



UMEÅ UNIVERSITET

VÄLKOMMEN TILL KUNSKAPSVECKAN

WELCOME



UMEÅ UNIVERSITET

AI om AI

Lärdomar från en försök med AI-genererad kurslitteratur

Jonathan Wedman
Fil. dr i Beteendevetenskapliga mätningar
Excellent lärare

Upplägg

- Bakgrund – Vad och varför?
- Utvecklingen
- För- och nackdelar med AI-genererad kurslitteratur
- Lärdomar
- Framtid



UMEÅ UNIVERSITY

Bakgrund

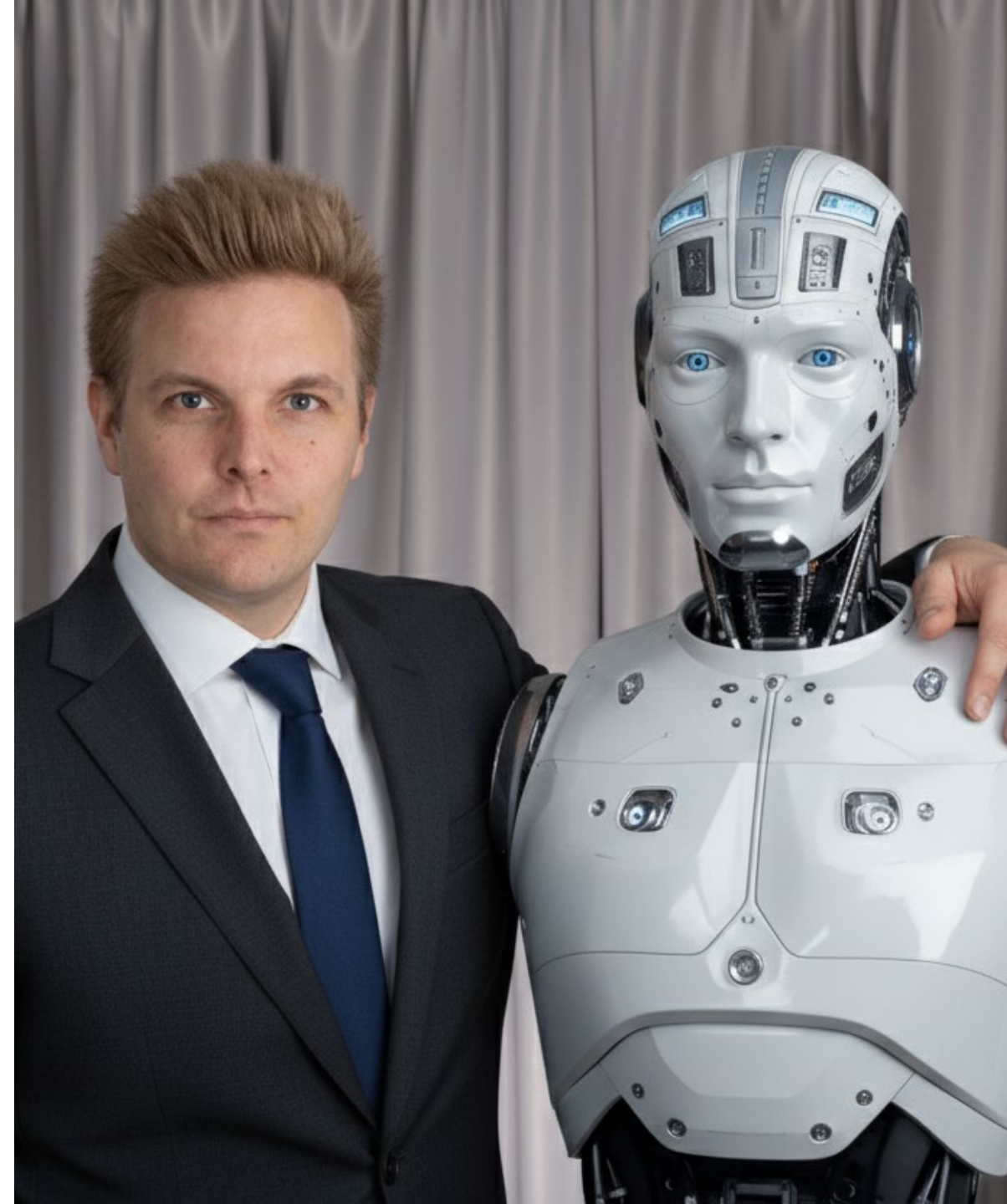
Bakgrund – Vad har jag gjort?

- Jag har promptat en bok som heter "AI om AI"
- ...och en ordlista med 70+ AI-relaterade begrepp och koncept



Bakgrund – Men varför?

- Intresse för AI och kursansvarig
 - Introduktion till AI i skolan, 7,5hp (2IT041)
 - Fördjupning om AI i skolan, 7,5hp (2IT042)
 - Att undervisa AI på gymn.: Tekniska perspektiv, 7,5hp (6PE335)
- Åtskilliga gästföreläsningar och fortbildningsdagar om AI
 - I bedömning
 - I undervisning
 - I forskning
- Svårt med traditionell kurslitteratur att hinna med
- Leva som jag lär





UMEÅ UNIVERSITY

Utvecklingen

Processen från idé till färdig bok

- **Stegvis promptning för "AI om AI"**

1. AI-litteracitet enligt Skolverket m.fl:
 1. Teknisk del
 2. Praktisk del
 3. Samhällelig/Etisk del
2. Skapa övergripande innehåll till varje del
3. Använda ChatGPTs *Deep research* för att skapa innehåll.

Skolverket

- **Deep Research om begrepp**

1. Vad säger Skolverket och internet om begrepp?
2. Använda ChatGPTs *Deep research* för skapa innehåll.



1. Skolverket (kommentarer till AI)

Ämnets karaktär och plats i programstrukturen

Ämnet artificiell intelligens är till sin karaktär tvärvetenskapligt. Innehållet i ämnet kan delas in i tre delar; en teknisk, en samhällsvetenskaplig och en tillämpad del. Undervisningen ska behandla kunskaper om AI och dess

Centralt innehåll, artificiell intelligens 1

Det centrala innehållet i ämnet återfinns under tre rubriker; artificiell intelligens, praktisk användning av AI och förhållningssätt till AI. Rubrikerna har framför

Bedömningsplan: Artificiell intelligens – Artificiell intelligens 1¹, 100 poäng (ARTART01)

(Teknisk/Tillämpad/Samhällsvetenskaplig del)

Betygskriterier (E/C/A)	Centralt innehåll	Kognitiv nivå och process	Bedömningsmetod
Eleven visar godtagbara/goda/mycket goda kunskaper om AI, dess tillämpningar, tekniker och beståndsdelar.	- Definition av AI samt centrala begrepp inom området. - Vikten av data, datakvalitet för AI och val av data. - Drivkrafter bakom utvecklingen av AI. - Översikt över användningen av AI, däribland prediktion, robotik, vision och generativ AI samt olika tekniker och metoder som ligger bakom dessa. - Översikt av tekniker för AI, däribland sökning, klassificering och objektigenkänning. - Översikt av metoder och algoritmer inom AI, däribland beslutsträd, regression samt övervakat och oövervakat lärande.	Minnas	Stängda och korta frågor
Eleven använder med viss/-/god säkerhet AI för att lösa enkla problem	- Översikt över processen för problemlösning med AI. - Enklare typ av problemlösning med hjälp av AI, till exempel klassificering, objektigenkänning, prediktion, tolkning och bearbetning av naturligt språk (NLP), enklare maskininlärning och användning av spelagent. - Principer för, och tillämpningar av, vanliga tekniker inom AI, däribland maskininlärning och robotik. - Metoder för enklare träning av AI.	Tillämpa² - Genomföra	Praktiskt prov
Eleven gör enkla/välgrundade/välgrundade och nyanserade jämförelser av mänsklig och artificiell intelligens.	- Jämförelse mellan hur enklare lösningar med AI fungerar och hur en människa löser samma problem. - Några situationer där AI är överlägsen människan och tvärtom.	Förstå³ - Jämföra	Essäfrågor
Eleven redogör på ett godtagbart/utvecklat/väluutvecklat sätt för lagar och andra bestämmelser som gäller för AI.	Normer samt lagar och andra bestämmelser som gäller AI och dess användning.	Minnas - Komma ihåg - Känna igen Förstå - Förklara	Stängda och korta frågor
Eleven för enkla/utvecklade/utvecklade och nyanserade resonemang om möjligheter och risker med AI samt dess konsekvenser för samhället.	- Etiska dilemman med användandet av AI, däribland transparens. - Demokratiska, sociala, ekonomiska, miljömässiga och säkerhetsmässiga möjligheter och risker med AI-användning samt dess konsekvenser för samhället.	Förstå³ - Förklara	Muntligt prov

2. Övergripande innehåll

Du är expert på AI-litteracitet i utbildning med 20 års erfarenhet av att utbildningsrelaterad forskning relaterad till AI och AI-litteracitet. Lista de viktigaste komponenterna för AI-litteracitet inom vart och ett av de tre områdena: Tekniska, Praktiska och Etiska aspekter av AI.



2. Övergripande innehåll

🔧 1. Tekniska aspekter av AI

Dessa komponenter rör förståelsen för hur AI fungerar på ett grundläggande plan, utan krav på att själv kunna programmera avancerade modeller:

- **Grundläggande begrepp inom AI**
T.ex. maskininlärning, neurala nätverk, algoritmer, träning och testning av modeller.
- **Datastrukturer och datakvalitet**
Förstå vikten av datasetens roll i AI-system, inklusive bias, datarensning och etik kring datainsamling.
- **Automatiseringsprinciper**
Kunskap om hur AI-system efterliknar mänskligt beslutsfattande och vilka tekniker som används (klassificering, prediktion, NLP, etc.).
- **Modellers funktion och begränsningar**
Känna till att AI inte är allvetande – förstå begrepp som överanpassning, black-box-problematik och probabilistiskt tänkande.
- **Skillnaden mellan smal och generell AI**
Medvetenhet om att dagens AI är "smal" och optimerad för specifika uppgifter.



3. Deep research

Skriv en utförlig rapport som inkluderar utförliga beskrivningar och förklaringar för allt inom "Tekniska aspekter av AI-litteracitet". Koppla rapporten till utbildning, framförallt till högre studier. Varje "Komponent" ska vara en egen huvudrubrik och varje del av "Innehåll" ska vara en egen underrubrik. Skriv utförligt och detaljerat om alltihop. Målgruppen är studenter vid universitetet.



3. Deep research

- Output
 - Arbetade i 14 minuter.
 - Ca **6500 ord** $\approx$ 13 sidor
 - Löpande **källhänvisningar**
 - Avslutande **referenslista** med klickbara länkar
 - Enkelt att exportera som **länk**, **.docx**, eller **.pdf**
- Repetera för **Praktiska** (6500 ord) och **Etiska** (13 000 ord)
- Kvalitet på texten: Ca **80%** av vad en perfekt lämpad mänsklig författare borde kunna producera.



1. Begreppslista

AI i undervisning
AI-agenter
AI-assistenter
AI-litteracitet
AI-organisatörer
AI-rättigheter
AI-rättsfall
AI-säkerhet
AI-transparent
Algoritm
Algoritmisk bias
Ansvarsbias
Artificiell generell intelligens (AGI)
Artificiell intelligens (AI)
Artificiell superintelligens (ASI)
Artificiella neurala nätverk
Attention-mekanismer
Automationsbias
Automatiserad bedömning
Autonoma AI-agenter
Autonoma AI-system
Bayesianska nätverk
Beslutsträd
Bias
Black box AI
Chatbotter
Data
Databias
Datakvalitet
Djupinlärning (Deep Learning)
Existerande risker
Explainable AI
Federerad inlärning (Federated Learning)
Few-shot learning
Förstärkningsbias
Förstärkningsinlärning (Reinforcement Learning)
Generativ AI
Grundmodeller (Foundation Models)
Hybrider
Intelligenta undervisningssystem
Interaktionsbias
Klassificering
Kunskapsrepresentation
Learning Analytics
Maskininlärning
Meta learning
Molekyl-situationer
Multimodal AI
Naturlig språkbehandling (Natural Language Processing)
Neurala nätverk (Neural Networks)
Objekt igenkänning
Övervakat lärande
Prediktion
Prompt Engineering
Regression
Retrieval-Augmented Generation (RAG)
Risker med AI
Robotik
Räddning
Self-supervised learning
Semantiska nätverk
Singularitet
Snäv AI (Artificial Narrow Intelligence)
Spolagent
Stora språkmodeller (Large language models)
Subsymbolisk AI
Superintelligens
Symbolisk AI
Sökning
Transfer learning
Transformer-arkitektur
Uelämningsbias
Vision
Zero-shot learning
Övervakat lärande



2. Deep research

Du är expert begrepp och terminologi inom AI, men 20 års erfarenhet av att forska på AI och undervisa om AI. Jag är en universitetslektor i Sverige som undervisar om AI, med särskilt fokus på AI för lärare i skola och högre studier. Mina studenter ska känna till centrala begrepp inom AI.

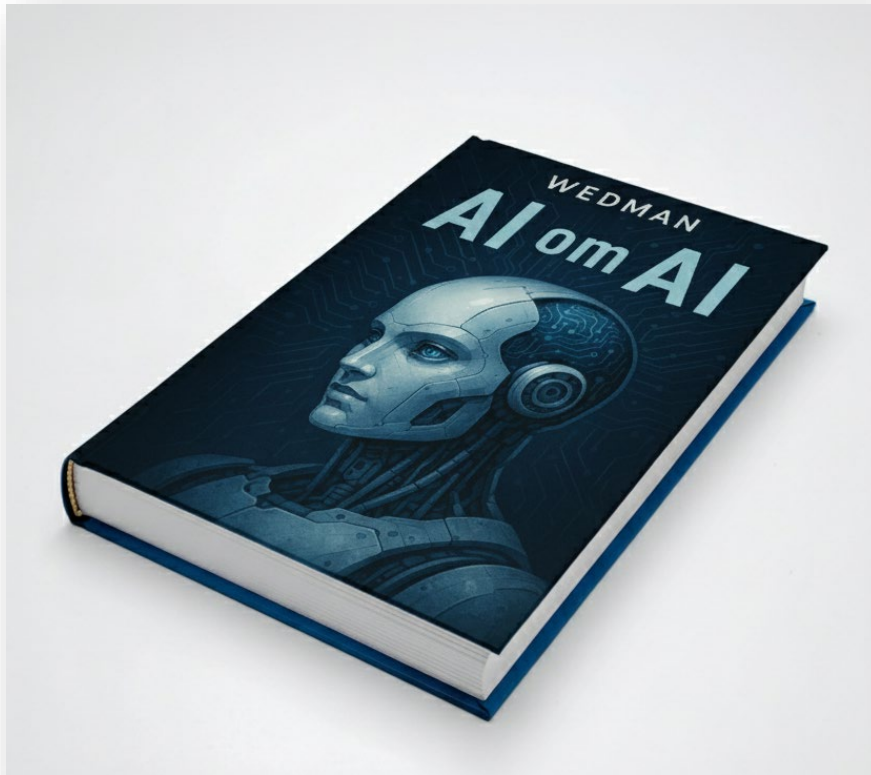
Du ska skapa en ordlista ("glossary") över centrala AI-begrepp genom att förklara vart och ett av följande begrepp utförligt. Inkludera källor och referenser. Vetenskapliga artiklar, rapporter och böcker är att föredra men även andra källor kan användas som stöd för att förklara begreppen. Begreppen du ska förklara är följande:

- AI i undervisning
- AI-agenter
- AI-assistenter ↓
- AI-litteracitet



Slutprodukterna

- Ca **26 000** ord lång text/bok + ca **29 000** ord lång Begreppslista



Bias

Definition: **Snedvridning** eller systematisk partiskhet i data eller beslut, som gör att vissa utfall avviker från det neutrala eller rättvisa – ofta så att en grupp **gynnas eller missgynnas** oproportionerligt ³³.

Beskrivning: Inom AI är *bias* ett paraplybegrepp för olika felbenägenheter som kan uppstå. Bias kan finnas i själva **datan** (t.ex. att en träningsdatamängd inte är representativ för verkligheten) eller i **modellen/algoritmen** (t.ex. att designen gör att vissa mönster förstärks på ett oönskat sätt). Resultatet är att AI-systemets output konsekvent slår fel i en viss riktning. *Bias behöver inte alltid vara "öräddvärd"* i normativ mening – i statistik kallas det bias även om en modell konsekvent överskattar ett värde. Men i AI-sammanhang fokuserar man ofta på bias som ger **diskriminerande effekter**, alltså att algoritmen beter sig annorlunda för olika demografiska grupper på ett icke önskvärt sätt. Bias kan introduceras på många steg: genom **insamling av data** (t.ex. över- eller underrepresentation av en grupp), genom **etikettering** (människa fördomar i träningsfacit), genom **modellval** (en viss modell kan anta förenklingar som inte håller för alla grupper), eller genom **användarinteraktion** (feedback-loopar, se **förstärkningsbias** och **interaktionsbias**). Att upptäcka bias kräver mätning – man måste analysera AI:ns utfall för olika subpopulationer. Att åtgärda bias kan innebära att justera datan (t.ex. "*de-biasing*" genom att balansera datasetet), att lägga till **rättvisaregularisering** i modellen eller att efterjustera resultaten. **Rättvisa (fairness)** är närbesläktat med bias-begreppet: målet är att minimera bias för att uppnå algoritmiskt rättvisa beslut.

Exempel: Ett AI-system för kreditbedömning visade sig ge markant lägre kreditpoäng till lånesökande från vissa postnummerområden, trots liknande ekonomisk status. Denna bias spårades till att träningsdatat innehöll historiska lån som påverkats av redlining – AI:n hade "*lärnt sig*" ett diskriminerande mönster. Banken fick ingripa och korrigera modellen och datan för att eliminera denna bias, så att bedömningen blev rättvisare.

Källor: IBM noterar att algoritmisk bias "*ofta reflekterar eller förstärker existerande socioekonomiska, rasliga och könsmässiga bias*" ³³ och kan "*leda till skadliga beslut eller åtgärder, samt underminera förtroendet för AI*" ³⁴. Enligt en översikt av Dooley (2019) uppstår bias i AI bl.a. genom "*skiev eller begränsad träningsdata, subjektiva programmeringsbeslut eller felaktig tolkning av resultat*" ⁷¹. För att hantera detta framhåller experter vikten av **AI governance**, transparens och kontinuerlig utvärdering för bias ⁷².



UMEÅ UNIVERSITY

Nyttan

Upphovsrätt?

- Kort tolkning: **AI-genererat = Ingen upphovsrätt**
 - Upphovsrätt kan enbart uppstå för verk som skapats av människor. //Digg.se (2025)
 - Myndigheten för digital förvaltnings ("Digg") tolkning av 1 kap. 1 § upphovsrättslagen: *Den som har skapat ett litterärt eller konstnärligt verk har upphovsrätt till verket*
- Men åsikter går isär: **Kanske, det beror på**
 - Fri sammanfattning av Kronér och Mårtensson Hjälmborg (2024)



AI-genererade verk – upphovsrätt eller inte?

Utvecklingen av artificiell intelligens fortsätter att skapa såväl möjligheter som utmaningar för lagstiftare och kreatörer. En särskilt komplex fråga är hur upphovsrättsligt skydd kan tillämpas på verk genererade med hjälp av AI. Siri Mårtensson Hjälmborg och Mathilda Kronér gör en internationell utblick, och tittar närmare på några av de avgöranden som hunnit komma i Tjeckien, Kina och USA - samt resonerar om vilken betydelse bedömningarna gällande vikten av mänsklig medverkan kan ha för svensk del.



Siri Mårtensson
Hjälmborg
Partner, Morris Law



Mathilda Kronér
Biträdande jurist, Morris Law

Följden

Vem som helst får göra ***vad som helst*** med texten

- Dela med andra
- Ladda upp i AI-verktyg
 - Sammanfatta (text+ljud)
 - Redigera eller bygga vidare på
 - Översätta
 - Förenkla (eller krångla till)
 - Skapa föreläsningar / lektioner
 - Konstruera prov samt instuderingsmaterial
 - Ge återkoppling
 - Med mera...



Min nytta HT25

- Började försiktigt med **160 provfrågor** till de tre kurserna.
 - Kom ihåg: Manuell handpåläggning behövs alltid
- Känslan av frihet
 - Jag får göra vad jag vill med texten
 - Studenterna får göra vad de vill med texten



Provfråga – före och efter

1. Vad innebär god generalisering för en modell?
 - a) Den presterar väl på nya, tidigare osedda data
 - b) Den minns exakta träningsfall men misslyckas på nya
 - c) Den undviker testning för att spara resurser

2. God generalisering innebär att en modell presterar väl på...
 - a) nya, okända data
 - b) gamla, kända data
 - c) exempel från träningsdata

Vad tyckte kursdeltagarna?

- Slutgiltigt svar i februari 2026
- ...men av 103 kursdeltagare som använt texterna i två månader har **0 klagomål** inkommit.





UMEÅ UNIVERSITY

Farhågor

Farhågor

- **Överanvändning**
 - AI-texter när det finns bättre mänskliga
- **Alltför olika undervisning**
 - 100 lärare i samma ämne gör 100 varianter av en text
- **Ekonomisk press från finansiärer**
 - Gratis är gott, särskilt för den som annars hade betalat
- **Färre människor vill författa läromedel**
 - Varför lägga ner ett år när AI gör 80% så bra på en halvtimme?
- **Fler fel**
 - Mindre mänsklig kontroll ger fler fel i texterna
- **Elever/Studenter tappar motivation**
 - Varför ska man behöva läsa en massa "AI-trams"?



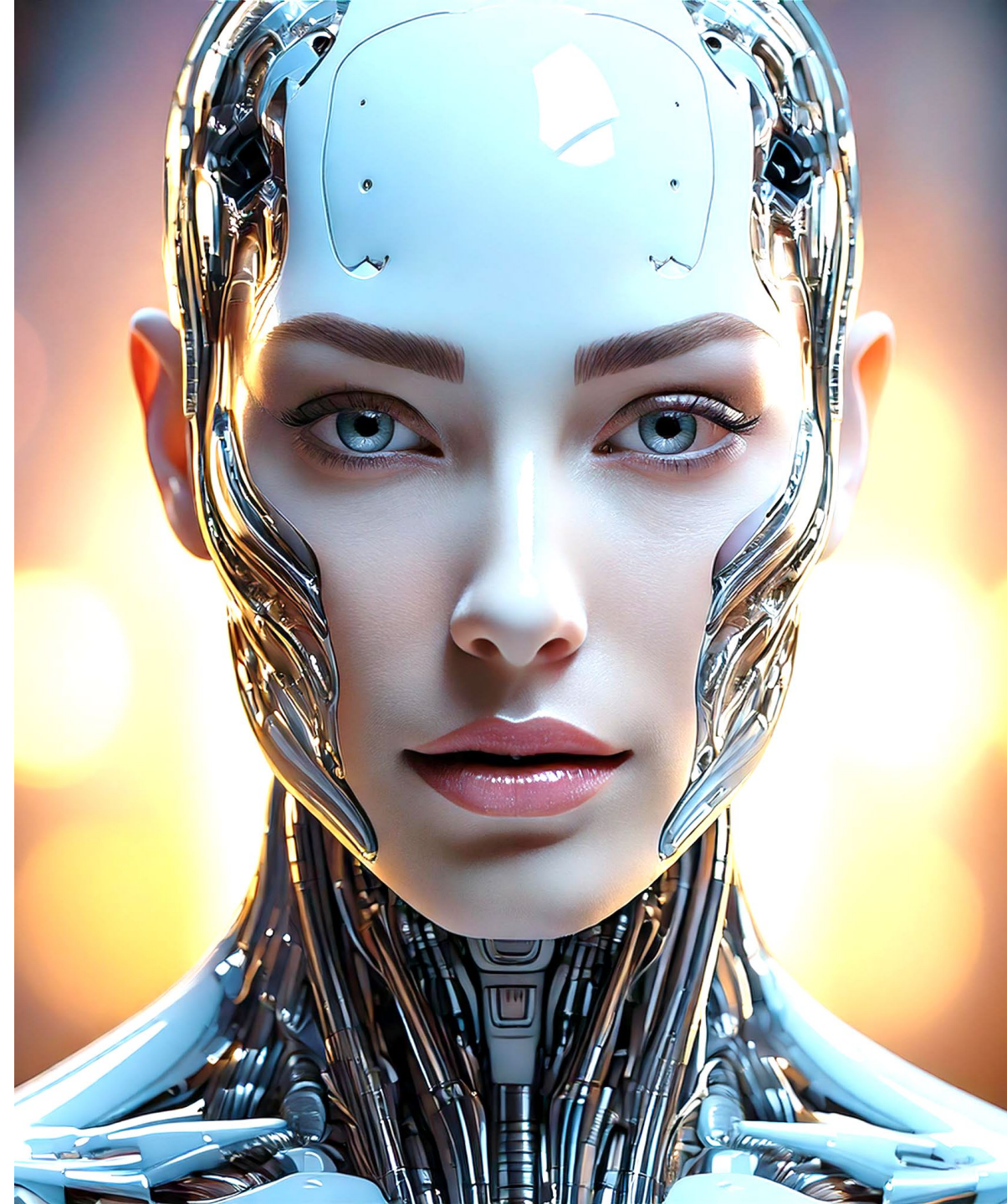


UMEÅ UNIVERSITY

Lärdomar

Lärdomar/reflektioner

- Överlag mycket trevliga texter!
- Lyxigt att ha en ordlista med 70+ begrepp man själv har valt ut och som är utförligt förklarade och exemplifierade i en kontext man själv valt.
- Någon veckas mänsklig puts hade varit önskvärt
 - "Kunskapskrav"
 - Ord som fallit bort
 - Inga bilder/figurer/tabeller





UMEÅ UNIVERSITY

Framtid

Min nytta efter HT25 (om boken gillas)

- **Instuderingsmaterial**

- Kortsvarsfrågor (testbaserat lärande)
- Mer omfattande frågor kompletta med svar och rättningsmall

- **Föreläsningar i urval**

- AI-litteracitet
- Olika typer av bias
- Noder och lager

- **Praktiska övningar**

- Lärarledda workshoppar
- Övningar för kursdeltagarna att göra på egen hand



UMEÅ UNIVERSITY

Sammanfattning

Sammanfattning och slutsats

- **27K** om AI litteracitet + **29K** ordlista
 - “Wedman”?
- Hög kvalitet men **rätt** person kan bättre
- Ca fyra timmars arbete plus lite väntan
 - Och ca fyra timmar att läsa igenom
- Ingen upphovsrätt = total frihet för alla
- Risk för felanvändning
- Studenterna verkar gilla(?)
- Stort pedagogisk värde nu och framöver

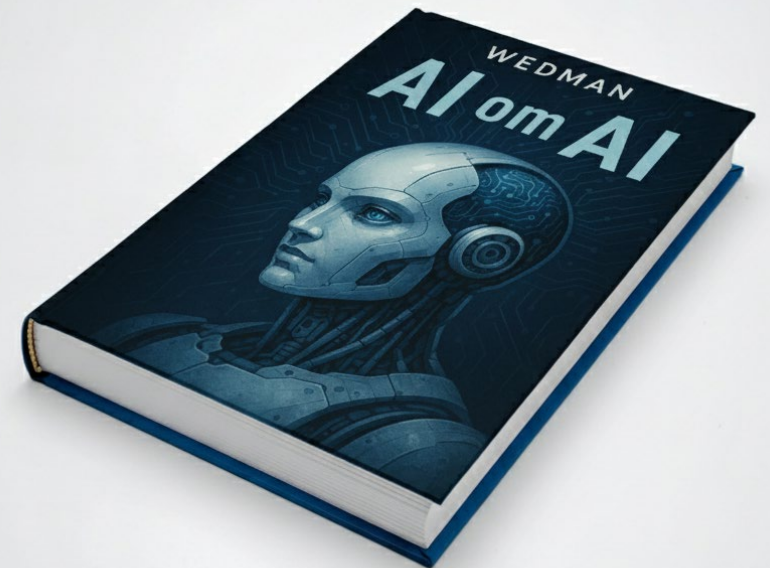
Bias

Definition: Snedvridning eller systematisk partiskhet i data eller beslut, som gör att vissa utfall avviker från det neutrala eller rättvisa – ofta så att en grupp **gynnas eller missgynnas** oproportionerligt ³³.

Beskrivning: Inom AI är *bias* ett problem som finns i själva **datan** (t.ex. att en **modell** eller **algoritmen** (t.ex. att den är att AI-systemets output konsekvent har en normativ mening – i statistik kallas detta **AI-sammanhang** fokuserar man på att identifiera och korrigera sådana **bias** som inte följer de många steg: genom **insamling** och **etikettering** (människliga fördomar och förenklingar som inte håller för **förstärkningsbias** och **interaktion**). AI:ns utfall för olika subpopulationer genom att balansera datasetet), och resultaten. **Rättvisa (fairness)** är en uppgift att uppnå algoritmsamt rättvisa beslut.

Exempel: Ett AI-system för kreditbetyg i vissa postnummerområden, trots att innehåll historiska lån som påverkar fick ingripa och korrigera modellens rättvisare.

Källor: IBM noterar att algoritmer **rasliga och könsmissiga bias** ³³ förteordet för AI ³⁴. Enligt en begränsad träningsdata, subjektiva hantera detta framhåller experter för bias ⁷².





UMEÅ UNIVERSITY

TACK!

jonathan.wedman@umu.se

Fil. dr i Beteendevetenskapliga mätningar