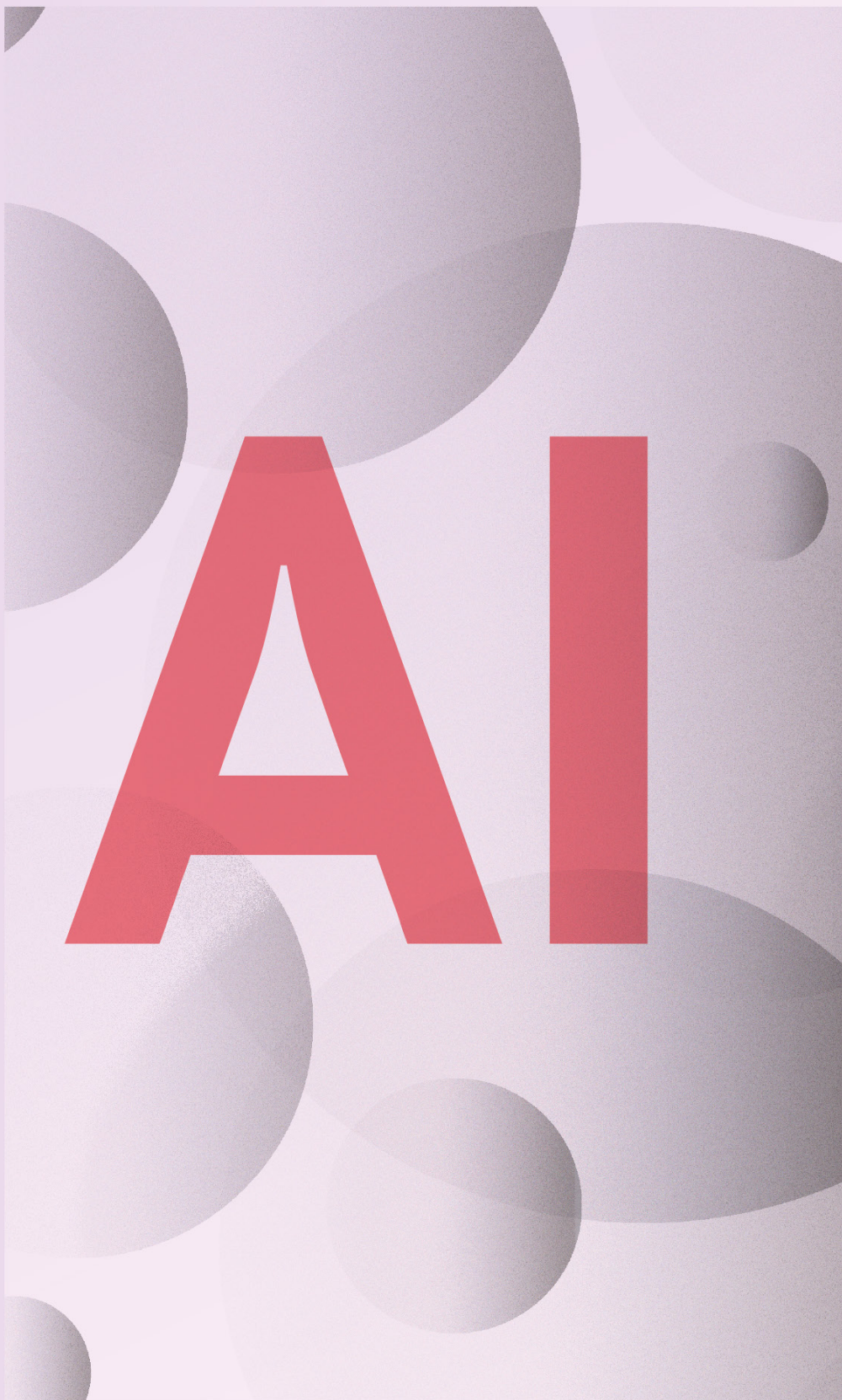


Begynnelsen av ett paradigmskifte

En rapport från Teknikföretagen



Publicerad Oktober 2023

Teknikföretagen

AI är här för att stanna

De flesta av oss har haft att göra med väldigt kraftfull AI under flera år. AI har styrvt vilka bilder du får se på Instagram, vilka inlägg du får se på Twitter och vilka annonser som dyker upp i ditt Facebook-flöde. Så varför pratar alla helt plötsligt om AI? En viktig anledning är att det som kallas generativ AI det senaste året har tagit stora kliv framåt. Med generativ AI menas AI som kan generera till exempel text, bilder och ljud – som alla kan verka förvånansvärt mänskliga, och flera sådana tjänster har under det föregående året blivit tillgängliga för allmänheten.

Den kanske mest omtalade av dessa generativa AI, ChatGPT, blev i februari 2023 den snabbast växande tjänsten någonsin, med 100 miljoner användare efter bara två månader. Det kan jämföras med TikTok som fick samma antal användare först efter nio månader, Instagram som behövde mer än två år, och Facebook, som behövde mer än fyra år för att nå 100 miljoner användare. Att denna nya generativa AI-teknik fullkomligen tagit världen med storm råder det alltså inga tvivel om.

Med denna utveckling har också frågetecken kring säkerhet och oönskade konsekvenser av en snabb AI-utveckling fått stort medialt utrymme. Är superintelligent AI slutet för mänskligheten, eller innebär det *bara* ett i grunden radikalt annorlunda samhälle där AI kan göra allt som en människa kan, men mycket bättre?

Frågorna ovan kommer vi inte besvara – däremot för vi resonemang om dem utifrån vad namnkunniga personer på AI-scenen tycker och tänker. Rapporten går inte heller på djupet vad gäller tekniska detaljer, det är helt enkelt för komplicerat. I stället vill Teknikföretagen med rapporten bidra till kunskapshöjning hos industrin, politiken och den intresserade allmänheten. För att hantera ett så komplicerat ämne mot en så bred målgrupp vill vi berätta den populärvetenskapliga versionen om hur AI fungerar, vad för sorts AI det finns och vad dagens AI-tjänster kan göra (del 1 och 2). Därefter följer en diskussion om de eventuella riskerna med AI och vad beslutsfattare bör ha i åtanke när AI-utvecklingen går framåt i en rasande fart (del 3). Avslutningsvis resonerar vi om industrins och näringslivets tillämpning av AI (del 4).

Gustav Wiel-Berggren
Policystrateg Teknikföretagen

Del 1 – Så fungerar AI

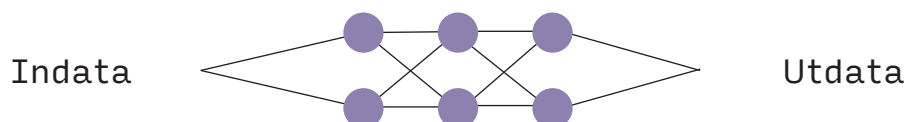
En enkel förklaring av något komplicerat

Du är mer lik en AI än du kanske tror – i alla fall din hjärna. Den mänskliga hjärnan är uppbyggd av neuroner som är sammankopplade i ett nätverk och skickar och tar emot nervimpulser – ett neuronät. Detta försöker man efterlikna i Artificiella neuronät – alltså en AI-hjärna. I detta avsnitt förklarar vi, grovt förenklat, principen bakom AI.

Ett artificiellt neuronät är expert på att känna igen mönster, och ur detta följer en lång rad förmågor. Men i dess mest grundläggande funktionalitet kan AI reduceras till detta: *en indata matas in i neuronätet, sedan sker beräkningar som resulterar i en utdata*. Till exempel kan du fråga röstassistenten i din telefon hur vädret blir imorgon (indata), där-efter sker beräkningar som resulterar i ett svar (utdata). Men indatan skulle lika gärna kunna vara en bild, en video eller en skriftlig fråga – och även utdatan kan variera i format.

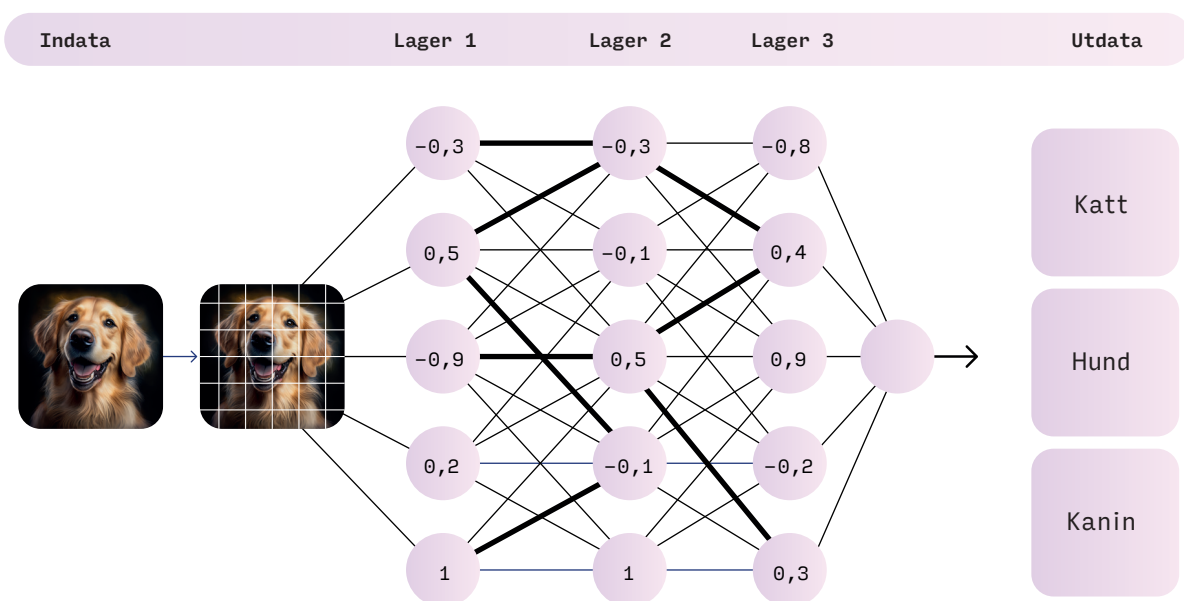
Indata → Beräkningar → Utdata

Så hur sker egentligen de beräkningar som kan ge AI en närmast övermänsklig förmåga? Enklast kan vi beskriva neuronerna i det artificiella neuronätet som noder med siffervärden som kan ändras (se lila cirklar i figuren nedan). Beroende på vilken indata som matas in så ändrar noderna i det första lagret värde, och beroende på hur kopplingarna till nästa lager av noder ser ut, så uppdateras värdena. När datan har passerat genom alla lager av noder har beräkningarna till slut resulterat i en utdata.



Illustrationen nedan fördjupar resonemanget genom att ta ett exempel med en AI som kan känna igen bilder. Om vi matar in en bild på en *golden retriever* i vår exempel-AI så kommer bilden först delas in i många olika mindre rutor. Beroende på färgen på dessa rutor aktiveras noderna i det första lagret olika mycket, genom att ge dem ett värde mellan -1 och 1. Värdet

skickas sedan vidare genom lagren av noder och antar nya värden beroende på hur kopplingarna noderna emellan ser ut (så kallad *viktning*). I andra änden av neuronnätet kommer sedan en utdata, som gissar vad bilden föreställde. I illustrationen nedan har neuronnätets beräkningar lett fram till att bilden sannolikt föreställer en hund. Neuronnätet i vårt ex-



empel har därför gissat rätt!

Men om neuronnätet i illustrationen i stället hade gissat fel, att bilden föreställer en katt, så rättas felet. Viktningen av kopplingarna mellan noder uppdateras, så att beräkningarna landar i att bilden föreställer en hund. Det här är vad som kallas att *träna en AI*: vi matar in många bilder på många olika hundar i neuronnätet, och

varje gång den gissar fel så påpekar vi detta – så att neuronnätet kan uppdatera viktningen mellan noderna och beräkningarna i stället producerar en utdata som säger att bilden mest sannolikt föreställer en hund. Ett neuronnät som tränats på stora mängder data blir därför väldigt duktig på att gissa, för ju mer träning, desto bättre gissningar.

Världens mest avancerade mjukvara?

Innan vi går vidare måste vi prata om GPT och GPT-4. I juni 2017 publicerade AI-forskare på Google en vetenskaplig studie om en ny AI-arkitektur: en så kallad *generative pre-trained transformer* – senare förkortat GPT.¹ Dessa faller inom kategorin språkmodeller, och tränas på stora mängder text. Året därpå släppte det amerikanska företaget OpenAI sin första GPT-modell. Och det är här stora kliv tas mot dagens avancerande generativa AI. För AI kräver stora mängder träningsdata, och tidigare språkmodeller var beroende av betydande manuell etikettering av den data man använde för att träna. Men med GPT-1 valde OpenAI ett annat tillvägagångssätt. Genom att först träna AI:n på stora mängder data och sedan finjustera resultaten med en mindre mängd noga utvald träningsdata kunde mängden träningsdata ökas avsevärt.²

För att illustrera hur avancerad en AI är talas det ibland om hur många *parametrar* den har. En parameter kan liknas vid hur relationerna mellan noder i det artificiella neuronätet ser ut.

OpenAI:s första version, GPT-1, har 117 miljoner parametrar. När ChatGPT släpptes, byggd på GPT-3.5, hade den imponerande 175 miljarder parametrar. I mars 2023 släppte de GPT-4, och vissa källor hävdar att GPT-4 har upp till 170 tusen miljarder parametrar. OpenAI har dock

inte själva bekräftat någon siffra, i stället har de snarast dementerat att det skulle handla om 1 000 miljarder parametrar.³

ChatGPT blev, som nämns i rapportens inledning, den tjänst som snabbast i världen nått 100 miljoner registrerade användare. Den har bidragit till att lyfta upp AI i rampljuset och fört med sig en flodvåg av AI-nyheter till radionyheter och dagstidningar. GPT-4 är, som OpenAI:s vd Sam Altman säger i en intervju med AI-forskaren Lex Fridman, förmodligen världens mest avancerade mjukvara⁴ – och den finns tillgänglig för alla som vill testa den. Det är heller inte ovanligt att en konversation med GPT-4 är förknippat med en personlig heureka-upplevelse.

Nivåer av intelligens

Det finns flera olika sätt att kategorisera AI, men en pedagogisk uppdelning som lämpar sig väl i detta sammanhang är *svag* och *stark*. Utifrån de exempel på modern AI som förekommer senare i denna rapport hade det inte varit konstigt att utgå från att dagens AI faller inom kategorin stark, men så är faktiskt inte fallet. Detta trots att dagens AI i många fall är så imponerande att de kan få en att tappa andan.

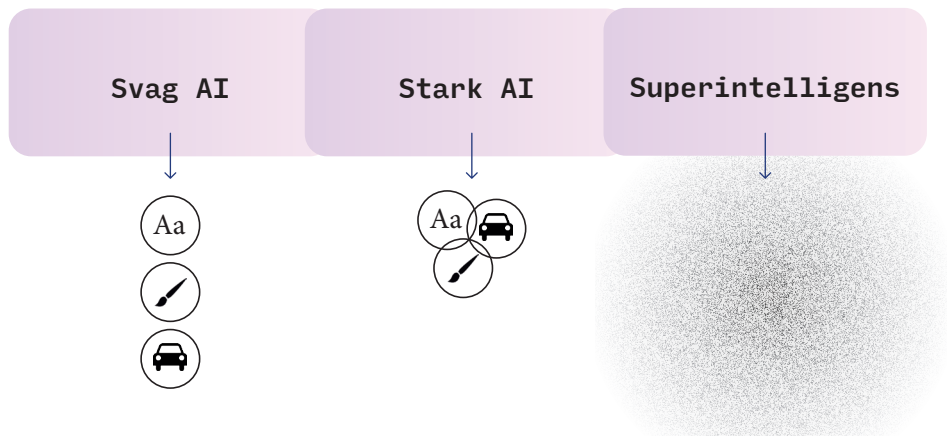
Svag AI är beteckningen på en AI som är skapad för att utföra en specifik uppgift, därför betecknas den ibland – kanske mer träffsäkert – även som *snäv* AI. Den är helt enkelt begränsad till den eller de uppgifter

för vilken den är programmerad. En svag AI som är byggd för att generera text, skapa grafik eller köra en bil kan göra just detta och den kan i många fall göra det betydligt bättre än en människa – men den kan inte lösa, eller ens förstå, andra problem vi ställer den inför.

Stark AI däremot är en AI på den nivå vi hittills endast sett i filmer, som HAL9000 Stanley Kubricks *2001: Ett rymdäventyr* (1968) eller Alicia Vikanders karaktär Ava i filmen *Ex Machina* (2014). Det är en AI som kan ta sig an samma intellektuella uppgifter som människor, en generell intelligens som inte behöver tränas i ett särskilt syfte, utan som har kapacitet att fungera just som en människa. En viktig term att känna till här, som blir alltmer använd när man talar om utmaningar och eventuella

hot med AI, är AGI – *artificiell generell intelligens*.

Men även om kategoriseringen av AI framför allt brukar handla om svag/snäv respektive stark/AGI så talas det i mer filosofiska sammanhang även om ytterligare en nivå: ASI, *artificiell superintelligens*, en nivå av intelligens som vida överträffar mänsklig kapacitet inom alla tänkbara fält. Det brukar också spekuleras i att en AGI, med tanke på att den sannolikt är bättre än en människa på att programmera, snart kommer ha skapat en superintelligens.



Del 2 – Hur bra är dagens AI?

Vad kan AI göra?

Att i tryckt format försöka berätta vilka förmågor AI har är att göra sig själv en otjänst. Utvecklingen går nämligen så pass snabbt framåt att man riskerar att bli inaktuell om en månad, eller ännu värre, i morgon. Världens techjättar arbetar för fullt med att utveckla nya AI-tjänster: programvaror många arbetar med varje dag, som Microsoft Word, Excel, Teams och Outlook har redan distribuerats i testversioner som har GPT-4 integrerat. Adobe Photoshop har nyligen släppts med en inbyggd generativ AI, och fler program vi använder dagligen kommer sannolikt snart ha funktioner som med talspråk låser upp även de mest avancerade funktioner i våra digitala verktyg. För den som har ett kontorsjobb kommer AI förändra hur vi arbetar – och ha stor potential att göra oss mer effektiva och bättre på våra jobb.

Trots att nya AI-tjänster dyker upp mest hela tiden kommer detta avsnitt ge några exempel på hur pass kraftfulla några av dagens AI-system faktiskt är. Det har sannolikt kommit nya och imponerande AI-tjänster när du läser denna skrift, men förhoppningsvis kan detta avsnitt bidra till att förmedla hur pass genomgripande effekter AI kommer ha på samhälle och näringsliv.

Slutet för programmeraren?

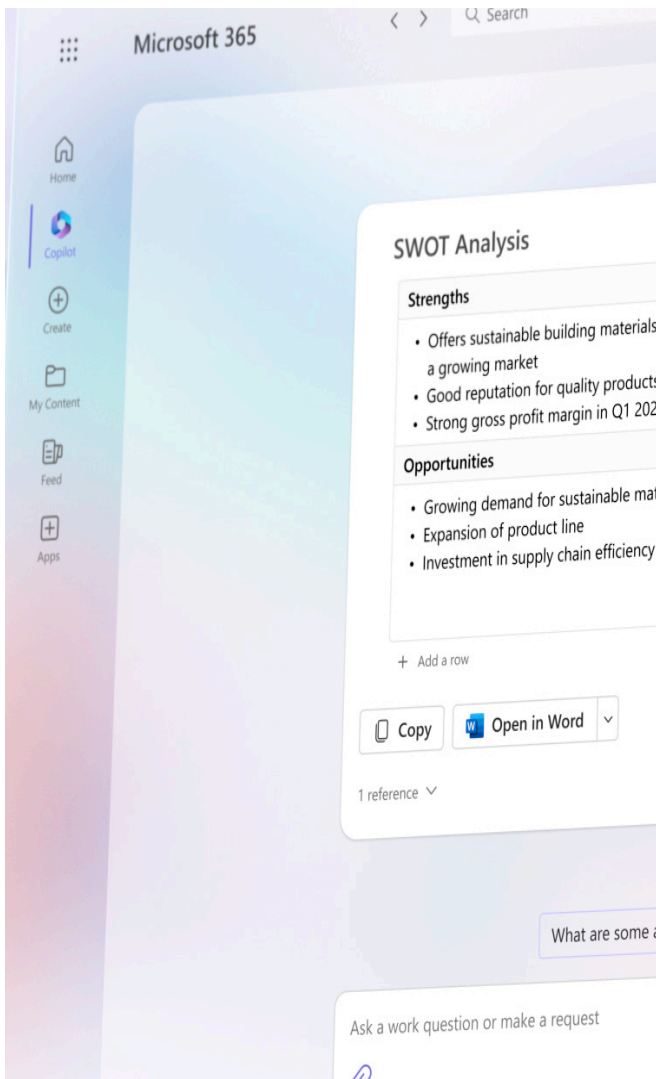
De flesta som vet hur man programmerar appar till en iPhone har sannolikt en högskoleutbildning eller ett långt programmeringsintresse i ryggsäcken. Därför är det imponerande att GPT-4 kan göra lekmän till programmerare.

Men hur skapar man då en iPhone-app? Genom att ställa frågan till ChatGPT presenterades en pedagogisk text om hur man installerar programmet Xcode, som behövs för att skapa appar till Apple-produkter. Därefter fick GPT-4 följande instruktion:

”Koda en iPhone-app som slumpmässigt väljer fem filmer från The Movie Database (TMDb) och presenterar dem en och en med en swipe-funktion. Appen ska hämta relevant information från TMDb. Appen ska visa en bakgrundsbild från filmen samt filmomslaget och titel, sedan i mindre text länk till trailer, filmens betyg, premiärdatum och en kort beskrivning av filmens handling (max 230 tecken)”.

ChatGPT svarade med ett långt sjok kod i språket Swift och beskrev hur denna kod skulle kopieras till Xcode. Här stötte den dock på problem: i stället för att visa en färdig app svarade Xcode med felmeddelanden. Det krävdes något som kallas för en API-nyckel, och så fanns det ett par fel i koden. ChatGPT bad om ursäkt för missen och skrev en instruktion för hur man skaffar en API-nyckel och rättade felen utifrån felmeddelanden från Xcode. Efter en kort dialog med vad som upplevs som en mänsklig programmeringsassistent presenterades en fullt fungerande app i Xcode.





Kommer kontorsarbete någonsin bli sig likt?

Tänk om du efter ett videomöte kunde be Microsoft Teams att skriva en projektplan utifrån det ni precis pratat om under den senaste timmens videomöte och skicka denna i ett mejl till din chef, ihop med en Powerpoint-presentation som argumenterar för att just detta projekt ska få grönt ljus.

Det går faktiskt redan, men funktionaliteten har bara släppts till få utvalda. Med Microsoft 365 Copilot har Microsoft integrerat GPT-4 i Office-paketet.⁵ En förväntad effekt är att fler kommer bli bättre på att verkligen använda de digitala verktygens fulla kapacitet – under Microsofts keynote, där de presenterade 365 Copilot berättade de att de flesta av deras kunder endast använder 10 procent av programmets funktionalitet – men med Copilot läses hela utbudet av avancerande funktioner bara genom att prata med programmet.⁶

Kreativ Intelligens

Att skriva kod eller transkribera ett videomöte är imponerande, men det är också uppgifter som inte nödvändigtvis kräver så mycket kreativitet. Dokumentation kring kod finns det gott om, och att transkribera handlar ju bara om att höra vad som sägs och översätta till text. Men ChatGPT kan även lösa kreativa uppgifter som den aldrig förr ställts inför, och har till och med en förmåga att skämta. Det visade Teknikföretagens tidigare vd i ett blogginlägg där han bad ChatGPT att skriva ett skämt och en haiku om svensk industri.⁷

KW

Skriv en haiku om svensk industri

GPT

Maskiner som ljuder,
Svensk industri blomstrar,
Framtiden byggs här

KW

Skriv ett skämt om tillverkningsindustrin

GPT

Varför var tillverkningsindustrin alltid trött på morgonen? För den var alltid uppe och körde hela natten!

“Imagine Darth Vader on the runway during fashion week”

En snabbt växande kategori AI-tjänster är de som kallas diffusionsmodeller. Bland dem finns tjänster som Stable Diffusion, Midjourney och DALL-E 2, som kan skapa en oändlig mängd nya fotorealistiska bilder såväl som stilistiska illustrationer utifrån en text. Flera av de framstående tjänsterna är tränade på datamängden LAION-5B, där 5B står för antalet bilder i datamängden: fler än fem miljarder.⁸ Bilden nedan är skapad med detta avsnitts rubrik som instruktion i AI-tjänsten Midjourney.

Diffusionsmodeller tränas genom att visas en bild och sedan successivt addera brus till bilden. Därefter får AI:n i uppgift att utgå från bruset och successivt redu-

cera bruset tills den har lyckats återskapa originalbilden. När modellen tränats färdigt och ska skapa en helt ny och originell bild genererar den först slumpmässigt brus, och reducerar successivt bruset tills den skapat en bild som överensstämmer med textbeskrivningen.⁹ Resultatet är att AI redan idag kan skapa i stort sett alla bilder du ber den om.

Men finns det någon nytta med detta? Man kan tänka sig att dessa tjänster snart kommer kunna utmana formgivare och illustratörer och kanske skapa ett nytt budgetsegment inom grafiska tjänster, där AI stått för en stor del av den kreativa processen. Kanske kommer alla med en idé om hur något ska se ut kunna bli sin egen art director.



Bild från Midjourney

Bilder skapade med Midjourney



A picture of a happy golden retriever



A child's drawing of a golden retriever



The pope as a rapper on a dimly lit stage



The pope in a space suit from the Catholic church

Vems röst kan du lita på?

I mitten av april släppte världsartisterna Drake och The Weeknd en ny låt. "Heart on my sleeve" tog fansen med storm och fick på kort tid 15 miljoner spelningar på sociala medieplattformen TikTok. Men det dröjde inte lång tid innan skivbolaget Universal Music utövade påtryckningar mot TikTok, YouTube och Spotify för att ta bort låten från deras tjänster. Anledningen? Trots att det låter precis som att Drake rap-

par och The Weeknd sjunger, så är det inte deras röster. Det är en AI som har tränats att låta som dem, och ingen av artisterna har något med låten att göra.¹⁰

Förutom text och bild har det alltså gjorts stora framsteg även inom röstimiterande AI, som kan lära sig hur en specifik person låter genom att lyssna på dennes röst. Därefter kan den läsa vilken text som helst med denna persons röst, alternativt byta ut en persons röst mot en annan persons röst i en ljudfil.

Den idag kanske mest avancerade röstimiterande AI:n heter VALL-E och är framtagen av forskare hos Microsoft. Med 60 000 timmar ljudböcker, uppläst av 7 000 olika engelsktalande personer så har VALL-E fått tillgång till träningsdata som är flera hundra gånger större än motsvarande system.¹¹ Tack vare det omfattande träningsmaterialet så har VALL-E blivit extremt duktig på att imitera röster, och behöver bara höra en tre sekunder lång inspelning av en röst för att kunna läsa upp vilken text som helst med just denna röst.

Att AI kan lära sig en röst är dock förvisso inte bara av ondo. Tvärt om har Apple nyligen släppt en tjänst för att låta din telefon lära sig din röst. Syftet är att de som har svårt att prata, eller av medicinska skäl kommer att bli av med sin röst, ska kunna använda fortsätta använda sin röst och behålla en del av sin identitet med *en text till tal*-funktion i sin telefon.

En vårdsrevolution i antågande?

Artificiell intelligens kommer ha stor inverkan på sjukvården – både som verktyg i den kliniska vården och som verktyg i medicinsk forskning. Förutom de exempel som nämns nedan finns en lång rad andra forskningsprojekt där AI verkar kunna bidra till stora genombrott. Inte minst en AI vid namn AlphaFold, påstår vissa, kommer förändra *allt*.¹² Genom att ha löst det så kallade proteinvecklingsproblemet har AlphaFold kanske givit forskare nyckeln till att förstå det biologiska livets bestånds-

delar och hur vi med denna kunskap kan angripa sjukdomar eller skraddarsy enzymer för att, till exempel, bryta ned plast. Ännu är detta fortfarande i ett tidigt stadie, men forskare menar att man ”verkligen kan tala om en revolution”.¹³

Sverige först i världen med AI för mammografiscreening

”AI-algoritmen hittar mönster i bilderna som vi mänskliga radiologer inte gör. Den känner igen mönster i mammografibilden som vi inte kan se med våra egna ögon”. Det säger Karin Dembrower, bröstradiolog på Capio S:t Görans Sjukhus och forskare som undersökt möjligheten att använda AI för att hitta bröstcancer. Studien, som gjorts i samarbete med Karolinska Institutet, har undersökt hur en AI presterar vid screening av mammografibilder. Resultatet är mycket positivt. Med en AI-assistent hittar man mer cancer utan att öka antalet onödiga undersökningar. Nu blir S:t Görans Sjukhus först i världen med att använda AI för att i större utsträckning hitta bröstcancer vid mammografiscreening.¹⁴

AI som vapen mot multiresistenta bakterier

Enligt en brittisk regeringsrapport som granskat antibiotikaresistens skulle antalet som dör på grund av antibiotikaresistens kunna uppgå till 10 miljoner människor årligen redan 2050 – fler än antalet som dör i cancer varje år.¹⁵ Risken att antibiotikaresistens blir en av de vanligaste dödsorsakerna globalt inom bara några decennier är högst reell.

Men nu har AI bidragit till ett medicintekniskt genombrott. Forskare vid MIT och McMaster University har utsatt den multiresistenta bakterien *A. baumannii* för 7 500 olika kemiska föreningar och noterat om bakteriens tillväxt fortsatt eller avstannat. Därefter matade de in de kemiska föreningarnas molekylstruktur i sin AI-modell, ihop med information om huruvida den hade påverkat bakteriens tillväxt. Därmed var träningsdatan klar.

För att sedan hitta en antibiotika som kan döda *A. baumannii* lät de AI-modellen analysera molekylstrukturen i 6 680 kemiska föreningar den tidigare inte sett. Analysen tog mindre än två timmar och ringade in 240 potentiella strukturer som forskarna sedan testade i labbmiljö. Testen kokade ned resultatet till nio antibiotikum, varav en var extremt effektiv på att döda *A. baumannii*, utan att ha någon effekt på andra bakterier. Den effektiva kemiska föreningen hade tidigare testats som ett

diabetesläkemedel, men genom att granska dess molekylstruktur kunde AI-modellen ringa in just denna som en med potential att vara effektiv mot *A. baumannii*.¹⁶

Just *A. baumannii* beskrivs av forskarna i studien som ”allmänhetens fiende nummer 1” bland multiresistenta bakterier och omskrivs av The Guardian, i sin rapportering om denna nyhet, som ”en dödlig superbakterie”.¹⁷ Upptäckten kan nog därmed bedömas som i grunden revolutionerande i kampen mot multiresistenta bakterier.

Forskningsrevolution

Vid Lawrence Berkley National Laboratory (LBNL) i San Francisco, och mer specifikt det så kallade A-labbet, finns en AI som helt på egen hand försöker skapa nya material för användning inom bland annat batteriteknik.

Men hur fungerar egentligen en autonom materialforskningsrobot? Först och främst är den tränad på mänsklighetens samlade kunskaper i kemi, och precis som med andra AI kan den se samband som är svåra för oss människor att upptäcka. Utifrån sina kunskaper hittar den sedan på recept för nya material med specifika egenskaper. Till sin hjälp har den en robotarm, 200 material i pulverform, en smältugn och flera olika gaser. Enligt forskarna i projektet har A-labbet ökat antalet nya material de tar fram hundrafaldigt jämfört med när verksamheten bedrevs av AI:ns mänskliga kollegor.¹⁸

Nu kan AI läsa dina tankar

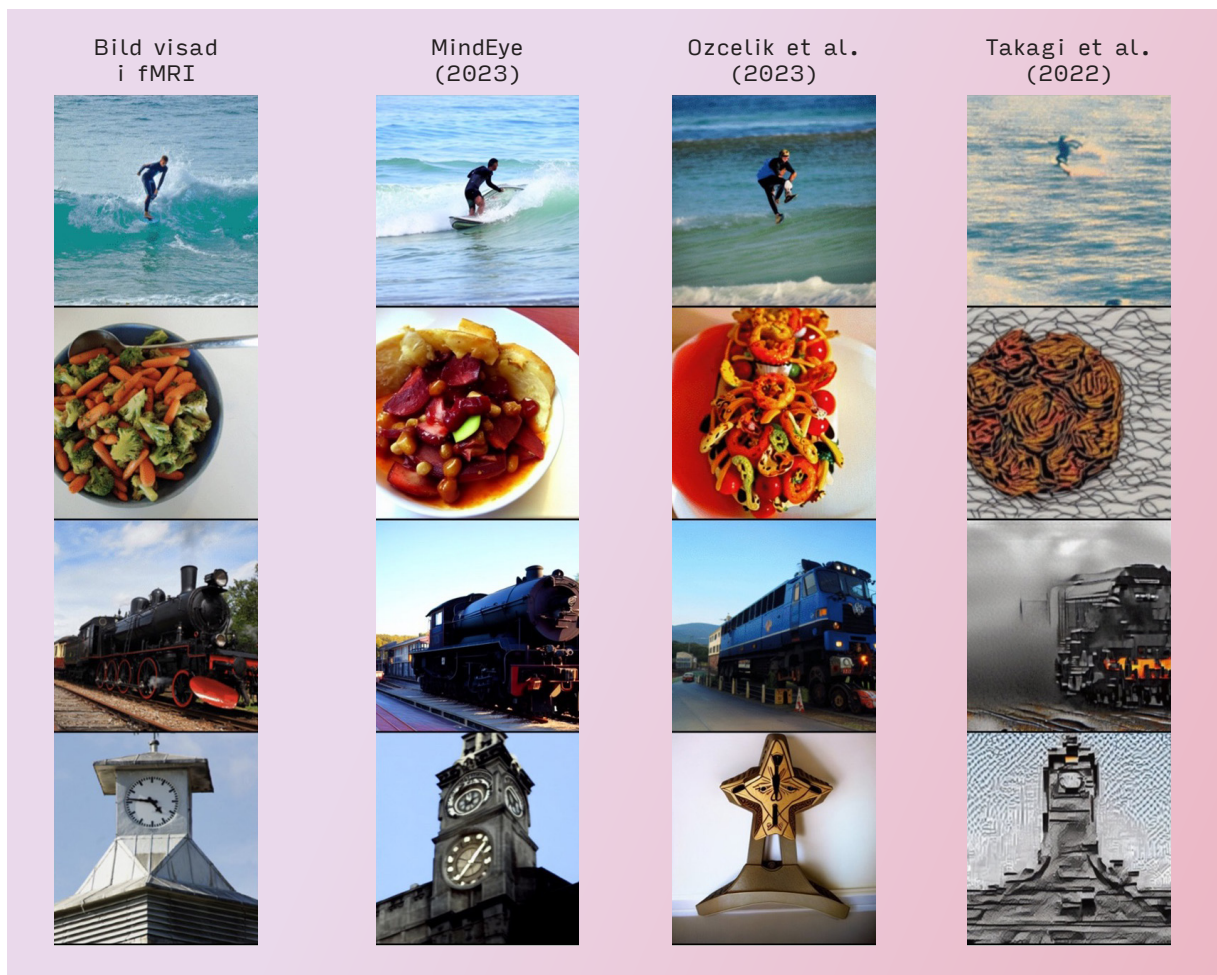
En AI fungerar som bekant så pass bra för att neuronät tränas för att känna igen mönster. Det är så den kan veta vad en bild föreställer, lära sig fakta och identifiera anomalier på en röntgenplåt. Men forskare har också tränat AI på att känna igen mönster i hjärnans blodflöde vid så kallad funktionell magnetröntgen, fMRI, och resultatet är chockerande. AI kan nu, endast genom att studera din hjärnaktivitet åter skapa det du ser framför dina ögon.

Genom att låta en person titta på en bild medan hjärnaktivitet mättes med fMRI kunde forskare vid Osaka University 2022 åter skapa mycket snarlika bilder med hjälp av en diffusionsmodell. Den vänstra

kolumnen av bilderna nedan presenterades för en person i en fMRI-undersökning, och övriga kolumnerna har genererats av forskarlag genom en diffusionsmodell, endast genom att studera blodflödet i personens hjärna.¹⁹

Sedan forskarna vid Osaka, Takagi et al., presenterat sina resultat 2022 har mycket hänt på området. Amerikanska forskare presenterade den 30 maj 2023 MindEye, som med större träffsäkerhet än Osaka-forskarna tacklat samma problem – och fått än bättre resultat.²⁰

Men det är inte allt. Forskare vid University of Texas at Austin har gjort stora fram-



steg vad gäller att avkoda hjärnaktivitet och omvandla det till text med en språkmodell liknande ChatGPT.

Genom att samla 16 timmar hjärnaktivitet från tre testpersoner som fick lyssna på podcasts, kunde forskare låta en AI analysera vilken hjärnaktivitet som hängde samman med vilket ord i podcasten. Därefter testades språkmodellens förmåga att avkoda tre olika aktiviteter: en inre monolog, alltså när en person talar med sig själv genom tankekraft utan att säga något, hur testpersonen uppfattar talat språk samt hjärnaktiviteten från att se ett filmklipp utan ljud.

Bäst var AI:n på att avkoda hur testpersonerna uppfattade talat språk, med 72–82 procents träffsäkerhet. Träffsäkerheten gäller dock inte ord för ord, till stor del handlade det snarare om att budskapet var detsamma, även om AI:n ibland också

prickade rätt ord. AI:n var likväl duktig på att avkoda testpersonernas inre monologer med 41–74 procents träffsäkerhet. Sämst presterade den när det kom till att avkoda filmklipp utan ljud, med 21–45 procents träffsäkerhet av testpersonens tolkning av filmklippen.²¹

Även om träffsäkerheten varierar från hög till låg så är det viktigt att ha i åtanke den kraftigt begränsade datamängden, för ju mer träningsdata desto bättre resultat, och i fallet med forskarna vid University of Texas at Austin hade de bara 16 timmars träningsdata. Hur resultatet hade sett ut om datamängden liknade VALL-E:s datamängd i storlek, med 60 000 timmar från 7 000 personer, kan vi idag endast spekulera i. Framtiden lär dock visa vilket resultat det skulle innebära.



Del 3 – Hot och politik

Kan AI verkligen utgöra ett hot mot vår existens?

Våren 2023 skapade ett antal forskare och entreprenörer rubriker när de i ett öppet brev uppmuntrade AI-utvecklare att pausa arbetet i sex månader. Brevet, som togs fram på initiativ av den svenske fysikern, kosmologen och AI-forskaren Max Tegmark, fick signaturer bland annat från Elon Musk och Apple-medgrundaren Steve Wozniak.²² Flera av brevets tusentals undertecknare anser att AI, bortom alla fantastiska möjligheter, också är ett verkligt hot mot vår existens. Brevsignaturerna är heller inte ensamma, en ny studie visade att hälften av tillfrågade AI-forskare uppskattade risken att AI utrotar mänskligheten är större än 10 procent.²³

Diskussionen om AI och dess eventuella hot mot människan blir lätt filosofisk – och känns nästan löjeväckande. Först och främst börjar diskussionen med ett stort *om*. För det är endast *om* människan lyckas skapa en generell intelligens med ett eget medvetande och en egen agens som frågeställningar kring hur vi samexisterar med denna potentiellt överlägsna intelligens faktiskt blir relevanta. Här finns det en tydlig uppdelning bland forskare inom AI-fältet, och på båda sidor står tunga namn. Förutom redan nämnda undertecknare av paus-brevet så har även Geoffrey Hinton, ofta kallad den artificiella intelligensens gudfader, lämnat sin tjänst på Google för

att varna för AI-utvecklingen.²⁴ Likaså varnar den tidigare OpenAI-medarbetaren Paul Christiano, som ledde OpenAI:s arbete med att se till att ChatGPT:s svar överensstämde med mänskliga intressen (så kallad AI alignment), för att det finns en 10–20 procentig risk att vi får ett AI-maktovertag som leder till att många, eller de flesta människorna, kommer att dö.²⁵

Sam Altman, vd för OpenAI, skriver i ett blogginlägg att de blir mer och mer vaksamma och försiktiga allt eftersom deras system närmar sig AGI-nivå.²⁶ Altman beskriver hur de har strukturerat OpenAI för att inte släppa en tjänst som är ”catastrophically dangerous”, och att han välkomnar öppna standarder för AGI, bland annat om hur man avgör att en AGI är säker att tillgängliggöras för allmänheten.²⁷

Samtidigt finns det de som inte alls håller med om att AI är ett hot på sådant sätt som Hinton eller Tegmark gör. En av de mest framstående är måhända Yann LeCun, AI-forskningschef på Facebooks moderbolag Meta, som i sina sociala mediekonkanaler avfärdar risker med AI som kraftigt överdrivna. I återkommande trådar på X (Twitter) beskriver han det som helt självklart att superintelligenta maskiner kommer vara i linje med mänskliga värderingar och att mänsklighetens undergång som konsekvens av AI är en helt orimlig tanke.²⁸ Han kallar även tron på en självförbättrande AI, alltså en AGI som leder till en superintelligens och sedan tar över världen, för löjeväckande – och motiverar det bland annat med att en AI vars beräkningska-

pacitet ökar exponentiellt också kommer att behöva exponentiellt mer hårdvara och elektricitet. Varför skulle vi ge den detta? Och inte nog med det – denna AI måste ges en vilja och en egen agens som gör att den faktiskt vill ta över världen.²⁹

Även den svenske ordföranden för det Svenska AI-sällskapet, Fredrik Heintz, är kritisk till idén om att vi riskerar att tappa kontrollen över AI i och med den nu snabba utvecklingen. I stället ser han en risk att spetsutvecklingen är koncentrerad till ett fåtal amerikanska bolag och att spelplanen riskerar att bli ojämlig.³⁰

Det finns andra hot än mördarrobotar

Teknikskiften medför ofta nya stora möjligheter för mänskligheten. Hur vi sedan den industriella revolutionen förändrat sätten på vilka vi producerar, transporterar och kommunicerar är goda exempel på det. Men i kölvattnet av teknikutvecklingen kan det även visa sig att ny teknik medför nya utmaningar. Det handlar inte sällan om behov av lagstiftning kopplat till säkerhet. Med kameror kom behov av lagstiftning kring kameraövervakning och integritet. Med internet och den ökande hanteringen av personuppgifter av företag som vill kommunicera med dig uppstod behov av en skarpare personuppgiftslagstiftning med GDPR.

Med sociala medier fick vi nya sätt att hålla kontakten med vänner, bekanta och främlingar – men techjättarnas innehålls-

modererande AI blev snart experter på att styra innehållet så att vi maximerar tiden i dessa appar. Trots allt det goda som sociala medier medfört blev vi snart varse om teknikens baksidor. Sociala medieplattformen Instagram kunde i egna studier se att ökad användning av appen, framför allt bland unga kvinnor, också ledde till ökad psykisk ohälsa. En granskning gjord av Wall Street Journal, med läckta dokument från Facebook, visade att företaget haft kännedom om att Instagram är ”toxiskt för tonårsflickor”, men att bolaget inte agerat.³¹ Därutöver finns det gott om studier som visar att sociala medier och de bilder och videos som visas för oss är sammankopplat med depression, självkänsla och ätstörningar.³²

Kopplingen till AI för dessa problem kan kanske tyckas långsökt. Det är ju trots allt användare som lägger upp innehåll – men plattformarnas innehållsmodererande AI har, eller kan ha, en förödande roll att spela. Organisationen CCDH, Center for Countering Digital Hate, publicerade en studie av TikToks AI, och kunde konstatera att den som visat intresse för användare och inlägg kopplat till självskadebeteende och ätstörningar snart kom att ha ett flöde fyllt av just sådant innehåll. Det ligger i linje med affärsmodellen – att maximera spenderad tid i en app.³³

Ovan nämnda konsekvenser av AI är dock sådant som visat sig efter år av interaktion med innehållsmodererande AI. Vilka konsekvenserna av den nya vågen av generativ AI blir återstår att se.

Vad behöver politiken veta?

Ännu är AI-tekniken ung, och även om den kommer att påverka samhället i grunden är det svårt att idag sja om hur politiken kommer behöva agera för att samhället bättre ska kunna möta upp de möjligheter AI erbjuder, såväl som hantera de utmaningar som är och kommer bli verklighet. Det är därför svårt att ge skarpa policyrekommendationer vad gäller regulativa behov – vi vet ännu inte vilka utmaningar vi kommer ställas inför och omfattningen av dessa. Detta avsnitt fokuserar därför främst inte på politiska förslag, utan i stället på vad politiken måste vara medveten om för att kunna agera när behoven uppstår.

Sverige måste hitta sin nisch

När det kommer till spetsutveckling av AI är konkurrensen stor – och konkurrenterna i många fall globala techjättar. Det handlar om namnkunniga företag som OpenAI, Microsoft, Google, Apple och Meta, som investerat betydande summor i utvecklingen av AI. Google och Meta rapporteras vara de som investerat mest, med belopp runt 300 respektive 200 miljarder kronor.³⁴ OpenAI, som startades 2015, gjorde det med en satsning på motsvarande 10 miljarder kronor, därefter har Microsoft investerat ytterligare 100 miljarder kronor i bolaget.^{35, 36}

I Sverige lägger vi omfattande medel på AI-forskning, även om det inte går att jäm-

ställa med storleken på de amerikanska företagens investeringar. Innovationsmyndigheten Vinnova har satsat 100 miljoner kronor på AI Sweden, ett center för forskning och innovation kopplat till AI, bland annat för att påskynda tillämpning av AI och ökad kompetens. I Sverige har vi även sedan 2015 WASP, Wallenberg AI, Autonomous Systems and Software Program, som investerat närmare fem miljarder kronor i AI-forskning ihop med flera universitet och högskolor.³⁷ Även om det i Sverige tidigt satsats mycket pengar på AI är det svårt att mäta sig med de enorma resurser som investeras av amerikanska techbolag. När entreprenören Elon Musk intervjuades om sina planer på att skapa en konkurrent till OpenAI/Microsoft och Google menade han att den som på allvar vill konkurrera med techjättarna på AI-fronten behöver topptalanger, träningsdata och beräkningskapacitet – och inte minst det senare har blivit hårdvaluta. Musk säger i samma intervju att den beräkningskapacitet som krävs innebär hårdvara för minst 2,5 miljarder kronor och att OpenAI i arbetet med GPT-5 sannolikt använder hårdvara för mellan 12 och 20 miljarder kronor.³⁸ Det är därför svårt att se hur den svenska AI-utvecklingen ska kunna matcha den amerikanska endast utifrån kostnader för, och tillgången till, de avancerade grafik-kort som är bäst lämpade att använda för AI-träningen. Att tillgången till globala topptalanger och träningsdata också är trånga sektorer försvårar utmaningen än mer.

Vad Sverige kan bli ledande i är tillämpningen av AI för att öka produktivitet och stärka näringsliv och samhälle. Ett bra exempel på detta är hur Sverige med Capio S:t Görans Sjukhus blivit först i världen med att använda AI vid mammografiscreening.

Här har den statliga forskningspolitiken en roll att spela. Sveriges kunskaps- och forskningspolitik behöver en strategisk inriktning som inte utgår ifrån politiska mål om prestigeområden – vi behöver en forskningspolitik som tar sin utgångspunkt i Sveriges reella möjligheter att ta positioner i global framkant. I detta är AI en viktig del, men för att bli framgångsrik måste denna strategi låta möjligheter och behov vara formulerade av industrin i nära samarbete med forskningsinstitut och akademi. En ny svensk forskningsstrategi i nästa forskningsproposition behöver alltså allokera mer resurser åt att låta just industrin, instituten och akademien fokusera på de områden dessa tre aktörer bedömer att Sveriges framgångar kan bli störst. Då kan vi också ta ledarpositioner inom AI-användning.

Arbetsmarknaden kommer att förändras

AI kommer ha stor inverkan på arbetsmarknaden. Vissa jobb kommer på sikt sannolikt bli ersatta, samtidigt som AI kommer bli ett viktigt verktyg i andra yrken. Det handlar om uppgifter som att föra anteckningar i ett möte, skriva en

sammanfattning av en text eller korrekturläsning. Det handlar om automatiserat beslutsfattande och ett minskat behov av handläggare. Det handlar om urvalsprocesser där en AI snabbt kan sälla mellan ansökningar. Men likaväl handlar det om att kreativa uppgifter kan delegeras en AI. Ens manus till en muntlig presentation kan med AI-hjälp ges en fängslande dramaturgi och en diffusionsmodell kan på sekunder skapa mängder av designförslag för webbsidor, logotyper, illustrationer och marknadsföringskampanjer.

Ett fält där AI redan blivit ett viktigt verktyg för många yrkesverksamma är inom programmering. Programmeringsplattformen Github publicerade i september 2022 en studie där de utvärderat sin AI-programmeringsassistent Copilot, och denna visar att effektiviseringsmöjligheterna är enorma. I studien delades 95 programmerare slumpmässigt in i två grupper – en som fick använda Copilot och en som skulle lösa uppgiften själva. De som fick använda AI-assistenten löste uppgiften på 1 timme och 11 minuter i genomsnitt, medan de som skulle lösa uppgiften själva gjorde motsvarande arbete på 2 timmar och 41 minuter. Studien visade även att tre av fyra som använde Copilot kunde fokusera på mer tillfredsställande delar av sitt arbete, och fler än nio av tio uppgav att repetitiva uppgifter kunde hanteras snabbare.³⁹

Ur ett arbetsmarknadspolitiskt perspektiv är teknikskiften och arbetsmarknadens förändring ett ständigt faktum. Därför har vi redan idag system för omställning,

varför särskilda arbetsmarknadspolitiska åtgärder i alla fall i detta skede vore förhastade. Däremot är det viktigt att vara uppmärksam på att ny tillämpning av AI på kort tid skulle kunna förändra ett yrke i grunden och att omställningssystem snabbt behöver kunna dimensioneras för att hantera sådana utmaningar.

Samtidigt är det möjligt att många av de yrken i vilka AI kommer att bli viktiga assistenter också är yrken där det idag råder stor kompetensbrist. En av de mest akuta förändringarna på arbetsmarknaden skulle därför kunna vara att bristen på till exempel programmerare och utvecklare minskar avsevärt.

Kompetensfrågan i centrum

En viktig del i att bli världsbäst på tillämpning av AI i näringsliv och samhälle är att vi har rätt kompetens på forskarnivå. Det hänger tätt samman med den statliga forskningspolitiken och dess bidrag till att industri, akademi och institut kan skapa gemensamma forskningsarenor där ny mark bryts.

Likaväl är det viktigt för den breda tillämpningen av AI att fler i samhället får grundläggande kunskap om AI. Här kan den finska utbildningen Elements of AI, som är en öppen och gratis AI-utbildning, tjäna som inspirationskälla – särskilt som denna nu finns på svenska. I Finland har mer än 2 procent av befolkningen, över 110 000 personer, gått utbildningen. Motsvarande

insatser skulle även kunna göras i Sverige för att öka kunskapsnivån i befolkningen, och på sikt skapa bättre förutsättningar för implementering av AI-tjänster brett i samhället – en sorts Hem-PC-reform för den artificiella intelligensens teknikskifte.

Vilket innehåll kan man lita på?

En av de på kort sikt måhända största konsekvenserna av att AI blir bättre på att skriva som människor, imitera våra röster och skapa visuella kopior av oss är att det kommer bli svårare att veta vad du kan lita på i digitala kanaler.

Sam Altman, OpenAI:s vd, menar att AI redan med dagens kapacitet kan skapa stor skada genom att användas för att sprida propaganda i sociala medier.⁴⁰ En AI kan ges i uppdrag att skapa och sprida trovärdig desinformation i syfte att gynna eller skada människor, företag eller politiska processer. Samma AI skulle kunna generera falska bilder som stöd för sin falska information. AI skulle också kunna utnyttja sociala mediernas egna innehållsmodererande AI för att se till att en viss desinformation lyfts upp ur bruset och visas för fler mottagare.

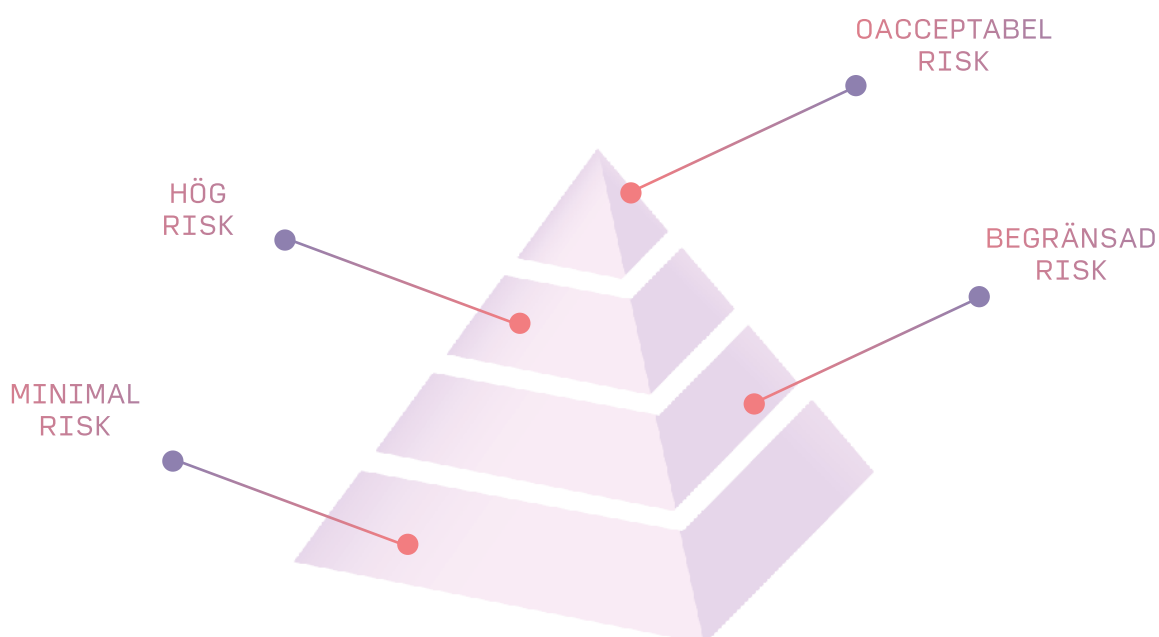
Trots att tillämpningen av AI ännu får anses vara i sin absoluta linda finns redan exempel på detta. Den 22 maj 2023 spreds en bild i sociala medier som visade en explosion utanför det amerikanska försvarsdepartementets högkvarter Pentagon.

Bilden skapade omedelbart oro på den amerikanska börsen, innan uppgifterna kunde dementeras och bilden konstateras vara AI-genererad.⁴¹ Det finns även exempel på bedrägerier där kriminella använt röstimiterande AI för att över telefon lura sig till pengar genom att låtsas vara en familjemedlem i nöd.⁴²

I mars 2023 lyfte Europol, EU:s organ för samarbete kring brottsbekämpning, just upp bedrägeri och desinformation som två viktiga områden för rättsväsenden att nära följa utvecklingen för att förutse och förebygga oönskade effekter av den snabba AI-utvecklingen.⁴³ En konsekvens av generativ AI och dessa snabba framsteg skulle kunna vara att alltmer av den digitala media vi konsumerar sker genom olika filter. Det kan handla om ett uppsving för redaktionell media eller genom olika sätt att identifiera antingen AI-genererat innehåll, eller genom att med BankID eller annan identifikation bekräfta en mänsklig avsändare.

Det här är AI-akten

Behöver svenska lagstiftare oroa sig för AI-undergången? Om den kommer så ja, men inte ur ett lagstiftningsperspektiv – de juridiska bitarna hanteras av EU genom AI-akten. Denna förväntas klubbas under slutet av 2023 och tillämpningen börjar gälla leverantörer av AI-tjänster två eller tre år efter detta. För att inte dra all AI över en domedagskam har EU valt att dela upp AI i fyra riskgrupper: *oacceptabel risk*, *hög risk*, *begränsad risk* och *minimal risk*. När lagstiftningen träder i kraft kommer det att innebära stora förändringar i hur AI används och regleras, om de ska tillåtas verka inom EU. Alla AI-system som lanseras i EU måste bedömas och klassificeras utifrån dessa kategorier, oavsett hur harmlösa de kan verka.⁴³



OACCEPTABEL RISK

AI som anses utgöra ett hot mot människor bedöms utgöra en oacceptabel risk och blir därmed förbjudna inom EU. Denna kategori är till stor del skapad för att förbjuda AI inom övervakning, identifiering och rangordning av individer i realtid. Som exempel kan nämnas kameraövervakning på offentliga platser där en AI kan identifiera dig med hjälp av ansiktsigenkänning, eller en arbetsgivare som använder AI för att göra en bedömning av en arbetsökandes person genom att låta AI tillskriva individen en ”social poängsättning” utifrån dennes interaktioner på sociala medier eller socioekonomiska status. Kort sagt blir både social poängsättning och biometrisk identifiering i realtid och på distans förbjudet. Vissa undantag kan dock ske vid grov brottslighet så länge identifiering inte sker i realtid.

Till oacceptabel risk räknas även AI som innebär kognitiv beteendemanipulation av personer eller sårbara grupper. Det kan handla om leksaker med AI som kan få barn att agera på ett sådant sätt att de utsätter sig för faror, eller särskilt sårbara grupper som utnyttjas med hjälp av AI.

AI-system i högrisk-kategorin, nivån under oacceptabel risk, innefattar alla system som skulle kunna innebära en negativ påverkan på säkerhet eller människors grundläggande rättigheter, och kommer att innefatta produkter och tjänster som tillsammans har flera miljarder användare dagligen. Det handlar om produkter inom ramen för EU:s produktsäkerhetslagstiftning – som AI i flygplan, bilar, hissar eller leksaker – men det handlar också om innehållsmodererande AI som styr vad vi ser i våra sociala kanaler.⁴⁴

HÖG
RISK

- 1) Biometrisk identifiering av personer – t.ex. ansiktigenkänning.
- 2) Hantering av kritisk infrastruktur – t.ex. styrning av elsystem eller trafikljus.
- 3) Utbildning och yrkesutbildning – t.ex. rättning av prov.
- 4) Sysselsättning, förvaltning av arbetstagare – t.ex. hantering av CV i rekrytering.
- 5) Tillträde till varor eller tjänster – t.ex. kontroll av kreditvärdighet.
- 6) Brottsbekämpning – t.ex. för analys av bevis.
- 7) Migration, asyl och gränskontroll – t.ex. kontroll av resedokument.
- 8) Rättsprocesser och demokratiska processer – t.ex. AI för att tolka lagen.
- 9) Innehållsmodererande AI på sociala medie-plattformar – t.ex. AI som avgör vilket innehåll du får se i dina sociala kanaler (adderades av parlamentet 23 maj).

AI inom någon av de nio kategorierna ovan kommer genomgå en gedigen myndighetsprövning, i vilken bland annat riskbedömning, robusthet och säkerhet testas. Bland kraven ingår även att data sparas så att resultat av AI-användningen kan följas upp. Dessa system måste också ha en löpande mänsklig övervakning för att minimera risken med användningen. Dessa system kommer sedan finnas registrerade i en högriskdatabas.⁴⁵

BEGRÄNSAD
RISK

System med begränsad risk innefattar AI till exempel chatbottar, som ChatGPT. Här finns det endast ett transparenskrav – det måste framgå att man chattar med en AI.

I kategorin minimal risk finns sådana AI som filtrerar skräp i din mejlkorg. Majoriteten av AI-system som för närvarande finns i EU placeras idag i denna kategori.

MINIMAL
RISK

Vad händer då om en AI-leverantör trots den rigorösa kontrollfunktionen ändå lyckas orsaka personskada? Det blir en fråga för *direktivet om skadeståndsansvar*, som hanteras först när AI-akten har klubbats.

Kommer AI-akten sinka EU:s utveckling?

Europa ligger efter när det kommer till utveckling av ledande AI-system, och det finns flera faktorer – inte minst AI-akten, som riskerar att placera oss i en än sämre situation. Redan idag dominerar AI-sektorn av USA och Kina. Under 2022 stod USA för mer än 65 procent av alla privata investeringar i AI, medan Kina och EU stod för 18 respektive 15 procent.⁴⁶

Sett till vetenskapliga artiklar ligger Kina långt före alla konkurrentländer, med drygt 40 procent av publicerade vetenskapliga artiklar 2021. Det kan jämföras med EU+UK på drygt 15 procent och USA på 10 procent. Sett till antal vetenskapliga artiklar mellan 2010 och 2021 toppas listan av nio kinesiska universitet, följt av amerikanska Massachusetts Institute of Technology, på tionde plats. Med tanke på skillnader i styrelseskick blir såväl privata investeringar som forskningsartiklar något vanskliga mått för en precis jämförelse mellan USA och Kina – sammantaget visar dock datan att just dessa länder ligger i AI-utvecklingens framkant, och att EU i många fall ligger långt efter.⁴⁷

Lägg till detta AI-akten, genom vilken AI-system måste klassificeras och i många fall genomgå rigorösa tester – med risk för stora skadeståndskrav när de väl släpps på marknaden. Det är förstås svårt att på förhand bedöma riktigheten och nödvändigheten i de hårt ställda krav AI-akten innebär för en uppsjö AI-tjänster, särskilt som det finns forskare och experter som menar att AI kommer innebära en ofrånkomlig undergång, likaväl som det finns forskare och experter som menar att hoten är kraftigt överdrivna. Något som däremot kan sägas med större säkerhet är att EU inte har en ledande position inom AI, varken vad gäller forskning eller den privata sektorns investeringar. Med AI-akten och vad den medför i form av pålagor för den som vill utveckla en AI-tjänst innebär det också en uppenbar risk att det företag som vill utveckla AI-tjänster också placerar denna utveckling utanför EU.

Global utblick

Det är inte bara inom EU som politiken försöker skapa regelverk kopplat till AI, tvärt om sker det just nu världen över. Genomgående är att det handlar om att försöka dra nytta av AI, samtidigt som risker ska minimeras. Därutöver pågår arbete för att redan nu få företag att ingå frivilliga överenskommelser innan ny lagstiftning varit på plats.

- *USA: Frivilliga åtaganden innan lagstiftning kommit på plats*
Biden-Harris-administrationen har under sommaren och sensommaren 2023 fått de största AI-utvecklarna i USA att gå med på en rad frivilliga åtaganden för att försäkra sig om att AI-tjänster ska vara säkra för slutanvändare och samhälle. Företagen åtar sig också att investera i cybersäkerhet och interna säkerhetsrutiner och väntas också leva upp till höga transparenskrav för att vinna allmänhetens förtroende. I ett uttalande skriver Vita huset att detta frivilliga åtagande är en viktig brygga för att tidigt reglera AI innan lagstiftning kan vara på plats.⁴⁸
- *AI Pact*
Lagstiftning tar tid – det gäller inte minst AI-akten. Därför har företrädare för EU redan nu försökt få till stånd en AI-pakt, ett frivilligt regelverk där företag som Google och OpenAI accepterat principer om ansvarsutkrävande och transparens.⁴⁹
- *G7 Hiroshima AI Process*
I maj 2023 träffades ledarna för G7-länderna i Hiroshima för ett första möte om ländernas förhållningssätt till AI och framförallt generativ AI. Syftet är att koordinera regelverk och insatser kopplat till AI, inte minst för att se till att generativ AI också lever upp till G7-ländernas förväntningar om AI och demokratiska principer. Innan 2023 är slut ska G7 mötas igen för att ta fram ett policyramverk.⁵⁰
- *Kinesiska statens riktlinjer för generativ AI*
Den kinesiska myndigheten Cyberspace Administration of China har under första halvan av 2023 tagit fram riktlinjer för utvecklare av generativ AI, vilken innebär att tjänsterna måste genomgå säkerhetstestning av myndigheter innan de kan släppas för allmänheten. Granskningen innebär bland annat att tjänsterna måste ligga i linje med kommunistpartiets värderingar. Utvecklare av AI-tjänster blir med dessa riktlinjer också ansvariga för legitimiteten i träningsdatan (utifrån kommunistpartiets värderingar får man gissa). Därutöver måste utvecklare registrera samtliga användares ”identiteter och relaterad information”. De utvecklare som inte efterlever regelverket riskerar att bli föremål för brottsutredning.⁵¹

Det kinesiska regelverket skulle i teorin kunna sätta hämsko på den kinesiska utvecklingen av generativ AI och stora språkmodeller. Träning av stora språkmodeller kräver nämligen så pass stora mängder data att det är svårt att kontrollera allt innehåll i datamängden (datamängden Common Crawl som använts bland annat av ChatGPT innehåller information från 240 miljarder webbsidor). Om utvecklare riskerar kännbara straff om en AI uttrycker värderingar som inte ligger i linje med kommunistpartiets värderingar kan resultatet dels avskräcka från att utveckla sådana tjänster, dels kräva att enorma resurser för att kontrollera träningsdatan.

Ur ett konkurrenskraftsperspektiv är det förstås att föredra om så många länder som möjligt kan anamma liknande regelverk. Ur det perspektivet kan det vara positivt att EU:s AI-akt både är omfattande och har kommit långt i processen. Om AI-akten kan agera inspiration för andra länders och regioners kommande regelverk har vi måhända en fördel av att tidigt ha lärt oss regelverket.

Del 4: Industrins tillämpning av AI

Så kommer AI användas i industrin

Många frågar sig hur industrin kommer att påverkas av AI – eller hur tillämpningen av AI kommer att se ut inom industrin. Det är lätt att gå händelserna i förväg och tro att AI över en natt kommer att förändra det mesta. Mer rimligt är nog att likna tillämpningen av AI i industrin med internets intåg – över en natt hände inte så mycket, men på 20 år har arbetssätt, digitala verktyg, försäljningskanaler, affärsmodeller samt varu- och tjänst-utbud förändrats i grunden. I detta kommer vissa företag gå före medan andra halkar efter. För företag inom tillverkningsindustrin eller tekniktjänster skulle utvecklingen kunna se ut enligt följande:

- 1) AI för effektivisering av kontorsarbete och fysiska miljöer
- 2) Automatisering av uppgifter där datan primärt består av text eller bilder
- 3) Utvecklat tjänsteinnehåll där AI är en integrerad del av varan eller tjänsten
- 4) AI som står bakom forskningsupptäckter och teknologiska framsteg

1) I ett tidigt skede kommer AI kunna bidra till effektivisering. Som tidigare nämnt i rapporten handlar det om digitala verktyg som gör dig lite mer effektiv i kontorsarbetet, oavsett om det handlar om mötesanteckningar, dataanalys, programmering eller research. Men det handlar också om hur AI kan styra och effektivisera fysiska platser. I industriella miljöer är termen *digitala tvillingar* väl etablerad, och beskriver hur en digital kopia av en fabrik kan simulera förbättringar och effektiviseringar i produktionsflöden. Likväl finns det redan idag företag som erbjuder AI-styrning för ventilation och uppvärmning av fastigheter för att dels ge ett bättre inomhusklimat, dels sänka energiför-

brukningen. Förutom de AI-tjänster som kommer ingå i våra vanliga digitala arbetsverktyg så kommer industrin och näringslivet brett erbjudas en rad nischade AI-tjänster som innebär personella eller resursmässiga besparingar.

2) Därefter kommer AI sannolikt leta sig längre in i den administrativa delen av företagen. Granskning av stora mängder text i rekryteringssyfte, granskning av avtal eller för att bistå med översättning kan vara områden där AI-verktyg kommer kunna avlasta eller ersätta medarbetare. Stödfunktioner så som kundtjänst kommer även det i högre grad att hanteras av AI. Även kvalitetsgranskning som skett genom okulär besiktning kan ersättas av en kamera och en AI.

3) Lite längre fram kommer AI bli en del av tjänsteutbudet kopplat till tillverkningsindustrins produkter. Det är svårt att se en framtid där en bil eller en industrirobot inte styrs med AI-hjälpmiddel. Det finns förvisso redan självkörande bilar, som dessutom är förvånansvärt bra – och det sätter fingret på något som komplicerar utvecklingen av AI-tjänster på denna nivå: nämligen att olika lagstiftning i olika regioner gör det olika attraktivt att investera i forskning och utveckling av AI beroende på var ett företag har sitt säte.

Generativ AI kan även komma att bli regel snarare än undantag i datorspel, där så kallade NPC:er (Non-Player Characters) fram till idag huvudsakligen varit förprogrammerade att kunna säga endast en eller ett par meningar. Snart kanske vi ser spel där NPC:er kan styras av generativ AI, men givits en egen persona samt egna agendor och intressen. Det skulle i teorin kunna öppna upp för oändliga variationer i spel som tidigare varit låsta till spelutvecklarnas möjlighet att lägga resurser på de NPC:er som populera dagens datorspel.

4) Med ytterligare framsteg inom AI – och när fler företag med stora forskningsresurser inom olika fält blir bättre på att använda AI kan vi kanske också räkna med fler stora forskningsgenombrott. Precis som AI används för att hitta effektiv antibiotika mot multiresistenta bakterier så går det att tänka sig att den AI som tränas på stora mängder data inom specifika forskningsfält, liksom materialforskningsroboten vid LBNL, kan identifiera sådant människan inte haft resurser eller förmåga att hitta. Redan idag är AI-drivna robotlaboratorier i full gång både vad gäller just materialforskning och medicinsk forskning. Det är förmodligen inte en långsökt gissning att alla företag med stora forsknings- och utvecklingsbudgetar också kommer att använda AI i innovations- och effektiviseringssyfte.

Så många använder AI

Även om det senaste årets utveckling kopplat till generativ AI inneburit en upphaussning av AI generellt i samhället så är AI inte ett nytt fenomen. Statistiska Centralbyrån har vid två tillfällen mätt hur vanligt det är att företag och offentlig sektor använder AI, och resultaten kanske kommer som en överraskning. Bland svenska företag brett använde 6,5 procent av företagen AI, medan mer än 30 procent av stora företag (>250 anställda) använde AI i verksamheten. Syftet med användningen av AI var primärt att förbättra en existerande produkt eller tjänst, men nästan lika vanligt var användning av AI för att antingen utöka kunskapen om relationer med kunder eller använder, alternativt att utveckla en ny produkt eller tjänst.⁵²

I Teknikföretagens medlemsundersökning om digital mognad i teknikindustrin uppgav strax under 8 procent att de använder sig av AI i verksamheten – precis som i SCB:s undersökning fanns dock en stora skillnader mellan stora respektive små företag.⁵³

Även om utvecklingen av AI-tjänster såväl som AI-forskning till stor del sker i stora amerikanska techbolag eller på kinesiska universitet så sticker Sverige ut i en parameter i Stanfords AI-index från 2023: 1,20 procent av jobbannonser i Sverige uttrycker önskemål om kunskaper inom AI. Det kan låta lite, men i själva verket är det femte högst i världen.⁵⁴ Någon ytterligare analys av vad för AI-kunskaper som efterfrågas i Sveriges presenteras inte i denna data, däremot skulle det kunna indikera att det svenska näringslivet ser AI som en viktig del i framtidens produkt- och tjänsteerbjudande och att många svenska företag vill dra nytta av att vara tidiga i detta teknikskifte.

Slutnoter

- 1 arXiv:1706.03762 [cs.CL] <<https://arxiv.org/abs/1706.03762>>
- 2 Radford et al. Improving Language Understanding by Generative Pre-Training <https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf>
- 3 <<https://github.com/openai/gpt-3>> <<https://medium.com/@mlubbad/the-ultimate-guide-to-gpt-4-parameters-everything-you-need-to-know-about-nlps-game-changer-109b8767855a#4cf9>>
- 4 AI-forskaren Lex Fridman vid MIT intervjuar OpenAI:s vd Sam Altman
- 5 <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>
- 6 The Future of Work With AI - Microsoft March 2023 Event <<https://www.youtube.com/watch?v=Bf-ds9CcRU>>
- 7 Klas Wählberg på LinkedIn <<https://www.linkedin.com/pulse/verkligheten-eller-kartan-klas-wahlberg/>>
- 8 Databasen LAION-5B <<https://laion.ai/blog/laion-5b/>>
- 9 Diffusion Models: A Practical Guide <<https://scale.com/guides/diffusion-models-guide>>
- 10 Dagen Nyheter 2023-04-20 <<https://www.dn.se/kultur/ai-lat-av-drake-och-weekend-tas-bort/>>
- 11 VALL-E Voice Synthesizer <<https://valle-demo.github.io/>>
- 12 Forbes 2021-10-03 <<https://www.forbes.com/sites/robtoews/2021/10/03/alpha-fold-is-the-most-important-achievement-in-ai-ever/>>
- 13 Forskning och Framsteg 2022-10-30 <<https://fof.se/artikel/2022/9/ai-vecklar-upp-proteinernas-hemligheter/>>
- 14 Capiro S:t Göran 2023-05-16 <<https://capiostgoran.se/nyheter/2023/ai-ska-hjaltill-att-hitta-brostdcancer/>>
- 15 <<https://www.who.int/news-room/fact-sheets/detail/cancer>>, <https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf>
- 16 MIT News 2023-05-25 <<https://news.mit.edu/2023/using-ai-scientists-combat-drug-resistant-infections-0525>>
- 17 The Guardian 2023-05-25 <<https://www.theguardian.com/technology/2023/may/25/artificial-intelligence-antibiotic-deadly-superbug-hospital>>
- 18 Science.org, 2023-04-20 <<https://www.science.org/content/article/ai-driven-robots-start-hunting-novel-materials-without-help-humans>>
- 19 Science.org, 2023-03-07 <<https://www.science.org/content/article/ai-re-creates-what-people-see-reading-their-brain-scans>>
- 20 Takagi, Nishimoto, 2023 <<https://sites.google.com/view/stablediffusi>>

- on-with-brain/>
- 21 Scotti, Banerjee, Goode et. al., 2023 <<https://medarc-ai.github.io/mindeye/>>
Nature Neuroscience, 2023-03-15 <<https://www.nature.com/articles/s41593-023-01304-9.epdf>>
- 22 Future of Life Institute, 2023-03-22 <<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>>
- 23 AI Impacts 2022 <<https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/>>
– undersökningen har rönt viss kritik från de som hävdar att faror med AI är kraftigt överdrivna, men vi gör bedömningen att den ändå är värd att nämna.
- 24 Dagens Nyheter, 2023-05-01 <<https://www.dn.se/varlden/ais-gudfader-varnar-for-mordarrobotar/>>
- 25 Bankless podcast <<https://www.youtube.com/watch?v=GyFkWb903aU>>
- 26 <https://openai.com/blog/planning-for-agi-and-beyond>
- 27 Ibid.
- 28 Yann LeCun på X/Twitter <<https://twitter.com/ylecun/status/1651449929842274309>>
- 29 20VC Podcast med Yann LeCun E1014, 15 maj 2023
- 30 Dagens Nyheter, 2023-05-11 <<https://www.dn.se/ekonomi/har-tar-sverige-uppkampen-mot-ai-jattarna/>>
- 31 Wall Street Journal, 2021-09-14 <<https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>>
- 32 American Psychological Association: How can we minimize Instagram's harmful effects? 2 december 2021. <<https://www.apa.org/monitor/2022/03/feature-minimize-instagram-effects>>
- 33 AP News, 2022-12-15 <<https://apnews.com/article/technology-health-eating-disorders-center-government-and-politics-0c8ae73f44926fa3daf66bd7caf3ad43>>
- 34 Exo Insight, 2023-03-22 <<https://insight.openexo.com/how-much-is-invested-in-artificial-intelligence/>>
- 35 CNN Business, 2015-12-12 <<https://money.cnn.com/2015/12/12/technology/openai-elon-musk/>>
- 36 Bloomberg, 2023-01-23 <<https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai#xj4>>
- 37 WASP Sweden <<https://wasp-sweden.org/about-us/>>
- 38 Under Wall Street Journal CEO Council sade Musk att det krävs hårdvara för åt minstone 250 miljoner dollar, samt att OpenAI använder sig av uppskattningsvis 30 000 till 50 000 Nvidia H100-kort, som kostar uppskattningsvis 350 000 till 450 000 kronor styck.
- 39 Github Blog <<https://github.blog/2022-09-07-research-quantifying-github-copi>>

- lots-im pact-on-developer-productivity-and-happiness/>
- 40 AI-forskaren Lex Fridman vid MIT intervjuar OpenAI:s vd Sam Altman
- 41 AP News, 2023-05-23 <<https://apnews.com/article/pentagon-explosion-misinformation-stock-market-ai-96f534c790872fde67012ee81b5ed6a4>>
- 42 Washington Post, 2023-03-05 <<https://www.washingtonpost.com/technology/2023/03/05/ai-voice-scam/>>
- 43 Europol <<https://www.europol.europa.eu/media-press/newsroom/news/criminal-use-of-chatgpt-cautionary-tale-about-large-language-models>>
- 43 EU AI-akten <<https://data.consilium.europa.eu/doc/document/ST-8115-2021-INIT/sv/pdf>>
- 44 Parlamentets tillägg till AI-akten, 2023-05-23 <<https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>>
- 45 EU-Kommissionen <<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>>
- 46 Stanford AI Index 2023 <<https://aiindex.stanford.edu/report/>> s. 191
- 47 Ibid. s. 44
- 48 White House Fact Sheet, 2023-09-12 <<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/>>
- 49 Politico, 2023-05-26 <<https://www.politico.eu/article/big-tech-rumble-europe-global-artificial-intelligence-debate-ai-pact/>>
- 50 Reuters, 2023-05-20 <<https://www.reuters.com/world/g7-calls-developing-global-technical-standards-ai-2023-05-20/>>
- 51 Reuters, 2023-04-11 <<https://www.reuters.com/technology/china-releases-draft-measures-managing-generative-artificial-intelligence-2023-04-11/>>
- 52 Statistiska Centralbyrån: AI-användning i företag och offentlig sektor <https://scb.se/contentassets/ea0e9cccd58343e7a07fe4c055f8fad2/nv0116_2022a01_br_nvftbr2301.pdf>
- 53 Teknikföretagen: Digital mognad i tillverkningsindustrin, 2022 <https://www.teknikforetagen.se/globalassets/rapporter--publikationer/digitalisering/digital-mognad_2022.pdf>
- 54 Stanford AI Index 2023 <<https://aiindex.stanford.edu/report/>> s. 173

Teknikföretagen