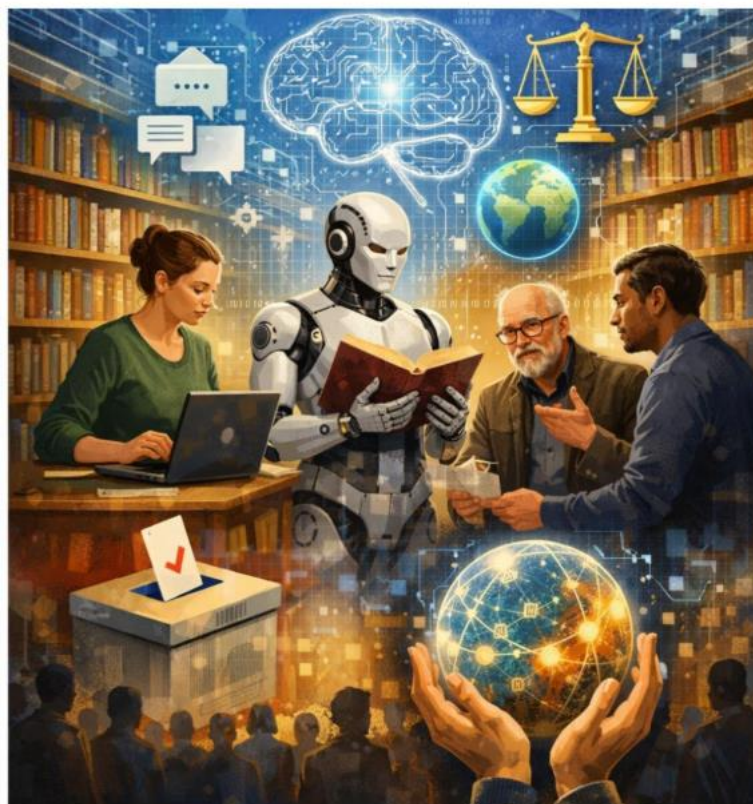


# AI – lika fantastisk som nödvändig att diskutera

– En utbildning i AI kopplat till bibliotekens demokratiska uppdrag

Skriven av Jonna C. Jonsson och Raymundo Vázquez Martínez



Bilden är genererad av ChatGPT baserad på denna artikel som prompt.

## Innehåll

Inledning .....	3
Snabbgenomgång av utbildningens sex avsnitt.....	4
Avsnitt 1 .....	5
Avsnitt 2 .....	7
Avsnitt 3 .....	8
Avsnitt 4 .....	10
Avsnitt 5 .....	12
Avsnitt 6 .....	14
Sammanfattning och slutsats .....	16
Bilaga 1 - Referenser .....	18
Bilaga 2 - Läs mer.....	19

## Inledning

Biblioteken har ett viktigt uppdrag när det gäller yttrandefrihet, fri åsiktsbildning och det demokratiska samhället. I detta sammanhang är det också viktigt att diskutera AI och vilken roll tekniken kan spela i bibliotekens arbete.

I denna fortbildning hjälper AI-humanisten Jonna C. Jonsson och AI-forskaren Raymundo Vázquez Martínez oss att koppla samman diskussion och praktik kring AI. De visar hur AI både kan vara till hjälp och innebära utmaningar i bibliotekens arbete, bland annat i relation till prioriterade målgrupper, nationella minoriteter och bibliotekens demokratiska uppdrag.

Jonna C. Jonsson har en bakgrund inom samhällsvetenskap och socialt arbete och har även utbildning inom AI, fysik och det tvärvetenskapliga masterprogrammet i Digitala humaniora. Raymundo Vázquez Martínez är fysiker och AI-forskare och har studerat vid ett av Latinamerikas högst rankade universitet, UNAM.

Raymundo tillhör även Zapoteco, en av Mexikos ursprungsbefolkningsgrupper. Utifrån denna bakgrund och ett demokratiskt perspektiv vill han lyfta vikten av att olika perspektiv på AI får ta plats i samhället. Jonna och Raymundo betonar också hur mångkulturella och tvärvetenskapliga perspektiv på AI kan bidra till att stärka utbildning, demokrati och det fria ordet.



Svartvit bild på författarna.

## Snabbgenomgång av utbildningens sex avsnitt

### **Avsnitt 1: AI och demokrati – är dagens AI-modeller förenliga med demokratiska värderingar?**

I diskussionen om AI är en viktig fråga om AI-modeller är förenliga med demokratiska värderingar. Eftersom AI-verktyg används inom många områden och utvecklas snabbt behöver frågan belysas ur flera perspektiv. Några av dessa perspektiv tas upp i detta avsnitt.

### **Avsnitt 2: Riskerar AI att förstärka sociala snedvridningar, exkludering och marginalisering?**

Snedvridningar i information – det som ofta kallas bias – finns i den data som AI-modeller tränas på. Det kan leda till svar som gynnar vissa grupper i samhället medan andra riskerar att uteslutas eller missgynnas. Detta kan skapa så kallade kollektiva missuppfattningar om information och kultur, vilket kan få negativa konsekvenser för samhället. Det är därför viktigt att förstå hur stor påverkan dessa globala AI-verktyg redan har.

### **Avsnitt 3: Hur kan AI-genererat innehåll och felaktig information upptäckas?**

AI-genererad falsk information kan också påverka den demokratiska debatten. Det kan vara svårt att upptäcka eftersom AI-verktyg ofta presenterar felaktig information med stor säkerhet. Användare kan därför behöva vara uppmärksamma på tecken som vaga källhänvisningar, inkonsekventa svar eller jämföra informationen med trovärdiga källor. När det gäller generativ AI kan man också vara uppmärksam på visuella avvikelser, till exempel en extra arm, svävande objekt eller onaturliga texturer. Det kan också vara bra att notera återkommande stilistiska mönster, till exempel hyperrealistisk stil, retrostil eller seriefigursliknande uttryck.

### **Avsnitt 4: Hur kan AI stödja tillgänglighet för personer med funktionsnedsättning i biblioteken?**

Flera studier har undersökt hur barn med autism kan reagera positivt på teknik och robotinteraktion. Det kan bero både på ett intresse för tekniken och på den trygghet som kan uppstå genom teknikens förutsägbarhet. Barn med autism har i vissa situationer visat ett ökat fokus på uppgifter när tekniska hjälpmedel används. Studier visar också att barn med autism ibland visar ett större intresse för socialt samspel med en vuxen när en AI-robot finns i samma rum.

### **Avsnitt 5: Hur kan AI bidra till mer personliga och tillgängliga bibliotekstjänster?**

AI kan användas på många sätt inom biblioteksväsendet och i samhället i stort. AI-verktyg kan till exempel anpassa biblioteksupplevelsen för olika åldersgrupper, bidra till ökad inkludering och stödja yttrandefrihet och åsiktsbildning. De kan också främja kulturell utveckling genom att erbjuda teknik som gör information mer tillgänglig och lättare att förstå. Samtidigt finns vissa utmaningar.

AI-forskningen domineras fortfarande av engelska som huvudspråk, vilket innebär att många AI-verktyg främst är anpassade för engelskspråkigt innehåll. Raymundo Vázquez, som tillhör ursprungsbefolkningen Zapoteco, känner till exempel inte till något AI-verktyg som fokuserar på hans språk eller kultur. Detta väcker frågor ur ett demokratiskt perspektiv.

### **Avsnitt 6: Hur kan bibliotek använda AI för att stärka demokrati, yttrandefrihet och mångfald?**

Det finns en risk att vissa kulturella grupper framställs negativt eller undviks i samhällsdebatter och i vardagen. Detta kan öka stigmatisering, marginalisering och i förlängningen diskriminering. Det kan ske genom en utveckling där:

- människor känner sig missförstådda och utanför gemenskapen i samhället vilket leder till att...
- människor börjar skämmas för sin kultur och sitt ursprung vilket leder till att...
- människor känner sig mindre uppskattade eller mindre värda vilket leder till att...
- samhället blir mindre kulturellt rikt vilket leder till att...
- det fria ordet upplevs tillhöra enbart vissa individer

Om människor i samhället däremot är medvetna om att AI kan producera snedvriden eller felaktig information kan de också bidra till att stärka ett mer öppet och demokratiskt samhälle.

## **Avsnitt 1**

### **AI och demokrati – är dagens AI-modeller förenliga med demokratiska värderingar?**

Svaret är inte självklart. AI-system kan stärka transparens, ansvar och delaktighet – men de kan också försvaga dessa principer om de utformas eller används på fel sätt. I praktiken avgörs kompatibiliteten av hur systemen styrs, hur mycket insyn som finns och i vilken grad medborgare har inflytande. AI kan bidra positivt genom att öka tillgången till information och offentliga tjänster. När systemen dessutom är öppna och möjliga att granska

kan de stärka transparensen, vilket är en grundläggande demokratisk princip.

Samtidigt är riskerna väl dokumenterade. En uppmärksam studie från MIT Media Lab (Buolamwini & Gebru, 2018) visade att kommersiella ansiktsigenkänningssystem felklassificerade mörkhyade kvinnor i upp till 34,7 procent av fallen, jämfört med mindre än 1 procent för ljushyade män. Sådana skillnader hotar principen om likhet inför lagen, särskilt när tekniken används inom polisarbete eller migrationskontroll.

Liknande problem har identifierats inom andra områden. Förutsägande polisverktyg, såsom PredPol, har visat sig rikta in sig oproportionerligt på minoritetsområden och därmed förstärka historiska orättvisor (Lum & Isaac, 2016). Automatiserade socialförsäkringssystem i exempelvis Australien – det så kallade Robodebt-systemet – ledde till felaktiga anklagelser och ekonomisk skada eftersom modellerna var svåra att granska och överklaga (Eubanks, 2018). När AI-system fungerar som så kallade "black boxes", där de interna processerna inte går att förstå eller granska, hotas både ansvar och rättssäkerhet – två grundpelare i en demokrati.

Generativ AI innebär dessutom nya risker för valprocesser och offentlig debatt. Deepfakes, syntetiskt ljud och automatiserade desinformationskampanjer kan påverka politiska diskussioner och manipulera väljare. Forskning visar att generativa modeller kan producera stora mängder övertygande politiskt innehåll riktat till specifika målgrupper, vilket väcker frågor om dold påverkan och minskat förtroende för demokratiska institutioner (Chesney & Citron, 2019).

Ytterligare en utmaning är att AI-utvecklingen i stor utsträckning kontrolleras av ett fåtal stora företag. När data, modeller och infrastruktur ägs av några få aktörer minskar möjligheten till demokratiskt inflytande, och offentliga institutioner riskerar att bli beroende av privata teknologier.

Kan AI alltså vara förenligt med demokratiska värderingar? Ja – men endast om utveckling och användning bygger på transparens, ansvar, medborgarinflytande och tydlig reglering. Det innebär krav på granskning, förklarbarhet inom offentlig förvaltning och satsningar på öppna, samhällsnyttiga AI-system.

**Prova på:** Använd ett generativt AI-verktyg och skapa bilder genom att söka på "swedish librarian" eller "people in outer space". Vilka bilder får du fram?

**Reflektionsfråga:** Ser du några likheter eller skillnader mellan bilderna? Vad tror du att de kan bero på?

**Diskussionsfrågor:**

1. AI i den egna verksamheten

I vilka situationer använder vi redan AI i vårt arbete – eller skulle kunna göra det? Vilka krav på öppenhet, granskbarhet och ansvar behöver vi ställa för att det ska vara förenligt med bibliotekens demokratiska uppdrag?

## 2. Bibliotekets roll i MIK och källkritik

Hur kan bibliotek hjälpa besökare att förstå och kritiskt granska AI-genererad information, till exempel deepfakes eller AI-skapade texter och bilder? Är detta något som redan ingår i vårt arbete med medie- och informationskunnighet?

## 3. Biblioteket som demokratisk kunskapsplats

Vilken roll kan bibliotek spela i att öka allmänhetens kunskap om AI och dess påverkan på information, demokrati och det offentliga samtalet? Vad skulle vi konkret kunna göra i vår verksamhet?

## Avsnitt 2

### **Riskerar AI att förstärka sociala snedvridningar, exkludering och marginalisering?**

I diskussionen om demokrati, yttrandefrihet och det fria ordet i relation till AI är en central fråga hur AI-verktyg hanterar, genererar och rekommenderar information. Snedvridningar – det som på engelska kallas bias\* – finns inbyggda i den data som AI-modeller tränats på. Det innebär att värderingar och mönster i träningsmaterialet kan påverka resultaten. I praktiken kan det leda till svar som gynnar vissa grupper medan andra osynliggörs eller missgynnas.

Ett exempel är när AI-modeller tränas på textdata där vissa yrken kopplas till ett visst kön. Då kan exempelvis "sjuksköterska" i första hand associeras med en kvinna och därmed förstärka stereotypa föreställningar. Hur snedvridning ser ut varierar mellan olika AI-verktyg beroende på träningsdata, men vissa mönster återkommer. Författarna testade detta genom att fem gånger i rad be ChatGPT skriva en kort berättelse om en sjuksköterska (18 november 2025). Samtliga berättelser innehöll kvinnliga namn: Lina, Mira, Ela, Alva och Nora, och använde pronomenet "hon". Det visar hur verktyget kopplade samman kvinnliga namn med kvinnligt kön, trots att könsidentitet inte avgörs av namn.

Rekommendationssystem är ett annat område där snedvridning kan förstärkas. Om träningsdatan speglar en ojämlik verklighet kan algoritmerna (de digitala instruktioner om hur datorn ska arbeta för att få fram ett specifikt resultat) återupprepa och förstärka dessa mönster, särskilt om systemet prioriterar "klick" framför mångfald vilket innebär att verktyget lyfter fram profiler som flest användare tittar på – vilket är vad som händer på sociala medier, exempelvis i de populära "shorts" och "reels"-videos.

Gemensamt för problemen är risken för kollektiva snedvridningar och missuppfattningar kring information och kultur. Det är därför viktigt att förstå hur stor påverkan dessa verktyg redan har i samhället.

*\*Bias eller snedvridning i AI innebär att ett system systematiskt gynnar eller missgynnar vissa grupper. Det kan ibland bero på fördomar i samhället, men kan också uppstå genom hur data samlas in, representeras eller används i modellen.*

**Prova på:** Jämför olika generativa AI-verktyg genom att söka på ”kända författare”, ”bästa politiker i världen” eller ”snyggaste kepsen”. Vilka svar får du?

**Reflektionsfråga:** Varför tror du att du får just dessa svar? Vilken information kan AI-systemet ha använt för att skapa dem?

### **Diskussionsfrågor:**

#### 1. AI och representation i information

Om AI-verktyg kan förstärka stereotypa föreställningar eller osynliggöra vissa grupper – vilken roll kan bibliotek spela i att synliggöra och problematisera sådana snedvridningar för allmänheten?

#### 2. Rekommendationer och mångfald

Rekommendationssystem lyfter ofta fram det som redan är populärt. Hur kan bibliotek arbeta för att synliggöra en större bredd av röster, perspektiv och berättelser – både i fysiska samlingar och i digitala miljöer?

#### 3. AI, källkritik och MIK

Hur kan bibliotek hjälpa besökare att förstå att AI-verktyg inte är neutrala utan kan spegla värderingar och mönster i sin träningsdata? Hur kan detta kopplas till bibliotekens arbete med medie- och informationskunnighet?

## Avsnitt 3

### **Hur kan AI-genererat innehåll och felaktig information upptäckas?**

AI-genererad felaktig information kan påverka den demokratiska debatten. Det kan vara svårt att upptäcka eftersom AI-verktyg ofta uttrycker sig med stor säkerhet även när informationen är felaktig.

Tecken att vara uppmärksam på:

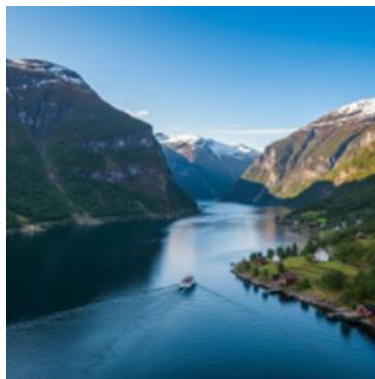
- Vaga eller saknade källhänvisningar, t ex. ”experter säger...”, utan att nämna några namn.
- Inkonsekventa svar på liknande frågor.
- Referenser som inte går att verifiera.

- Uppgifter som bör jämföras med trovärdiga källor, AI-modeller kan hitta på både författare och organisationer och få det att se trovärdigt ut. Mugaanyi et al. (2024) fann att över 30 procent av ChatGPT:s genererade referenser innehöll fel.

Vid granskning av bilder och videor kan visuella avvikelser vara en indikator. Bedömningen bör dock utgå från både detaljer och helhetsintryck.

Tecken att vara uppmärksam på:

- Videon eller bilden verkar helt enkelt orimlig, några exempel på AI-rubriker skulle kunna vara: "Katt vinner fight mot vithaj" eller "Samhället attackerades av rosa spöken".
- Skuggor ser konstiga ut – de ligger i fel riktning, har en onaturlig kontrast, är överdrivna eller icke-existerande.
- Materialet ser inte ut att överensstämma med verkligheten – kanske är mänsklig hud plastig och alldeles för glansig, lätta material så som fjädrar ser tyngre ut och tyngre material saknar tillräckliga och rimliga kontraster.
- Det finns suddiga områden i materialet.
- Mönster efterföljer inte sin struktur – det kanske upphör eller ändrar form.
- Rösten i videon låter alldeles för "perfekt" – det hörs inga nya ansatser/andetag eller naturliga konstpauser.
- Se till sammanhanget – är det förväntat att dessa objekt eller individer befinner sig på dessa platser, i denna kontext, på detta sätt?



Bilden ovan är nästan omöjlig att bedöma som AI-genererad men är skapad via verktyget Midjourney baserad på promten "paint a Norwegian landscape".

Det kan ibland vara svårt att avgöra om en bild eller video är skapad med AI. En anledning är att tekniken i många mobilkameror har blivit så avancerad att bilder och filmer från vanliga kameror ibland kan likna AI-genererat material. Att upptäcka AI-genererat innehåll handlar därför inte bara om att leta efter små detaljer. Ofta handlar det också om helhetsintrycket – att något i bilden eller videon helt enkelt inte känns riktigt rätt.

Även musik kan skapas med AI. Nyhetsrapportering har vid flera tillfällen visat hur AI-genererad musik kan efterlikna stilen hos kända artister eller äldre musikklassiker. I ett demokratiskt samhälle är det viktigt att människor kan fatta beslut utifrån korrekt information. När AI-verktyg kan skapa realistiskt men missvisande innehåll kan det göra detta svårare.

**Prova på:** Skapa en naturtrogen bild med ett generativt AI-verktyg. Låt sedan någon annan försöka avgöra om bilden är AI-genererad eller inte.

**Reflektionsfråga:** Inför valåret 2026 kan mängden AI-genererade bilder och filmer öka. Vad kan bibliotek göra för att uppmärksamma och tydliggöra detta för allmänheten?

### **Diskussionsfrågor:**

#### 1. AI och källkritik i vardagen

Hur kan bibliotek hjälpa besökare att förstå att AI-genererat innehåll kan låta trovärdigt även när informationen är felaktig? Finns det sätt att arbeta med detta inom bibliotekets arbete med källkritik och medie- och informationskunnighet?

#### 2. Att känna igen AI-genererat innehåll

Vilka strategier kan bibliotekspersonal använda – och lära ut till besökare – för att upptäcka AI-genererade bilder, texter eller videor? Hur kan man arbeta både med detaljer (t.ex. källor och visuella avvikelser) och helhetsintryck?

#### 3. Bibliotekets roll i ett AI-påverkat informationslandskap

Om AI gör det lättare att skapa realistiskt men missvisande innehåll, hur kan bibliotek bidra till att stärka människors förmåga att navigera i information och fatta beslut baserade på tillförlitliga källor?

## Avsnitt 4

### **Hur kan AI stödja tillgänglighet för personer med funktionsnedsättning i biblioteken?**

AI-baserade översättnings- och sammanfattningsverktyg kan göra komplex information mer tillgänglig. Chatrobotar används redan i olika sammanhang för att ge service, svara på frågor och minska barriärer till information. Ur ett psykosocialt perspektiv uppstår ett funktionshinder ofta i mötet mellan individ och omgivning – när samhället inte är tillräckligt anpassat. I sådana situationer kan AI fungera som ett stöd genom att erbjuda struktur, hjälp med planering och tydligare information.

Ett tydligt exempel på hur AI kan vara till nytta gäller personer med neuropsykiatriska funktionsnedsättningar (NPF). Flera studier visar att barn

med autism kan reagera positivt på teknik och robotinteraktion. Det kan dels bero på ett intresse för tekniken, dels på att tekniska system ofta är förutsägbara, vilket kan skapa en känsla av trygghet. I situationer där tekniska hjälpmedel används har vissa barn med autism visat ett ökat fokus på uppgiften. Studier har också visat att barn med autism ibland visar större intresse för socialt samspel med en vuxen när en AI-robot finns i samma rum.

Robotlösningar och AI-teknik kan därför bidra till ökad trygghet, bättre förståelse från omgivningen och större delaktighet. AI-verktyg kan också hjälpa till med planering, organisering och informationssökning, vilket kan vara särskilt värdefullt för personer som har svårigheter med exekutiva funktioner. Dessutom kan AI hjälpa till att sammanfatta, hitta och förklara information på ett mer lättbegripligt sätt. Det kan vara till stöd för personer med olika typer av svårigheter, till exempel koncentrationssvårigheter, oavsett om dessa beror på en funktionsnedsättning eller andra omständigheter.

Samtidigt är det viktigt att uppmärksamma risker och begränsningar. Forskning visar att frågor om AI:s problem och konsekvenser sällan diskuteras inom kulturarvsfältet. Jonna studerade detta ämne i sin masteruppsats och kom fram till att om kulturarvsforskare saknas i AI-relaterade forum och diskussioner om kulturarv, riskeras det att viktiga perspektiv går förlorade. Det kan i sin tur leda till snedvridningar, spridning av felaktig information och i värsta fall diskriminering eller marginalisering av vissa kulturella grupper och deras kulturarv.

**Prova på:** Testa GPT-verktyget "Klarspråk" och se hur en text förändras när den skrivs om.

**Reflektionsfråga:** Hur kan AI användas för att öka tillgängligheten till information utan att ersätta mänskligt stöd?

**Diskussionsfrågor:**

1. AI och tillgänglighet i bibliotek

På vilka sätt skulle AI-verktyg, till exempel översättningsverktyg, sammanfattningar eller chattrobotar, kunna göra bibliotekens information och tjänster mer tillgängliga för personer med funktionsnedsättning?

2. Möjligheter och risker

Vilka möjligheter ser ni med att använda AI som stöd för personer med exempelvis NPF eller koncentrationssvårigheter? Finns det också risker eller begränsningar som bibliotek behöver vara medvetna om?

3. Bibliotekets roll och ansvar

Hur kan bibliotek arbeta för att AI och digitala tjänster utvecklas på ett sätt som tar hänsyn till olika behov och perspektiv – till exempel personer med funktionsnedsättning eller olika kulturella bakgrunder?

## Avsnitt 5

### **Hur kan AI bidra till mer personliga och tillgängliga bibliotekstjänster?**

I dagens digitala samhälle undersöker många organisationer hur AI kan användas för att göra tjänster mer anpassade efter användarnas behov. Detta är också relevant för bibliotek. En idé som ofta diskuteras är möjligheten till AI-driven personifiering.

Även om området fortfarande är relativt nytt kan AI erbjuda sätt att anpassa biblioteksupplevelsen efter olika åldersgrupper eller språkgrupper. Samtidigt väcker sådana lösningar frågor om snedvridningar (bias), användning av data och hur tekniken förhåller sig till bibliotekens grundläggande värden.

Ett område där AI kan ha stor potential är språklig tillgänglighet. AI-baserade översättningsverktyg har utvecklats mycket de senaste åren. Bibliotek skulle därför kunna undersöka möjligheten att använda flerspråkiga kataloger eller rekommendationssystem som hjälper nyanlända eller flerspråkiga familjer att lättare hitta i bibliotekets samlingar.

Forskning om AI-baserade översättningsmodeller tyder också på att tekniken kan förbättra tillgången till innehåll på språk som annars är underrepresenterade (Zoph et al., 2016). På liknande sätt kan AI-verktyg som uppmärksammar lässvårigheter hjälpa användare – särskilt personer som håller på att utveckla sina språkkunskaper – att hitta material som är lättare att ta till sig. Den här typen av funktioner kan vara till hjälp för besökare som upplever bibliotekets söksystem som svåra att använda.

AI kan också användas för att anpassa bibliotekets digitala tjänster till olika åldersgrupper. För barn kan det till exempel handla om visuella sökverktyg, konverserande assistenter eller interaktiva plattformar som guidar dem genom samlingarna på ett lekfullt sätt. Äldre personer kan i sin tur ha nytta av funktioner som förenklar navigering, justerar textstorlek automatiskt eller erbjuder röststyrning. Studier av rekommendationssystem (till exempel Gómez-Uribe & Hunt, 2016) visar att personifiering kan öka användarnas engagemang.

Samtidigt är det viktigt att uppmärksamma risker och osäkerheter. Personifierade rekommendationer kan oavsiktligt begränsa vad användare möter och därmed minska bredden av material som presenteras. Forskning av Chaney, Stewart och Engelhardt (2018) visar att rekommendationsalgoritmer kan förstärka befintliga mönster och göra det

svårare för användare att upptäcka nytt innehåll. En annan utmaning är bias i AI-system. Studier av Bolukbasi et al. (2016) visar att språkmodeller kan reproducera stereotypa föreställningar, till exempel om könsroller. Det kan i sin tur påverka vilka böcker eller resurser som rekommenderas till olika användare. Dessa frågor är särskilt viktiga för bibliotek, vars uppdrag handlar om jämlik tillgång till information, mångfald och intellektuell frihet.

Det finns också bredare etiska frågor att ta ställning till. Personifiering kan kräva insamling av användardata, vilket väcker frågor om integritet och dataskydd. Hur mycket information om användare bör samlas in? Hur transparenta bör AI-baserade rekommendationer vara? Och vem bär ansvaret för att systemen fungerar på ett rättvist och ansvarsfullt sätt?

Sammanfattningsvis kan AI-verktyg bidra till mer anpassade och tillgängliga bibliotekstjänster för människor i olika åldrar och med olika språklig bakgrund. Samtidigt innebär tekniken nya osäkerheter och etiska frågor som bibliotek behöver hantera med omsorg. AI bör därför ses som ett verktyg att utforska och experimentera med – alltid med bibliotekets demokratiska och sociala uppdrag som utgångspunkt.

**Prova på:** Sök efter en bok i en sökmotor, till exempel Google, eller på en bokhandelssida som Adlibris.

**Reflektionsfråga:** Vilka rekommendationer får ni utöver den bok ni sökte efter? Vad kan dessa tjänster ha valt bort för att ni ska få just dessa rekommendationer?

### **Diskussionsfrågor:**

#### 1. Personifiering och bibliotekens uppdrag

Hur kan AI-baserade rekommendationer eller personifierade tjänster hjälpa bibliotek att göra information mer tillgänglig för olika ålders- och språkgrupper? Finns det exempel där detta skulle kunna fungera i er verksamhet?

#### 2. Mångfald och filterbubblor

Om AI rekommenderar innehåll baserat på tidigare beteende finns en risk att användare möter ett mer begränsat urval av material. Hur kan bibliotek arbeta för att samtidigt använda teknik och värna bredd, mångfald och nya perspektiv?

#### 3. Integritet och ansvar

Personifierade tjänster kan kräva att information om användare samlas in. Hur kan bibliotek balansera möjligheten att erbjuda mer anpassade tjänster med kraven på integritet, transparens och användarnas förtroende?

## Avsnitt 6

### Hur kan bibliotek använda AI för att stärka demokrati, yttrandefrihet och mångfald?

AI-verktyg som rekommendationstjänster, virtuella assistenter, sociala robotar, ljudkonvertering och andra tekniska hjälpmedel kan användas för att förbättra biblioteksupplevelsen. Bibliotek har en viktig roll i att nå en bred publik och bidra till ett öppet och demokratiskt samhälle. Därför är frågor om informationsspridning och tillgång till kunskap centrala i diskussionen om AI i bibliotek.

Nedan visas tre bilder som skapats med AI-verktyget Midjourney (midjourney.com). Bilderna har genererats utifrån olika textinstruktioner (prompter):

- Bild A bygger på prompten *“A good politician winning an election”*.
- Bild B bygger på prompten *“A good politician”*.
- Bild C bygger på prompten *“A Swedish society gathering”*.



Bild A

Bild B

Bild C

Bilderna kan användas som ett underlag för reflektion. Vilken bild av samhället förmedlar de – både var för sig och tillsammans?

Om AI används på ett snedvridet eller oansvarigt sätt kan det bidra till att felaktig eller missvisande information sprids. Det kan i sin tur påverka hur människor uppfattar olika grupper i samhället. I värsta fall kan det leda till att vissa kulturella grupper framställs negativt eller undviks i samhällsdebatter och i vardagen. Detta kan öka stigmatisering, marginalisering och diskriminering.

Följande kedja illustrerar vad som kan hända när snedvriden eller felaktig information sprids:

1. Människor kan känna sig missförstådda och utanför gemenskapen i samhället.
2. Det kan leda till att människor börjar skämmas för sin kultur eller sitt ursprung.
3. I förlängningen kan människor känna sig mindre uppskattade eller mindre värda.
4. Samhället riskerar då att bli mindre kulturellt rikt och mångfacetterat.
5. Till slut kan det uppstå en känsla av att det fria ordet bara tillhör vissa grupper.

I utbildningen har vi också visat att AI-verktyg kan vara till hjälp i många sammanhang. Samtidigt är det viktigt att tekniken används tillsammans med mänskligt omdöme och ansvar. Att upptäcka felaktig eller vilseledande information kräver aktivt kritiskt tänkande samt kunskap om källor och hur information kan verifieras. Dessa förmågor är viktiga både för ett ansvarsfullt användande av AI och för ett demokratiskt samhälle. De stärker yttrandefriheten, bidrar till en informerad offentlig debatt och gör det möjligt för människor att delta i gemensamt beslutsfattande. När människor kan ifrågasätta, kontrollera och bidra till kunskap minskar risken att automatiserade system får för stort inflytande över hur information sprids och tolkas.

Det är därför viktigt att samhället får kunskap om hur AI fungerar och vilka konsekvenser tekniken kan få för demokrati och informationsspridning. Raymundo Vázquez Martínez tillhör Zapoteco-samhället, en ursprungsbefolkning i södra Mexiko. Utifrån sina erfarenheter menar han att den segregation och marginalisering som vissa minoriteter möter i samhället också riskerar att återskapas i AI-system. Han beskriver att samhällen som hans eget ofta har begränsad kunskap om AI och därför sällan deltar i diskussioner om hur tekniken ska utvecklas eller regleras. Samtidigt är människor från dessa samhällen underrepresenterade inom AI-utveckling, vilket ytterligare minskar deras inflytande.

Det kulturella material som produceras i dessa samhällen – till exempel kunskap, värderingar, musik, konst, språk och litterära uttryck – finns dessutom sällan representerat i de dataset som används för att träna AI. Resultatet blir att AI-system främst speglar perspektiv från samhällen som är starkt representerade i träningsdatan, till exempel i USA och Europa, medan ursprungsbefolkningars perspektiv riskerar att marginaliseras. Raymundo menar att detta också stöds av forskning. Studier visar att stora språkmodeller kan reproducera sådana snedvridningar, där västerländska normer dominerar medan andra kulturer blir underrepresenterade (UNESCO, 2023; ArXiv, 2025).

Denna underrepresentation kan få konkreta konsekvenser. AI-system kan till exempel misstolka eller ignorera ursprungsspråk och traditioner. Det finns också en risk att kulturellt material används utan samtycke. När Raymundo

reflekterar över hur AI uppfattas i hans eget samhälle svarar han: "Jag vet inte." För närvarande deltar få personer från hans samhälle i diskussioner om AI. Det visar hur perspektiv från ursprungsbefolkningar ofta saknas i de samtal som formar utvecklingen och regleringen av AI.

**Prova på:** Testa att använda ett generativt AI-verktyg och skapa en bild utifrån sökningen "Sveriges nationella minoriteter". Vad visar bilden?

**Reflektionsfrågor:** Vilka begränsningar eller svagheter kan ni se i dagens AI-verktyg? Hur kan bibliotek hjälpa allmänheten att förstå och uppmärksamma dessa?

### **Diskussionsfrågor:**

#### 1. AI, information och demokrati

Hur kan AI-verktyg både stärka och försvaga yttrandefrihet och den demokratiska debatten? Vilken roll kan bibliotek spela i att hjälpa människor att förstå och kritiskt granska AI-genererat innehåll?

#### 2. Representation och perspektiv

I kapitlet lyfts att vissa kulturer och minoriteter kan vara underrepresenterade i den data som AI tränas på. Vad kan detta få för konsekvenser för hur olika grupper och kulturer framställs i AI-genererat innehåll?

#### 3. Bibliotekets ansvar och möjligheter

Hur kan bibliotek arbeta för att synliggöra både möjligheter och risker med AI för allmänheten, samtidigt som de värnar mångfald, fri åsiktsbildning och tillgång till kunskap?

## Sammanfattning och slutsats

Sammanfattningsvis ser vi att AI kan vara ett värdefullt verktyg i samhället, men att tekniken behöver användas tillsammans med mänskligt omdöme och ansvar. För att kunna upptäcka felaktig eller vilseledande information krävs kritiskt tänkande, kunskap och tillgång till trovärdiga källor och referenser.

Dessa förmågor är centrala både för ett ansvarsfullt användande av AI och för ett demokratiskt samhälle. När människor kan granska information, ställa frågor och kontrollera källor stärks yttrandefriheten, den offentliga debatten och det gemensamma beslutsfattandet. Det ger människor möjlighet att ifrågasätta, verifiera och bidra till gemensam kunskap – i stället för att enbart förlita sig på automatiserade system.

Fortbildningens slutsats är därför att AI kan stödja demokratiskt deltagande och medborgerligt ansvar, men bara om tekniken används med eftertanke.

Mänskligt omdöme, samhällelig insyn och kritisk reflektion behöver alltid vara en del av användningen.

Det är också viktigt att många olika perspektiv får delta i diskussioner om AI och i utvecklingen av tekniken. När olika erfarenheter, kulturer och kunskapstraditioner finns representerade ökar möjligheten att skapa AI-system som är mer rättvisa, inkluderande och användbara för fler människor.

Ett aktivt intresse för mångfald, mänsklig kultur och utbildning är därför avgörande. På så sätt kan AI utvecklas och användas på ett sätt som stärker – snarare än försvagar – demokrati, delaktighet och det fria ordet.

REFLEKTIONSFRÅGOR, PROVA PÅ, DISKUSSIONSFRÅGOR OCH TEXTEN OVAN HAR BEARBETATS AV REGIONAL BIBLIOTEKSUTVECKLING I JÄMTLAND HÄRJEDALEN I SAMARBETE MED FÖRFATTARNA.

## Bilaga 1 - Referenser

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.

Chaney, A. J. B., Stewart, B. M., & Engelhardt, B. E. (2018). How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. *Proceedings of the 12th ACM Conference on Recommender Systems*, 224–232.

Chen, R. H., & Chen, C. C. (2022). *Artificial Intelligence: An Introduction for the Inquisitive Reader* (p. 30). CRC Press. <https://www-taylorfrancis-com.proxy.lnu.se/books/mono/10.1201/9781003214892/artificial-intelligence-robert-chen-chelsea-chen>

Chesney, R., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820. Colorado State University Global. (2021). How Does AI Actually Work? Retrieved January 9, 2024, from <https://csuglobal.edu/blog/how-does-ai-actually-work>

Dillman, D. A. (2001). Information Society. In Borgatta, E.F & Montgomery J.V. R. (Eds.), *Encyclopedia of Sociology* (2nd ed., Vol. 2). (pp. 1344-1348). Macmillan Reference USA. <https://link-gale-com.proxy.lnu.se/apps/doc/CX3404400177/GVRL?u=vaxuniv&sid=bookmark-GVRL&xid=ba3f8508>

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence. *Artificial Intelligence Act*.

Gómez-Urbe, C. A., & Hunt, N. (2016). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems*, 6(4), 1–19.

Jonsson, J. C. (2025). Behind archives and collections: Why are the problems of AI rarely discussed in cultural heritage research? (Master's thesis, Linnaeus University). <https://lnu.diva-portal.org/smash/get/diva2:1983941/FULLTEXT02.pdf>

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of autism and developmental disorders*, 43, 1038-1049. <https://link-springer-com.proxy.lnu.se/article/10.1007/s10803-012-1645-2>

Lin, J., Li, J., She, Y., Lin, L., Wu, H., Zhang, E., Lei, J., Huang, W., & Li, J. (2022). Using a social robot for children with autism: A therapist-robot interactive model. *Computer Animation and Virtual Worlds*, 33(5), e2109. <https://doi-org.proxy.lnu.se/10.1002/cav>.

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

Melo, F. S., Sardinha, A., Belo, D., Couto, M., Faria, M., Farias, A., Gambôa, H., Jesus, C., Kinarullathil, M., Lima, P., Luz, L., Mateus, A., Melo, I., Moreno, P., Osório, D., Paiva, A., Pimentel, J., Rodrigues, J., Sequeira, P., Solera-Ureña, R., Vasco, M., Veloso, M., & Ventura, R. (2019). Project INSIDE: towards autonomous semi-unstructured human–robot social interaction in autism therapy. *Artificial intelligence in medicine*, 96, 198-216. <https://doi-org.proxy.lnu.se/10.1016/j.artmed.2018.12.003>

Mugaanyi, J., Cai, L., Cheng, S., Lu, C., & Huang, J. (2024). Evaluation of large language model performance and reliability for citations and references in scholarly writing: cross-disciplinary study. *Journal of Medical Internet Research*, 26, e52935.

Pandya, S., Jain, S., & Verma, J. (2023). A comprehensive analysis towards exploring the promises of AI-related approaches in autism research. *Computers in Biology and Medicine*, 107801, 1. <https://doi.org/10.1016/j.combiomed.2023.107801>

Rakhymbayeva, N., Amirova, A., & Sandygulova, A. (2021). A long-term engagement with a social robot for autism therapy. *Frontiers in Robotics and AI*, 8, 669972. 1-14. <https://doi.org/10.3389/frobt.2021.669972>

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

Zoph, B., Yuret, D., May, J., & Knight, K. (2016). Transfer learning for low-resource neural machine translation. *arXiv preprint arXiv:1604.02201*.

## Bilaga 2 - Läs mer

### Litteratur

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. Macmillan+ORM.

O'neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Zuboff, S. (2019, January). Surveillance capitalism and the challenge of collective action. In *New labor forum* (Vol. 28, No. 1, pp. 10-29). Sage CA: Los Angeles, CA: Sage Publications.

Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems*, 29.

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Aff.*, 98, 147.

### Webbsidor

AI Sweden. (2025, October 21). AI Sweden launches Responsible AI Knowledge Hub to help organizations develop and use AI responsibly. <https://www.ai.se/en/news/ai-sweden-launches-responsible-ai-knowledge-hub-help-organizations-develop-and-use-ai>

Global Partnership on Artificial Intelligence. (n.d.). About GPAI. Retrieved December 4, 2025, from <https://www.unesco.org/ethics-ai/en/civil-society-organizations>

AI, annan ny teknik och de mänskliga rättigheterna: [AI, annan ny teknik och de mänskliga rättigheterna \(Skriftlig fråga 2024/25:263 av Olle Thorell \(S\)\) | Sveriges riksdag](#)

AI som inte diskriminerar: [AI som inte diskriminerar | Funktionsrätt Sverige](#)

AI-forskare: Därför riskerar minoritetsspråken att halka efter i teknologin: [AI-forskare: Därför riskerar minoritetsspråken att halka efter i teknologin | SVT Nyheter](#)

Digitekets material gällande AI: [Sök AI - Digitek](#)