



EN DEL AV SKRIFTSERIEN BEMA — OM LIBERALISMENS UTMANINGAR

#5

Individuell frihet i en optimerad beslutsmiljö

Vad händer med det öppna samtalet
när vi förlitar oss på digitala tvillingar?

REBECCA C. ÅNGSTRÖM

OHLIN institutet

bēma

En skriftserie om liberalismens utmaningar.

I antikens Athen var bēma den plattform man stod på när man presenterade sina argument inför folkförsamlingen. Ohlininstitutet erbjuder med skriftserien **bēma** en plattform för liberalt sinnade människor att framföra sina idéer. Skrifterna speglar således skribentens åsikter.



Bēma (grekiska: βῆμα) var i antikens Grekland en upphöjd plattform som användes för att tala från inför en församling eller domstol. Ordet betyder bokstavligen "steg" eller "plattform" och förekom i flera olika sammanhang, men gemensamt för dem alla var att det handlade om offentlig auktoritet, synlighet och ansvar.

I folkförsamlingen – exempelvis på Pnyx i Aten – stod talare på bēman när de skulle tala till folkförsamlingen. Det var en konkret plats för demokratisk deliberation – alltså diskussion och argumentation inför ett beslut. Den symboliserade medborgarens rätt att tala och delta i det gemensamma.

Den som stod på bēman trädde fram inför andra och måste stå för sitt ord. Retorisk skicklighet vägde tungt, men så gjorde även karaktär och ethos. Att tala från bēman var inte bara ett privilegium, utan ett uttryck för medborgarskap och förtroende. Den konkretiserar idén att beslut, sanning och makt ska formas inför andra, inte i det fördolda.

Individuell frihet i en optimerad beslutsmiljö

Vad händer med det öppna samtalet när vi förlitar oss på digitala tvillingar?

”Det är inte förutsägelsena som är det spännande, det riktigt magiska händer när AI kan avgöra den bästa vägen framåt!” Min vän datavetaren lutade sig fram över bordet och hennes ögon glittrade när hon berättade om sina förväntningar på framtidens AI-modeller. Året var 2019 och jag arbetade med ett forskningsuppdrag där jag undersökte hur AI implementeras och används i organisationer. Vi satt i ett litet mötesrum på hennes arbetsplats, jag med anteckningsblocket framför mig redo att höra hennes tankar om AI. Det var innan ChatGPT:s storstilade entré, men AI var redan ett väl etablerat samtalsämne. Böcker, poddar, och artiklar om framförallt maskininlärning publicerades i en strid ström. I fokus stod idén att AI, genom att analysera enorma datamängder, skulle kunna förutse en mängd olika händelser, från vardagliga skeenden till stora samhällsförändringar. Sådana förutsägelser skulle i förlängningen kunna omforma vår värld.

Hur en sådan förändring skulle kunna se ut var också en central del av den samtida diskussionen. Jag minns särskilt två exempel från den tiden. Det första presenterades i boken 'Prediction machines' (Agrawal, Gans, & Goldfarb, 2018) som publicerades samma år. Författarna beskrev hur mer precisa förutsägelser skulle kunna effektivisera både processer och fysiska miljöer i samhället. Som illustrativt exempel användes flygplatser, en plats som i hög grad formats av osäkerheter i tidsplanering. Många av oss lägger gärna till en extra halvtimme för att undvika stress om det skulle vara mycket trafik på vägen till flygplatsen eller om det skulle ta längre tid än väntat att ta sig genom säkerhetskontroller och terminaler. Resultatet blir ofta detsamma, att vi tillbringar en hel del tid med att vänta på flygplatsen. Det är i sin tur, enligt författarna, orsaken till att flygplatslounger, butiker och caféer vuxit fram för att både erbjuda resenärerna ett tidsfördriv och skapa affärsmöjligheter. Med bättre prediktioner av trafikflöden, både på vägen till flygplatsen och inom terminalerna, skulle mycket av denna osäkerhet kunna elimineras. Vi skulle kunna anlända precis i tid till vår avgång, och loungen och caféet skulle bli mer eller mindre överflödiga.

Det andra exemplet på hur AI skulle kunna påverka samhället handlade om hur prediktiva modeller skulle kunna förändra organiseringen av kritiska verksamheter. Det kanske mest iögonfallande exemplet handlade om prediktiv polisverksamhet (predictive policing), en typ av AI-lösningar som redan då tillhandahölls av företag som PredPol (senare Geolitica). Prediktiv polisverksamhet använder AI för att förutse var nästa brott sannolikt kommer inträffa, genom att kombinera data om bland

I fokus stod idén att AI, genom att analysera enorma datamängder, skulle kunna förutse en mängd olika händelser, från vardagliga skeenden till stora samhällsförändringar.

annat brottsstatistik, aktuella sociala evenemang som sportmatcher, väderförhållanden och närheten till barer. Att förutse brott möjliggör brottsförebyggande insatser, exempelvis genom att skicka patruller till områden där AI bedömer risken för brott som störst. Men detta innebär också att organisationen behöver utveckla processer som förflyttar fokus från reaktiv brottshantering till ett mer systematiskt proaktivt arbete (Shapiro, 2017; Waardenburg, Sergeeva, & Huysman, 2018). Prediktiv Polisverksamhet var och är alltjämt omdebatterat, bland annat då användandet av AI visat sig förstärka diskriminering och att uttolkning av AI-resultat i polisverksamhet inkluderat högst subjektiva bedömningar (Brayne & Christin, 2021; Lum & Isaac, 2016; O'Neil, 2016; Shapiro, 2017; Waardenburg et al., 2018)

Vi har inte kommit hela vägen till den framtid som förutspåddes, att implementera AI och skapa värde är ett betydligt långsammare arbete än vad många kanske tror. Men i många fall arbetar vi fortfarande mot samma mål: att skapa nytta genom att låta AI förutse och planera framtiden. De exempel som nämnts hittills låter kanske inte så omvälvande idag, men de är del av en större helhetstanke. Om, eller när, varje typ av verksamhet kan dra nytta av förutsägelser och fokusera sina resurser på proaktivt arbete skulle vår vardag förändras i grunden. Det var den visionen som fick min väns ögon att glittra när vi sågs 2019, och hon tog tanken ett steg längre. Hon beskrev AI som inte enbart förutser händelser, utan även överblickar ett stort antal möjliga scenarier för att bedöma deras sannolika utfall och därefter rekommendera den optimala vägen framåt. Eller med andra ord: när tekniken inte bara förutspår framtiden utan är delaktig i att forma den, beslut för beslut. Det hon beskrev går bortom prediktiv AI och kallades för 'preskriptiv (prescriptive) AI'.

'Preskriptiv AI', denna rådgivande form av AI, blev aldrig ett buzzword på samma nivå som generativ AI eller agentbaserad AI är idag. Men drömmen om att AI ska kunna utvärdera våra valmöjligheter och ge konkreta råd om hur vi bör handla lever i allra högsta grad vidare. Den längtan märks inte minst i hur AI tjänster positioneras som rådgivare, inte bara som stöd för att söka information, utan också för att fatta beslut. För att en AI ska kunna utvärdera olika möjliga utfall och ge sådana råd behöver den dock kunna simulera framtiden. Det är här den digitala tvillingen kommer in.

Den digitala tvillingen för optimerade beslut

Ett sätt att inom forskning och utveckling pröva olika scenarier för att identifiera den bästa vägen framåt är att använda datorsimuleringar, där alternativa händelseförlopp testas och analyseras. Sådana simuleringar kan byggas med hjälp av så kallade digitala tvillingar. En digital tvilling är en virtuell avbild av något fysiskt, det kan vara ett verktyg, ett system, en produktionsanläggning, eller en hel stad (Göteborgs Stad, 2026; IBM, 2020; Wikipedia). Genom att så långt som möjligt inkludera alla relevanta fysiska komponenter i den digitala modellen, till exempel med hjälp av sensordata, kan modellen spegla den fysiska verkligheten i realtid. Det gör det möjligt att testa, analysera och optimera olika scenarier i den digitala miljön, innan ändringar görs i den verkliga världen. Exempelvis, så används digitala tvillingar inom tillverkningsindustrin genom att modellera produktionsprocesser i virtuell miljö. Då kan verksamheten testa olika förändringar de vill göra i tvillingen, som justeringar i driften eller systemuppgraderingar, för att se hur ändringarna

Vi har inte kommit hela vägen till den framtid som förutspåddes, att implementera AI och skapa värde är ett *betydligt långsammare* arbete än vad många kanske tror.

kan komma att påverka produktionen i verkligheten. På så sätt kan verksamheten optimera sina beslut och minska osäkerheten.

En digital tvilling är också ett tekniskt verktyg som möjliggör den så kallade 'preskriptiva AI'n'. Med hjälp av den digitala tvillingen kan AI utföra tusentals simuleringar, analysera samtliga utfall för att identifiera vilket handlingsalternativ som med största sannolikhet leder till det bästa eller önskade resultatet, och ge råd om vägen framåt.

Det är viktigt att poängtera att digitala tvillingar, likt andra digitala verktyg, inte är ofelbara. Eftersom de bygger på data, modellering och olika antaganden är de känsliga för brister som låg datakvalitet, felaktiga eller förenklade modeller, eller bristande kalibrering med verkligheten. Om en sensor som mäter temperatur skickar felaktiga värden till en modell av en tillverkningsprocess fungerar fortfarande den digitala tvillingen, men slutsatserna blir missvisande eftersom de inte återger de verkliga förhållandena för processen.

Ofta talar man om digitala tvillingar av objekt eller immateriella ting, som tillverkningsprocesser eller trafikflöden. Men det finns även digitala tvillingar som skapas för att avbilda människor. Inom medicin utvecklas exempelvis digitala tvillingar av mänskliga kroppar (Söderlund Leifler, 2022). På samma sätt som en digital tvilling av en tillverkningsprocess byggs upp kring de faktorer som är relevanta för produktionen, modelleras en digital tvilling av en människokropp utifrån kroppens organ och biologiska processer. Tvillingen förses med data som är specifik för den enskilda individen, och kan sedan användas för att förutse hur en person kommer att svara på olika läkemedel eller behandlingar, samt användas som verktyg i läkemedelsutveckling (Söderlund Leifler, 2022).

Om en digital tvilling av en individ kan ge insikter om hur just den personen reagerar på läkemedel, kan man då inte tänka sig en digital tvilling för att testa hur en person reagerar på olika livsbeslut eller omständigheter i vardagen? Redan idag finns konsumenter som ser potentialen i att använda personliga digitala tvillingar för att simulera flera av livets svåra beslut. I en studie där tidiga användare av digital teknologi tillfrågades om hur de kan tänkas använda AI i framtiden uttryckte många ett stort intresse för att ha en egen digital tvilling, exempelvis för att "provköra" ett äktenskap och på så sätt bedöma relationens hållbarhet (Ericsson, 2024). En framtid som inte känns helt främmande med tanke på att en betydande andel av människor använder AI idag för att söka praktisk vägledning i vardagliga beslut (Chatterji, Cunningham, Deming, Hitzig, Ong, Yan Shan, & Wadman, 2025)

Personliga digitala tvillingar för att simulera individers egna liv finns än så länge bara i teorin, mig veterligen. Och vid en första anblick kan det låta lockande att optimera sin vardag. Vilken kost och träning passar mig bäst? Vilken utbildning bör jag välja? Vilka typer av jobb skulle jag trivas bäst på? Att i framtiden enkelt kunna besvara dessa frågor genom att simulera beslut kan upplevas ha stort värde, men väcker samtidigt frågor om vilken funktion beslutsfattande och eget omdöme har för individens personliga utveckling. Vad riskerar vi att förlora om vi kan låta maskiner optimera våra liv? Vi återkommer till den här frågan, men låt oss först rikta uppmärksamheten mot ett annat område där mänskliga digitala tvillingar kan vara på väg att bli verklighet, nämligen att simulera mänskliga populationer.

Om en digital tvilling av en individ kan ge insikter om hur just den personen reagerar på läkemedel, kan man då inte tänka sig en digital tvilling för att testa hur en person reagerar på olika livsbeslut eller omständigheter i vardagen?

Den mänskliga populationen i digital form

På samma sätt som en tvilling av en individ idag kan simulera interaktioner mellan organ och möjliga medicinska behandlingar, skulle en tvilling av en mänsklig population kunna utforska hur en bestämd grupp av människor kan tänkas reagera på händelser. I en sådan modell skulle varje individ vara representerad av en egen tvilling, och tillsammans skulle dessa individuella representationer (tvillingar) bilda en tvilling av populationen som helhet. Detta skulle möjliggöra simuleringar av olika framtidsscenarier för att förutse populationens samlade reaktion under olika förutsättningar.

Modeller för att simulera mänskliga populationer har använts sedan mitten på 1900-talet, i bland annat samhällsvetenskapliga och politiska syften (Chalmers, 2023). Men dessa modeller bygger på kraftigt förenklade psykologiska modeller vilket medfört att deras träffsäkerhet har varit begränsad (Chalmers, 2023). Med digitala tvillingar som bygger på verkliga individer skulle vi stå inför helt nya möjligheter.

Att gå djupare och simulera hela mänskliga populationens liv och leverne har kittlat fantasin hos många, från filosofer (Bostrom 2003; Chalmers, 2023) till dataspelsutvecklare. Ta till exempelvis populära spel som SimCity där spelaren fick bygga och förvalta en hel stad, och The Sims som låter spelaren zooma in på invånarna och deras sociala umgängeskrets. Bostrom (2003) och Chalmers (2023) har båda argumenterat för att om vi bara får tillräcklig datorkapacitet och avancerade beräkningsmetoder, kommer vi att kunna skapa virtuella världar som för sina invånare upplevs lika verkliga som vår egna. Därmed skulle vi också kunna bygga oerhört realistiska simuleringar av mänskliga skeenden.

Och visst skulle det vara otroligt spännande att simulera mänskliga samhällen, om inte annat bara för att se hur människor skulle interagera med varandra? Nyligen skapade AI-bolaget Altera en befolkning bestående av 1000 digitala agenter i spelet Minecraft (Altera, 2024; Hal Schwartz, 2024). Varje agent styrdes av en stor språkmodell (Large Language Model), och placerades i populationer av varierande storlek, från mindre grupper om 50-100 agenter till större sammansättningar om 500-1000 individer. Genom upplägget kunde Altera följa agenternas interaktioner och utvärdera deras framväxande beteende med hjälp av mått baserade på mänskliga samhällens utveckling. Agenterna, eller 'invånarna', läts agera autonomt och i realtid med varandra, och tillsammans började de utveckla beteenden som liknar människor, från att skapa professionella identiteter, till att sätta upp och följa gemensamma regler.

Under experimentets gång införde utvecklarna ett beskattningssystem som 'invånarna' kunde rösta om. Det resulterade i att de skapade olika ställningstagande grupper, för och emot beskattningen, som sedan debatterade förslaget. När grupperna tilläts bli större, från 500 agenter och uppåt, uppstod dessutom kulturer och religiösa tillhörigheter. Altera själva anser att experimentet bara var ett steg på vägen för att skapa ännu större AI simuleringar av mänskliga samhällen (Altera, 2024), vilka skulle kunna användas för att utforma exempelvis sociala policyer eller krishanteringsplaner (Hal Schwartz, 2024).

Att skapa simuleringar som inte bara bygger på generella språkmodeller, utan på specifika individer, skulle föra oss ännu närmare idén om en digital tvilling av en mänsklig population. En sådan tvilling skulle kunna bygga på personliga uppgifter som ålder, bostadsadress, familjesituation, utbildning, anställningshistorik, och en

Och visst skulle det vara otroligt spännande att simulera mänskliga samhällen, om inte annat bara för att se hur människor skulle interagera med varandra?

hel rad av attitydfrågor. Ett steg i den här riktning togs 2024 av forskare från bland annat Stanford University och Google DeepMind, som skapade en simulering som de hävdar kan replikera både individers och gruppers beteenden (Park et al., 2024). Simuleringen bygger, likt Alteras experiment, på stora språkmodeller men har kombinerats med två timmar långa intervjuer med 1052 verkliga personer. Syftet, enligt forskarna själva, är att ge beslutsfattare och forskare ett verktyg för att testa hur en mångfacetterad befolkning skulle reagera på nya politiska beslut, kampanjer, produktlanseringar eller oväntade kriser. När jag läser forskarnas artikel anar jag samma entusiasm som min vän datavetaren uttryckte 2019. Forskarna ser stora möjligheter, men har svårare att identifiera några betydande begränsningar när de pekar på potentiella användningsområden inom ekonomi, sociologi, organisationsforskning och statsvetenskap.

Kanske har dessa forskare och utvecklarna på Altera rätt, att simuleringar av individer och mänskliga populationer kan hjälpa oss hantera många av de utmaningar som finns i samhället idag? Om en digital tvilling av en population gör det möjligt att ”provköra” olika tänkbara händelseutvecklingar så kanske de kan hjälpa oss hitta bästa tänkbara utfall för beslut inom enskilda organisationer eller hela samhällen? En sådan lösning skulle kunna fungera som underlag för snåriga eller riskfyllda beslut. För frågor av en mer politisk natur, där avgränsningar och önskade mål kan stå emot varandra, ligger kanske värdet i att hitta en gyllene medelväg. Men det är snarare möjligheterna för optimering som skulle argumentera för användning av digitala tvillingar.

Min gissning är att exempelvis företag redan idag skulle kunna skapa förvänsvärt detaljerade digitala tvillingar av sin arbetsstyrka, baserat på den data som redan finns i interna system, som mejl, mötesinspelningar, kalenderposter och kommunikationsflöden. Med en sådan digital tvilling, baserat på representationer av anställda, skulle företaget potentiellt simulera olika framtida scenarier, till exempel hur en omstrukturering kan påverka produktiviteten och medarbetarnas nöjdhet. Samtidigt väcker det frågor om personlig integritet, övervakning av anställda och datadriven arbetsledning. Om den digitala tvillingen blir styrande, i den mening att beslut i praktiken anpassas efter tvillingens utfall, innebär det att arbetsstyrkan förväntas leva upp till ett digitalt ideal. I förlängningen blir det kanske ny form av Taylorism, där organisationen strävar efter att maximera produktivitet genom att optimera inte bara arbetsmoment och processer utan även arbetsstyrkan.

Andra möjliga exempel är politiska beslutsfattare som arbetar på förslag till samhällsreformer och vill förstå hur dessa kan komma att mottas av olika grupper, eller myndigheter som vill utarbeta informationsplaner för att nå ut till medborgare vid större kriser. Genom att testa antaganden i en digital tvilling av invånarna kan modellen fylla en liknande funktion som opinionsundersökningar och fokusgrupper. Samtidigt väcks i en samhällskontext frågan om hur en simulerad opinion bör vägas mot att faktiskt fråga väljare och medborgare. Om personliga data och intervjuer ligger till grund för tvillingen och fångar medborgares generella åsikter och ställningstagande, kan den digitala tvillingen då anses representera befolkningen i specifika frågor? Eller kan man se det som att medborgare i praktiken får komma till tals i fler frågor, just eftersom ”tillfrågandet” blir så lättillgänglig genom den digitala tvillingen?

Andra möjliga exempel är politiska beslutsfattare som arbetar på förslag till samhällsreformer och vill förstå hur dessa kan komma att mottas av olika grupper.

Oavsett tillämpningen kan vi nog dra slutsatsen att användandet av digitala tvillingar av mänskliga populationer skulle påverka hur vi fattar beslut, både vilka som involveras i beslutsprocessen och på vilka grunder beslut fattas.

Var finns fallgroparna?

Det är kanske framförallt lockelsen i att kunna förutspå framtiden och optimera beslut som motiverar att vi stannar upp och funderar på vilka konsekvenser en sådan utveckling kan få. Samtidigt som entusiasmen lyser från både forskare och framtidsspanare finns en rad frågor som skaver och gör framtidsvisionen obekvämt. Det rör framförallt inkludering, transparens, ansvar, agens och, i förlängningen, individuell frihet.

Om digitala tvillingar av mänskliga populationer kan identifiera de bästa tänkbara utfallen, eller åtminstone konsekvent ge rekommendationer som är lika goda som mänskliga bedömningar, vad händer då med beslutsfattarnas incitament att fråga berörda parter? Samtidigt, kommer individer vars digitala tvilling funnits representerade i den digitala populationen uppleva att de varit delaktiga i att ta fram tvillingens utfall? Kommer de ens kunna veta hur, eller om, deras digitala tvilling bidragit till de beslut som fattas? Och i vilken utsträckning blir det möjligt att ifrågasätta beslut som legitimeras genom simuleringar där ens digitala tvilling medverkat?

Avsaknaden av transparens förstärks av att många AI-modeller i praktiken fungerar som så kallade ”svarta lådor”, där det är svårt, för att inte säga omöjligt, att få insyn i AI modellens resonemang och hur enskilda faktorer har påverkat utfallet (Castelvecchi, 2016). Bristen på transparens försvårar i sin tur ansvarsutkrävande eftersom det blir oklart vad som orsakat eventuella fel (Nyholm, 2023). En parallell kan dras till diskussionen om självkörande bilar, där flera parter kan anses bära ansvaret om något går fel, inklusive den mänskliga föraren (om en sådan finns), biltillverkaren och den enskilde utvecklaren (Nyholm, 2023; Vallor, 2024).

Översatt till beslut som bygger på digitala tvillingar fås ytterligare implikationer. Om digitala tvillingar visar sig fatta bättre beslut än människor, kommer det då att uppfattas oansvarigt att inte konsultera och följa tvillingens rekommendationer? En tillförlitlig tvilling kräver en kontinuerlig synkronisering med verkligheten för att förbli relevant. Vem ansvarar i så fall för att en personlig tvilling i populationen hålls uppdaterad; utvecklaren som byggt modellen, användaren som nyttjar den, eller individen vars digitala representation ingår i tvillingen?

Fortsättningsvis uppstår frågor om vad som händer med oss ”fysiska” människor om vi blir gradvis exkluderade från själva dialogen? Det för oss tillbaka till vår tidigare fråga; Vad riskerar vi att förlora om vi optimerar allt för oss själva, men samtidigt exkluderas från processen att ta beslut vi annars skulle ha behövt brottas med? I boken *The AI Mirror* (Vallor, 2024) beskriver Shannon Vallor hur AI, trots sin imponerande kapacitet, egentligen aldrig blickar framåt. Istället, eftersom AI baserar sina resultat på historisk data, kan de enbart erbjuda en spegelbild av det förflutna. Att använda AI blir, enligt Vallor, som att gå baklänges in i framtiden och navigera genom att hålla upp en spegel för att hitta vägen framåt.

Vallor utforskar även filosofens Sellars begrepp ”space of reason” i relation till moral och AI. ”Space of reason” är den plats där vi människor resonerar, värderar argument och prövar idéer. Vallor menar att när vi överlåter beslutsfattandet till AI

**I vilken
utsträckning
blir det möjligt
att ifrågasätta
beslut som
legitimeras
genom
simuleringar
där ens
digitala tvilling
medverkat?**

riskerar vi att tappa just detta utrymme att reflektera, bedöma och ifrågasätta vad som är gott och lämpligt, samtidigt som vi riskerar att förlora utrymmet för fantasi och att föreställa oss nya möjligheter. Kort sagt när vi lämnar över beslut till AI begränsar vi vår egen förmåga att växa, resonera och tänka nytt.

Även om Shannon Vallor pratade om AI generellt, så är det inte svårt att överföra hennes resonemang på simulering med hjälp av digitala tvillingar baserade på mänskliga populationer. Här blir frågan om individuell frihet kontra organisationens eller samhällets strävan efter optimering plötsligt central. För även om simuleringar med digitala tvillingar av mänskliga populationer kan hjälpa beslutsfattare att pröva nya reformer eller förutse grupperns reaktioner på kriser, finns en risk att själva dialogen med medarbetare och medborgarna krymper. Det gäller särskilt om simuleringar börjar uppfattas som ett tillräckligt beslutsunderlag. Å ena sidan kan AI hjälpa oss att välja den mest gynnsamma, och kanske minst konfliktfyllda, vägen framåt. Å andra sidan riskerar vi att försvaga det gemensamma samtalet, kompromissens betydelse och rätten att påverka beslutens riktning, om människor inte ges möjlighet att komma till tals.

I korthet, om beslut i allt högre grad optimeras genom simulering, vilken dialog lämnas då kvar för medarbetare och medborgare att delta i? Kan det demokratiska deltagandet reduceras till en modell snarare än en process?

Vägen framåt

Och kanske är det just här den avgörande frågan uppstår: Vill vi ha ett samhälle där gemensamma beslut optimeras för att skapa bästa möjliga utfallet för så många som möjligt? Eller vill vi ha ett samhälle där vi värnar friheten att resonera och ifrågasätta och genom dialog utveckla våra ställningstaganden tillsammans? Samtidigt måste vi förhålla oss till en verklighet där deltagande och engagemang kan vara begränsat. Föreställ dig det glest besökta publika mötet där en kommun önskar skapa dialog med medborgarna kring den framtida stadsutvecklingen. Är det då mer rättvist att luta sig mot en digital tvilling där alla medborgare är representerade, än att förlita sig på den öppna diskussionen?

Även om vi gemensamt beslutar att en digital tvilling bara ska betraktas som en röst i en större dialog, kan det visa sig svårt att argumentera mot den. Vi har en benägenhet att justera våra beslut efter AI:s rekommendationer och förklaringar (Lockey, Gillespie, Holm & Someh, 2021), även när rekommendationerna går emot vår egen bedömning (Klingbeil, Grützner & Schreck, 2024). Ett sådant inflytande har dessutom visat sig uppstå även i politiska val (Lin, Czarnek, Lewis, White, Berinsky, Costello, Pennycook, & Rand, 2025). Det leder oss till slut att fundera kring den liberala idén om individuell frihet, och om den i framtiden kan komma att utmanas på ett nytt sätt, oberoende av partipolitisk hemvist och inställning till teknologisk utveckling. Hur ska frihet förstås och försvaras när både individens, organisationens och samhällets val kan formas och ledas av till synes tillförlitliga AI-system och tvillingar?

Precis som tidigare introduktioner av AI kommer utvecklingen mest sannolikt inte att gå så snabbt som många storslagna framtidsvisioner ofta utmålar. De första digitala tvillingarna av mänskliga populationer lär inte beröra samhällskritiska eller känsliga sammanhang, såsom riksdagsval eller nationellt försvar. I stället börjar utvecklingen troligen med mer vardagliga tillämpningar, exempelvis marknads-

Eller vill vi ha ett samhälle där vi värnar friheten att resonera och ifrågasätta och genom dialog utveckla våra ställningstaganden tillsammans?

föringsbolag som erbjuder tvillingar av populationer för opinionsundersökningar eller företag som skapar tvillingar av sin arbetsstyrka för enklare analyser.

Faktum är att utvecklingen redan påbörjats. Den typ av generativa agent-baserade simuleringar som presenterades ovan håller redan på att omsättas i praktiken, där plattformar som simile.ai erbjuder organisationer möjlighet att simulera hur verkliga människor kan tänkas agera. Möjligheten att använda LLM-drivna agenter som syntetiska populationer har även undersökts på policy- och samhällsnivå av EU:s Joint Research Centre (Hradec, Ostlaender & Bernini, 2023). Vad som sannolikt följer är en gradvis normalisering av verktygen. Steget in i organisationer och samhällsorgan kan tas via medarbetare och medborgare som uppmuntras att dela synpunkter genom att skapa en digital representation av sig själva.

Men trots att det kan dröja några år väcker möjligheterna med digitala tvillingar av mänskliga populationer flera följdfrågor, frågor som jag visat inte är helt enkla att besvara. Min förhoppning är att vi hinner fundera ut vad vi vill innan tekniken är här. För, handen på hjärtat, när du väl har tillgång till en tvilling du vet kan optimera dina val och visa dig vägen mot det bästa möjliga utfallet i svåra beslut, skulle du våga låta bli att fråga? Och kanske ännu svårare, skulle du våga välja annorlunda? Kan jag ana att dina ögon glittrar?

Tack

Jag vill rikta ett varmt tack till Joakim Wernberg för värdefull vägledning och kloka synpunkter under arbetet med att utveckla denna text.

I arbetet med denna text har jag använt AI-verktyg som stöd för språklig granskning och läsbarhet. Det analytiska innehållet, resonemangen och slutsatserna som dras är mina egna.

Referenser

- Agrawal, A., Gans, J., & Goldfarb, A. 2018. Prediction machines: The simple economics of artificial intelligence. Boston, Massachusetts: Harvard Business Review Press.
- AL, Altera., Ahn, A., Becker, N., Carroll, S., Christie, N., et al. 2024. Project Sid: Many-agent simulations toward AI civilization. arXiv.
- Bostrom, N. 2003. Are We Living in a Computer Simulation? The Philosophical Quarterly, 53(211): 243–255.
- Brayne, S., & Christin, A. 2021. Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts. Social Problems, 68(3): 608–624.
- Castelvecchi, D. 2016. Can we open the black box of AI? Nature, 538(7623): 20–23.
- Chalmers, D. J. 2023. Virtuella världar: Filosofiska problem. Stockholm: Fri tanke.
- Chatterji, A., Cunningham, T., Deming, D., Hitzig, Z., Ong, C., et al. 2025. How People Use ChatGPT: w34255. no. w34255, Cambridge, MA: National Bureau of Economic Research.
- Ericsson AB. 2024. 10 Hot Consumer Trends: The AI-Powered Future.
- Göteborgs Stad. 2026, February 10. Om Göteborgs digital tvilling. Om Göteborgs digital tvilling. <https://goteborg.se/wps/portal/start/goteborg-vaxer/sa-planeras-staden/goteborgs-digitala-tvilling/om-goteborgs-digitala-tvilling>.
- Hal Schwartz, E. 2024, December 2. AI characters find religion in Minecraft. Techradar. <https://www.techradar.com/computing/artificial-intelligence/ai-characters-find-religion-in-minecraft>.
- IBM. 2020, April 12. Digital Twin, IBM. What is a digital twin? <https://www.ibm.com/think/topics/digital-twin#:~:text=A%20digital%20twin%20is%20a%20virtual%20representation,critical%20variables%20that%20affect%20an%20asset's%20performance>.
- Klingbeil, A., Grützner, C., & Schreck, P. 2024. Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI. Computers in Human Behavior, 160: 108352.
- Hradec, J., Ostlaender, N., & Bernini, A. (2023). FABLES: Framework for Autonomous Behaviour-rich Language-driven Emotion-enabled Synthetic populations. Publications Office of the European Union

Lin, H., Czarnek, G., Lewis, B., White, J. P., Berinsky, A. J., et al. 2025. Persuading voters using human–artificial intelligence dialogues. *Nature*, 648(8093): 394–401.

Lockey, S., Gillespie, N., Holm, D., & Asadi Someh, I. 2021. A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions. Proceedings of the 54th Hawaii International Conference on System Science. Presented at the HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS).

Lum, K., & Isaac, W. 2016. To predict and serve? *Significance*, 13(5): 14–19.

Nyholm, S. 2023. This is technology ethics: An introduction. Hoboken: John Wiley & Sons, Inc.

O’Neil, C. 2016. Weapons of math destruction: How big data increases inequality and threatens democracy (First edition). New York: Crown.

Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., et al. 2024. Generative Agent Simulations of 1,000 People. arXiv.

Shapiro, A. 2017. Reform predictive policing. *Nature*, 541(7638): 458–460.

Simile. n.d. Simile.ai. <https://simile.ai>, February 23, 2026.

Söderlund Leifler, K. (2022, March 17). Digitala tvillingar för att förstå hur kroppen fungerar. Linköpings Universitet. <https://liu.se/nyhet/digitala-tvillingar-for-att-forsta-hur-kroppen-fungerar>

Vallor, S. 2024. The AI mirror: How to reclaim our humanity in an age of machine thinking. New York, NY: Oxford University Press.

Waardenburg, L., Sergeeva, A., & Huysman, M. 2018. Hotspots and Blind Spots: A Case of Predictive Policing in Practice. In U. Schultze, M. Aanestad, M. Mähring, C. Østerlund, & K. Riemer (Eds.), *Living with Monsters? Social Implications of Algorithmic Phenomena, Hybrid Agency, and the Performativity of Technology*, vol. 543: 96–109. Cham: Springer International Publishing.

Wikipedia. n.d. Digital Twin, Wikipedia. Digital Twin, Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Digital_twin, February 21, 2026.

OHLIN institutet

STIFTAT I BERTIL OHLINS ANDA.
VERKAR I SVENSK LIBERAL TRADITION.

Om stiftelsen Bertil Ohlininstitutet

Ohlininstitutet är en partipolitiskt oberoende liberal tankesmedja. Institutet är en plattform för att stimulera och fördjupa liberal idédebatt, det innebär att olika åsikter ges plats och att institutet inte nödvändigtvis delar de åsikter som framförs. Stiftelsen Bertil Ohlininstitutet grundades 1993 för att "stimulera vetenskaplig forskning kring de liberala idéerna och deras tillämpning på samhällsförhållanden i Sverige och världen."

Institutets finansiering

Stiftelsen Bertil Ohlininstitutet finansieras främst av bidrag från ägarstiftelser inom liberal press. Genom åren har vid enstaka projekt även andra finansiärer förekommit.

Internationellt samarbete

Stiftelsen Bertil Ohlininstitutet är medlem i den europeiska liberala tankesmedjan ELF (European Liberal Forum).

Föreläsningar

Varje år arrangeras Bertil Ohlinföreläsningen och Wbbleföreläsningen där svenska och internationella föreläsare ger ett bredare perspektiv på aktuella och tidlösa politiska problem.

Plattform för idéer

Stiftelsen Bertil Ohlininstitutet bidrar till att fördjupa och stimulera den svenska liberala idédebatten, samt ge nya liberala debattörer möjlighet att göra sig hörda. Det kan ske i olika format såsom böcker, skrifter, seminarier, poddar och filmer.

Uppsatsstävling

Stiftelsen Bertil Ohlininstitutet arrangerar årligen en uppsatsstävling inom samhällsvetenskapliga ämnen för studenter på C- och D-nivå.

Debattörspris

Bertil Ohlininstitutet delar årligen ut ett debattörspris. Priset syfte är att stimulera seriös debatt kring viktiga frågor.

Kontakt

ohlininstitutet.se

Andreas Westerberg, ordförande

kansliet@ohlininstitutet.se