



# LUND UNIVERSITY

## Vad innebär digitalisering och autonoma teknologier för individer och samhälle?

Larsson, Stefan

*Published in:*

Det demokratiska samtalet i en digital tid

2020

[Link to publication](#)

*Citation for published version (APA):*

Larsson, S. (2020). Vad innebär digitalisering och autonoma teknologier för individer och samhälle? I *Det demokratiska samtalet i en digital tid: En antologi om desinformation, propaganda och näthat* (s. 15-30). Den statliga kommittén Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet.

*Total number of authors:*

1

*Creative Commons License:*

Ospecificerad

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Vad innebär digitalisering och autonoma teknologier för individer och samhälle?

---

Stefan Larsson

# Vad innebär digitalisering och autonoma teknologier för individer och samhälle?

---

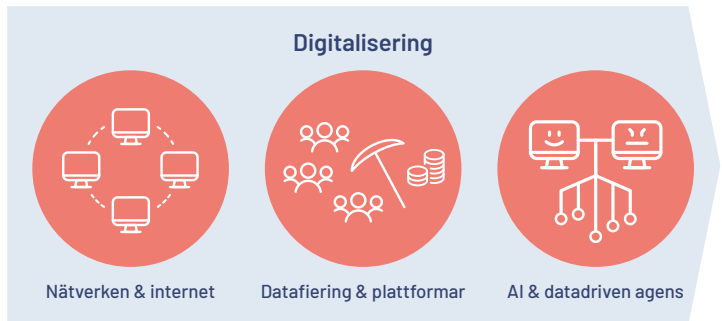
För att blicka framåt kan det ofta vara bra att först blicka bakåt. Genom datorisering, internet och digitala plattformar har vi nu åter hamnat i ett scenario med både farhågor och starka förhoppningar etablerade redan under 1900-talets mitt: den artificiella intelligensen (AI). Nu förstärkt av allt potentare och självständigare algoritmer, framdrivna av både allt kraftfullare processorer och en tämligen allomfattande *datafiering* av vår tillvaro. Jag vill i detta kapitel placera in dagens AI i sin digitala samtidshistoriska kontext – relationen till plattformar och internet som nätverk – för att därefter fördjupa bilden av de mänskliga farhågorna med autonoma teknologier. Jag vill visa på samspillet mellan databeroende AI och samhällets värderingar, för att kunna lyfta fram några av de demokratiska utmaningar som följer i kölvattnet av storskalig men vardaglig automation.

Innan en bedömning görs av vad AI och maskininlärning betyder för individer och samhälle ur ett demokratiskt perspektiv behövs en beskrivning av hur digitaliseringen utvecklats över tid.

## Digitalisering som utvecklingssteg

Dagens digitalisering har en förhistoria i flera steg. Grundläggande aspekter som elektrifiering och datorisering är naturligtvis förutsättningar, men jag vill här fokusera på i) nätverken, ii) plattformarnas datafiering och iii) de varianter av maskininlär-

ning som ofta refereras till som AI just nu. AI, och vad man kan kalla *datadriven agens*, behöver – menar jag – förstås i relation till plattformarnas tillväxt, datans roll och internet som nätverk. De samhällsvetenskapliga tänkare och forskare som tidigare fått genomslag för sina analyser av de olika stegen nedan har också beskrivit både internets och plattformarnas samhällsbetydelse, medan kunskapen om den moderna databeroende artificiella intelligensens samhällsbetydelse ännu är i sin linda.



Figur 1. Digitalisering som utvecklingssteg.

## Steg 1: Nätverken och internet

Man kan från ett samhällsperspektiv konstatera att det just är nätverken som många inflytelserika digitaliseringstänkare har fokuserat på. Här kan Manuel Castells<sup>1</sup>, Yochai Benkler<sup>2</sup> och Lawrence Lessig nämnas. Sistnämnde hänvisade 2006 till det kollaborativt uppbyggda Wikipedia som "the one surprise that teaches more than everything here".<sup>3</sup> Han gjorde det i en tid där begrepp som *crowdsourcing* och *wikinomics* myntades i samklang med idéer om hur "massans vishet" i nätverkets anda betraktades med stor optimism för samhällets och marknadens utveckling. Tecken på avfärden från denna nätverkens era kan skönjas i Lessigs personliga forskargärning. Från hur han under

ett decennium fokuserade på upphovsrättsliga frågor, där det teknikdrivet optimistiska utmanade sekelgamla föreställningar om kontroll och ägande av immateriella uttryck övergick i ett behov av att förstå det politiska systemets låsning till specifika industriers skyddsbehov. 2000-talets första decenniums digitalt drivna tillförsikt om kulturell och demokratiskt grundad frihet förbyttes därmed gradvis till insikter om samhällets trög-rörliga maktförhållanden. Det tycks ha fått Lessig att tvivla på att det (amerikanska) politiska systemet ens var byggt på demokratisk grund, och nådastöten för den digitala tillförsikten kan ha kommit vid det amerikanska presidentvalet 2016. Ett val som, tillsammans med den brittiska brexitomröstningen och skandalen runt Cambridge Analytica, kom att bli sinnebilderna för den nya tidens desinformation, politisk polarisering och riktad politisk reklam. Ska man förstå dagens förutsättningar för det demokratiska samtalet behöver rimligen kritiska insikter kring de digitala plattformarnas betydelse läggas till insikterna om nätverken och internet som struktur – och vilken roll allas vår data kommit att spela för digitala marknader såväl som för nyhets- och kunskapsförmedling överlag.

## Steg 2: Datafiering och plattformar

Om nätverken och internets tidiga tillväxt av många betraktades som en demokratiseringens och det mänskliga samskapandets tid så utmålas de digitala plattformarnas tillväxt av vissa som ett uttryck för en monopoliseringens tid. En betydande forskare inom detta fält är medievetaren José van Dijck som karaktäriserar förändringen som framväxten av ett *plattformssamhälle*.<sup>4</sup> Flertalet, med van Dijck, ser också en tydlig kommersialisering av det digitala, vilket inte minst syns på topplistan över de globalt högst börsvärderade företagen, med amerikanska teknikjättar som Apple, Microsoft, Alphabet (Google), Amazon och

Facebook i topp, och kinesiska tech-drivna Alibaba och Tencent som klättrare. Ett årtionde tidigare kännetecknades samma lista av olje- och gasföretag, finansiella konglomerat och medicinska bolag.<sup>5</sup>

Kvantifieringen och lagrandet av allt fler aspekter av vår tidigare icke-lagrade och kvalitativt upplevda tillvaro har nu medfört en "datafierad" tillvaro, med central betydelse för hur marknader fungerar och uppluckringen av hur gränser dras mellan privat och offentligt, kommersiellt och icke-kommersiellt. Här har uttryck som *uppmärksamhetsekonomi*,<sup>6</sup> *övervakningsekonomi*<sup>7</sup> och *plattformskapitalism*<sup>8</sup> föreslagits som begrepp att förstå digitaliseringens kommersiella organisering, med insamling och användning av individers data i centrum. Den skalbara tillväxten när det gäller plattformar visade sig också i mycket vara proprietär – dvs. privatägd och kontrollerad<sup>9</sup> – automatiserad och centraliserad, vilket har haft direkt betydelse för hur det demokratiska samtalet och mediedistribution tar sig uttryck, inte minst i termer av brist på insyn och transparens.

En central mekanism ligger i hur s.k. innehållsmoderering i sociala medier fungerar, dvs. hur man från privatägda plattformar hanterar sortering och distribution av vilka bilder, nyheter och vilken reklam som får spridas och hur. Samtidigt som det hela tiden sker i samspel med en större medieekologi. Hur denna moderering går till har dock konsekvenser för individers förutsättningar att få kunskap kring olika händelser och hur deras samtal går till, vilket även kan ha politiska följder. T.ex. lyfter medieforskare Tarleton Gillespie fram behovet av mer plattformstransparens och argumenterar ytterst för ett ökat ansvarstagande för de som designar plattformar, som under lång tid annars retoriskt själva har tonat ner sitt ansvar och

argumenterat för en neutral position, att de "bara" är en plattform.<sup>10</sup> Man kan konstatera att många av samhällets mest fundamentala frågor – inte minst demokratins förutsättningar i form av både politiska val och hur nyheter förmedlas – har kommit att debatteras och studeras i relation till digitala plattformars betydelse.

### Steg 3: AI och datadriven agens

Forskningen kring AI och maskininlärning har ökat kraftigt de senaste åren, och tillämpningar har blivit ett vardagligt inslag i våra smarta telefoner, översättningsapplikationer och musikrekommendationer. En rad avancerade utmaningar har visat sig vara möjliga att övervinna, påvisat av den kraftigt ökade precisionen i bildigenkänning, möjligheten att besegra en av historiens starkaste spelare av det komplexa brädspelet Go, eller att skapa fotorealistiska fejkbilder och fejkvideos genom s.k. generativa adverseriella nätverk (GANs).

Jag gör i rubriken ovan en åtskillnad mellan den artificiella *intelligensen* och en "datadriven agens". Jag gör det för att lyfta fram skillnaden mellan förmågan att lösa komplicerade problem med hjälp av stora mängder data och att vara intelligent i en mer mänsklig mening. Teknikfilosofen Luciano Floridi argumenterar för behovet av att förstå dagens AI som en väldigt potent problemlösardomän snarare än något medvetet, kännande eller på annat sätt kognitivt intelligent. Båda sidor representerar en historik inom AI-forskningen, där den ena sidan har haft en starkare utväxling än den andra.<sup>11</sup>

Samtidigt har ett forskningsbehov därmed vuxit rörande etiska och sociala utmaningar kopplade till tillämpad AI.<sup>12</sup> Behovet av styrning eller reglering har lyfts fram både på EU-nivå och av stora

plattformaktörer.<sup>13</sup> Den samhällsvetenskapliga forskningen har ännu inte hunnit ta samma helhetsgrepp kring nätverk och plattformar, men samtidigt sammanfaller många av AI-frågorna med diskussionen om digitala plattformar, som bl.a. av skalbarhets-skäl automatiserar mycket av sin verksamhet, inklusive normativt svåra avvägningar kring innehållsmoderering och vad som bedöms som relevant för vem eller vilken grupp. Det betyder att många av digitaliseringens frågor har kommit att sammanfalla med frågor om AI. Det betyder att en stor del av den digitalisering som träffar vardagens olika delar – handelns förändring, försäkringsbranschen, sjukvårdens utveckling, fordonen vi använder – stimulerar både förhoppningar och farhågor som kopplas till AI. Och, oaktat den inomvetenskapliga utvecklingen, finns det en utmaning i att säkerställa en ansvarsfull användning som kommer mänskligheten till nytta. Det som ibland lyfts fram som en dualitet med både löften och risker.

### **AI som möjlighet och utmaning: det prometianska löftet**

Vid Frankrikes lansering av en nationell AI-strategi i mars 2018 beskrev President Emmanuel Macron AI som "ett prometianskt löfte". Genom att frambringa Prometheus – titanen i grekisk mytologi som gav människorna förmågan att dra nytta av eld, mot Zeus vilja, vars straff han också ådrog sig – lyfte presidenten därmed fram en dualitet av både potential till enorm nytta för mänskligheten, såväl som medföljande risker. Han knöt därmed an till tidigare farhågor för autonoma teknologier som syns i både science-fictionlitteraturens Isaac Asimovs tre robotiklagar såväl som Mary Shelleys Frankenstein, som ju väcker en varelse till liv. Sistnämnda publicerades under tidigt 1800-tal med undertiteln "a modern Prometheus", och behandlar inte minst en av dåtidens stora innovationer: elektrifieringen.



## Potentialen i att se mönster

En förenklad beskrivning av maskininlärningens dynamiska utveckling är att det handlar om mjukvara som kan lära från exempel. Om man på förhand kan konstatera att många träningsbilder innehåller katter så kan en algoritm finna sambandet och sedan ställas inför nya bilder och sortera ut de bilder som innehåller just katter. Översatt till vården kan det betyda tidig identifiering av begynnande prostatacancer; i sociala medier kan det betyda upptäckt och censurering av falska nyheter, trollkonton och oönskat material som halshuggningar och barnpornografi; i bankernas säkerhetsarbete kan det betyda upptäckt av bedrägerier och avvikande kriminellt beteende. Medieflöden blir *individuellt relevanta* även för tjänster med miljarder samtidiga användare; musikrekommendationer tipsar om musik man aldrig hört men som faller en perfekt i smaken, fotoprogram sorterar upp familjens medlemmar genom en snabb analys av ansikten. Potentialen till skalbarhet, effektivisering och precision tycks oerhörd. Tilltron är, om inte annat, stark. Men, hur ser då utmaningarna ut?

## De oavsiktliga konsekvenserna

Det finns en tilltagande insikt i forskningen om oförutsedda och oavsiktliga samhällseffekter av autonoma system och AI. En utmaning har med de stora datamängderna att göra, i kombination med svårigheten att förklara mer exakt hur vissa typer av algoritmer finner mönster i både textmängder och bilddata-baser, och därmed vad ett visst utfall i detalj baseras på. Det kan vara svårt att förklara hur själva lärandet gick till, att utvärdera om något gick fel, och i så fall var eller varför. Härav uttrycket "black box", dvs. en svart låda utan insyn, vilket används för att beskriva en problematisk brist på förklarbarhet eller transparens. Uttrycket används både på algoritm-nivå såväl som på app-

licerad samhällsnivå.<sup>14</sup> I kombination med plattformars kommersiella intressen av att inte visa upp sina innovationer för konkurrenter eller aktörer som vill missbruka en viss applikation, och det faktum att datadrivna marknader blivit komplexa i sig – det är svårt att följa vart individers data reser – lyfts allt oftare behovet av mer transparens. Transparensfrågan är en del av de etiska utmaningar som förs fram om risker med s.k. samhällelig bias och behovet av ansvarsfull och tillförlitlig användning av AI-system.

### En studie av kommersiella ansiktsigenkänningstjänster visade att precisionen var avsevärt lägre för mörkhyade kvinnor än för ljus hy och manliga ansikten.

Det har t.ex. uppmärksammats att en av de mer använda bild-databaserna, ImageNet, som innehåller runt 14 miljoner taggade bilder, har flest bilder från ett fåtal länder, som USA och Storbritannien. Detta får konsekvenser för maskininlärning gällande kulturella uttryck, t.ex. klassificeras amerikanska bröllopsbilder med högre precision än etiopiska och pakistanska.<sup>15</sup> Dvs. databasens karaktär kommer att definiera vad som identifieras som en bröllopsklänning, i det här fallet den vita typ som är vanlig i USA, medan färggranna indiska bröllopsklänningar kategoriseras som "performance art" eller "costume".<sup>16</sup> En studie av kommersiella ansiktsigenkänningstjänster visade att precisionen var avsevärt lägre för mörkhyade kvinnor än för ljus hy och manliga ansikten.<sup>17</sup> En annan studie visade att precisionen i några av de bästa systemen för att upptäcka gångtrafikanter var olika för olika hudtoner,<sup>18</sup> vilket ledde till en diskussion om vad

konsekvenserna för autonoma fordons upptäckt av gångtrafikanter kunde bli. Ett problem ligger därmed i att träningsdatan inte är representativ för den mångfacetterade värld vi lever i. En annan studie visade t.ex. hur automatiserade jobbannonser med högre lön oftare skickades till män än kvinnor.<sup>19</sup> Här var troligen inte problemet att datan inte representerade världen på ett korrekt vis, utan snarare att världen innehåller ojämställdhet (och en rad andra strukturella utmaningar). Den automatiserade tjänsten speglade en könsmässigt skev arbetsmarknad. Här blir utmaningen snarare normativ – dvs, hur menar man att världen *borde* se ut – och vilket ansvar som kan krävas av de som designar en tjänst som samspelar med vårt skeva samhälle.<sup>20</sup> Det finns rimligen inte alltid en neutral position – man kan inte ”bara” vara en plattform om man reproducerar eller än värre förstärker samhällets sämre sidor.

Och här är det kombinationen av AI-funktionalitet och plattformarnas skalbarhet – vissa applikationer kan ha flera miljarder samtidiga användare – som formar både utmaningar och nyttor. Det finns t.ex. en kritik mot hur den delvis automatiserade innehållsmodereringen går till för de miljarder samtidiga användarna på sociala medieplattformar, för hur medierekommendationer på YouTube riskerar att bidra till sensationalism<sup>21</sup> och radikaliserings<sup>22</sup>. Det finns farhågor kring hur individuellt relevanta och automatiserade tjänster – s.k. ”personalisering” – kan leda till manipulation snarare än förstärkande av individers medbestämmande, om den dras för långt. Det finns även farhågor kring utfallet av en geopolitiskt motiverad kapprustning runt teknologier som ansiktsgenkänning och autonoma vapensystem. Nätverk av twitterbottar kan bidra till att polarisera politiska diskussioner och riktad politisk reklam kan riskera att manipulera vårt demokratiska medbestämmande. AI-system kan reprodu-

cera och förstärka vår ojämlikhet, vår rasism, vår faktaresistens och vår...kärlek till katter på internet.

## Digitalisering och autonoma teknologiers dualitet

AI-system kan, drastiskt uttryckt, vara döden förklädd i kod. Men samtidigt ska vi inte glömma att "döden" redan är förklädd i analoga kläder – i traditionellt styrda bilar och stressade mänskliga läkares felbeslut. AI och maskininlärningens kraftfulla metoder skulle också kunna representera den oförtröttlige mönsterfinnaren som kan finna ovanliga diagnoser där mänskliga experter går bet, eller den datorseende väjningsmjukvaran i ett fordon som undviker kollision. Att upptäcka det analogas skevheter kan rentav leda till ett organisatoriskt lärande, även om det i sig kan vara smärtsamt. En maskininlärningsexpert upptäckte under 2016 att franska asyldomare dömde helt olika, även när de rättsliga förutsättningarna inte skulle ge anledning till det. De var helt enkelt partiska på ett rättsosäkert sätt. Och konsekvenserna blev digra – för insynen. Tre år senare infördes en lag som förbjöd analys av individuella domares dömande.<sup>23</sup>

AI-system kan vara – de är redan – den personliga assistenten åt miljarder människor samtidigt, svarar på en ofattbar mängd frågor samtidigt och översätter alla språk, med en historiskt ojämförbar hastighet. Idén om individuell relevans är inte på något sätt förfelad i sig, allt hänger på i vilket syfte den s.k. personaliseringen utförs. En utmaning utöver styrning mot ansvarsfullhet och tillförlitlighet<sup>24</sup> ligger i behovet av att förstå vår samtida automatiserade digitaliserings betydelse för individ och samhälle. Hur det kommersiella internet, organiserat i digitala plattformar med monopolistiska datainsamlade ambitioner, påverkar samhällets demokratiska förutsättningar. Uppemot 80 år efter Asimovs robotiklagar och 200 år efter Shelleys Franken-

Vad innebär digitalisering och autonoma teknologier för individer och samhälle?

---

stein kan man konstatera att delar av utmaningarna i detta "prometianska löfte" har vi mött förut, medan skalan är ny. Det demokratiska samtalets förutsättningar behöver idag förstås i ljuset av storskalig, plattformsbaserad men vardaglig automation.

## Slutnoter

- 1 Se *The Rise of the Network Society* från 1996, som utgjorde första volymen i den tredelade *The Information Age: Economy, Society and Culture*.
- 2 Se *The Wealth of Networks* (2006).
- 3 *Code: And Other Laws of Cyberspace, Version 2*.
- 4 Vilket framförallt utvecklas i boken med samma namn, från 2018: *The Platform Society. Public Values in a Connective World*, författade av van Dijck, Poell & de Waal. För svenskt vidkommande, se också antologin *Plattformssamhället. Den digitala utvecklingens politik, innovation och reglering*, med medievetare Jonas Andersson Schwarz och undertecknad som redaktörer.
- 5 Även om börsintroduktionen av det Saudiarabiska oljeföretaget Saudi Aramco i december 2019 gjort att ett oljeföretag återtagit förstaplatsen.
- 6 Som i Hindmans bok *The Internet Trap: How the Digital Economy Builds Monopolies and Undermines Democracy* från 2018, eller Tim Wus artikel i *Antitrust Law Journal* med titeln "Blind spot: The attention economy and the law" från 2017.
- 7 Ett begrepp som är nära förknippat med Shoshana Zuboff, se exempelvis *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (2019).
- 8 Som i Nick Srniceks bok från 2017 med samma namn: *Platform Capitalism*.
- 9 Vilket, intressant nog, motsägs på mjukvaruutvecklingens område, med tillväxt av bl.a. GitHub och hanteringen av öppen mjukvara, som alltfler även av de större privata aktörerna är alltmer beroende av. Å andra sidan köptes just GitHub av Microsoft för 65 miljarder kronor under 2018, och vilken påverkan det har återstår att se.
- 10 Se *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media* från 2018. Jag tar inspiration från Gillespie och utvecklar framförallt relationen mellan plattformar och automation/AI i kapitlet "Sju nyanser av transparens: Om artificiell intelligens och ansvaret för digitala plattformars samhällspåverkan", i antologin om *Plattformssamhället*.
- 11 Se "AI i EU: Etiska riktlinjer som styrmedel" (2020).

- 12 Vilket vi visar i en flervetenskaplig genomgång av litteraturen på området för etiska, sociala och rättsliga utmaningar koppade till artificiell intelligens, se Larsson, Anneroth, Felländer, Felländer-Tsai, Heintz, Cedering Ångström, (2019) *HÅLLBAR AI*. Stockholm: AI Sustainability Center.
- 13 Jag analyserar just detta i kapitlet "AI i EU: Etiska riktlinjer som styrmedel" (2020), i Bakardjieva Engelbrekt, Michalski och Oxelheim (red.) *EU och teknologiskiftet*.
- 14 Se här exempelvis Frank Pasquales bok *Black Box Society* från 2015.
- 15 Se Shankar, S., Halpern, Y., Breck, E., Atwood, J., Wilson, J., & Sculley, D. (2017). No classification without representation: Assessing geodiversity issues in open data sets for the developing world. *arXiv preprint arXiv: 1711.08536*.
- 16 Enligt kommentar i *Nature* av Zou & Schiebinger, 2018, "AI can be sexist and racist – it's time to make it fair". <https://www.nature.com/articles/d41586-018-05707-8>
- 17 Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability and Transparency* (pp. 77-91).
- 18 Wilson, B., Hoffman, J., & Morgenstern, J. (2019). Predictive Inequity in Object Detection. *arXiv preprint arXiv:1902.11097*.
- 19 Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated Experiments on Ad Privacy Settings. *Proceedings on privacy enhancing technologies*, 2015 (1), 92-112.
- 20 För en diskussion om detta, se Larsson, S. (2019). The Socio-Legal Relevance of Artificial Intelligence. *Droit et Societe*, (3), 573-593.
- 21 Se Gillespie (2018).
- 22 Tufekci, Z. (2018). "YouTube, the Great Radicalizer". *The New York Times*, 10.
- 23 Langford, M. & Rask Madsen, M. (22 juni, 2019) "France Criminalises Research on Judges", <https://verfassungsblog.de/france-criminalises-research-on-judges/>
- 24 EU-kommissionen för fram tillit som en av fyra pelare för utvecklingen av AI, och dess AI-expertgrupp (High-Level Expert Group on AI) publicerade i april 2019 etiska riktlinjer för utvecklingen av tillförlitlig AI.