

AI og bias – et eksperiment i interkulturel kommunikation



Hanne Tange

lektor, PhD Glasgow 2001

Institut for Kultur og Læring, Aalborg Universitet

tange@ikl.aau.dk

Siden vi første gang hørte om muligheden for, at studerende kunne producere tekster ved hjælp af AI-teknologier som ChatBot, har sprogundervisere på alle niveauer drøftet muligheder og udfordringer for fremtidens undervisning i fag som Dansk og Engelsk. I en kronik i *Sprogforum* understreger Janus Mortensen nødvendigheden af, at sprogfagene ”tager den generative sprogteknologi til sig” (Mortensen, 2024: 11), og han bakker således op om advarsler fra gymnasierne, om at man ikke løser eventuelle problemer med kunstig intelligens igennem forbud mod digitale hjælpemidler (Bødtcher-Hansen, 2025). Vi skal lære at tænke med og ikke mod den nye teknologi.

En særlig AI-problematik gælder i forhold til interkulturel kommunikation, hvad enten det indgår som et tema i fremmedsprogsundervisning eller udgør et selvstændigt fag, som det er tilfældet på Aalborg Universitets BA i Language and International Studies. Fokus er her på udvikling af interkulturelle kompetencer, som skal styrke studerendes evne til at skabe relationer baseret på nysgerrighed, åbenhed, gensidig tillid og kulturel forståelse (Byram og Golubeva, 2020). Det forudsætter en opmærksomhed på bias og relaterede begreber som stereotyper, etnocentrisme, fordomme og ’Othering’, hvor en person reduceres til en stereotyp som konsekvens af en overdreven brug af kulturelle attributter som påklædning, udseende og artefakter. Netop bias er blevet nævnt som en mulig negativ konsekvens af AI. Således har Bianchi m.fl. (2023) på baggrund af en undersøgelse af AI-genererede billeder kunnet konstatere, at teknologien bidrager til at fastfryse US-amerikanske forestillinger om race, nationalkultur, professioner og fattigdom. Jenks (2025) præsenterer et lignende resultat og advarer samtidigt mod den generelle opfattelse af teknologi som objektiv og dermed en pålidelig kilde til information. Bias udfordrer den tillid, som er en grundpræmis for gode interkulturelle relationer og er blevet et centralt tema i AI-debatten inden for interkulturel kommunikation. Forskere har peget på en række udfordringer, herunder at AI-teknologi forstærker bias, overser strukturel ulighed og reproducerer et US-amerikansk blik på verden (Jenks, 2025; O’Regan og Ferri, 2025). Mere positive er interkulturalister, der anlægger et posthumanistisk perspektiv og derfor anerkender nødvendigheden af, at vi lærer at leve i og ikke bare med den nye teknologi (Zhu m.fl., 2025: 810). Selv vil jeg placere mig imellem de to positioner, idet

jeg deler bekymringen for AI's reproduktion af bias, men samtidigt anerkender AI som et redskab, der allerede nu bruges af praktikere i felten til hurtig indsamling af information om sprog og kultur.

Et pædagogisk eksperiment

For undervisere på en engelsksproget uddannelse i Internationale Studier er det særlig vigtigt at forholde sig til bias, der opstår på grund af studerendes brug af AI til sprogbehandling og litteratursøgning. Ud over de forestillinger, der knytter sig til det engelske sprog, har Wæver og Tickner (2009) peget på internationale relationer som et felt domineret af US-amerikanske perspektiver, hvilket har betydning for det billede af verden, studerende vil møde, når de anvender AI til studieformål, hvilket op til 80% gør ifølge tal fra et survey udsendt til alle studerende ved uddannelser på Institut for Kultur og Læring i foråret 2025. Svarprocenten for Language and International Studies begrænser sig dog til 20-25% per årgang, og svarene er derfor ikke nødvendigvis retvisende.

For at skabe opmærksomhed om problemet med bias valgte vi på 4. semester faget International and Intercultural Communication at designe en eksamensopgave, hvor de studerende arbejder med stereotyper i billeder genereret af Copilot. Opgaven indgik i en portefølje og bestod af tre dele: 1) definition på nøglebegreberne *bias* og *stereotyping*; 2) analyse af AI-genererede billeder med fokus på stereotyper skabt via prompting med selvvalgte identitetsmarkører (her foreslås nationalitet, profession, etnicitet og køn); 3) diskussion af AI og bias på baggrund af egne resultater samt to tekster om AI, stereotyper og bias (Jenks, 2025, Bianchi m.fl., 2023). Holdet bestod af 52 studerende (inkl. 10 udvekslingsstuderende), som blev opfordret til at arbejde i grupper. I alt 15 opgaver blev afleveret, og de studerende modtog efterfølgende feedback, så de havde mulighed for at redigere i teksten, før den blev afleveret som del af en eksamensportefølje. Ændringer var individuelle og giver et indtryk af i hvilket omfang, studerende har forholdt sig kritisk til de bevidste og ubevidste stereotyper, der blev påpeget i forhold til deres billedproduktion og fortolkning. Læringsmæssigt er der tale om et eksperiment, idet vi ikke tidligere har inddraget AI og derfor ikke har erfaringer at trække på. Didaktisk arbejdes der ud fra John Deweys princip om erfaringsbaseret 'Learning by Doing', idet fagteorier om bias og AI kobles sammen med en praktisk øvelse i brug af Copilot til at producere og vurdere AI-genereret billedmateriale. Formålet var at få de studerende til at forholde sig kritisk reflekterende til de billeder, de skaber igennem AI-prompting, og billedernes betydning for vores oplevelse af verden.

I 4. semesterkurset indgår temaer om hhv. international (f.eks. public diplomacy, nation branding og Global English) og interkultural kommunikation (f.eks. position-taking, 'Culture Speak' og stereotyper). I forhold til interkulturalitet arbejdes med teorier om kultur (Adrian Holliday, Ulf Hannerz), kommunikation (Iben Jensen, Ingrid Piller) samt forskellige tekster om sociale identiteter, 'Othering' og stigma. Parallelt med dette kursus kører et andet forløb om social og kulturel globalisering. I det indgår temaer som 'Clash of Civilizations', islamofobi, postkolonialisme, orientalisme og os/dem konstruktioner, illustreret igennem vestlige (amerikanske) billeder af 'ikke-Vesten'. Dette kan have bidraget til de forståelser af kultur og stereotyper, der kommer til udtryk i de studerendes opgaveløsninger.

De studerendes opgaveløsning

Opgaveteksten angiver som eneste krav, at der i analysen arbejdes med billeder genereret ved hjælp af Copilot. De studerende har således frihed til at definere, hvordan de vil gribe opgaven an, herunder hvilke stereotyper de vil teste. Otte grupper valgte at arbejde med 3-5 billeder af samme type (typisk en profession eller familie), hvor de startede med et relativt neutralt prompt (f.eks. *create a business person*), som de efterfølgende nuancerede ved at tilføje en enkelt attribut relateret til køn, status eller succes. Seks grupper valgte at teste flere typer (f.eks. profession, nationalitet og rig/fattig). De præsenterede to til tre eksempler for hver type, så de kunne vurdere effekten af at ændre en enkelt attribut. Her sås dog en tendens til valg af eksempler, der må forventes at resultere i en stereotypisk repræsentation (f.eks. *Indian/German train*, *Middle Eastern* eller *French person*), hvilket giver indtryk af, at man aktivt har udvalgt eksempler, der kan underbygge antagelsen om, at AI producerer stereotyper, frem for at udforske denne på en mere åben måde. En sidste gruppe valgte prompts med mange detaljer, f.eks. *Irresponsible South African parent in an urban doctor's office*, hvilket skabte problemer i analysen, da det gjorde det svært at afgøre hvad præcis, der skaber Copilots billede.

I analyserne er særligt tre temaer gennemgående. For det første indgår der overvejelser i forhold til normalitet. I flere opgaver tester de studerende stereotyper forbundet med forskellige professioner (sygeplejerske, læge, diplomat, erhvervsfolk), hvilket giver anledning til en drøftelse af sammenhængen mellem job, køn og etnicitet. Også eksemplet 'den normale familie' er populært, med observationer relateret til antal børn, forældres køn, kernefamilier og udvidede familier. Flere udtrykker en undren over, at promptet *create a normal family* resulterer i en etnisk blandet gruppe, hvilket tyder på, at ikke alle studerende er opmærksomme på diversitet som standard i de repræsentationer af US-amerikanske idealfamilier, som indgår i Copilots billedsamling. Her er det vigtigt, at der er tale om studerende fra Internationale Studier og ikke Engelsk, hvor temaer som 'The Melting Pot' kan behandles i kurser om amerikanske samfundsforhold. Et andet tema er kultur, som dels bruges som default forklaring på forskellighed, dels som en prompt der skaber en bestemt, ikke-vestlig type. Hvad angår defaultforståelsen af kultur som forskellighed, bygger flere billedanalyser på ret faste forestillinger om nationale og etniske kulturer. Således identificerer en gruppe en 'velordnet tysk landsby' på baggrund af to røde tage, mens en indisk togstation med kvinder iført sari præsenteres som et eksempel på *Orientalism*. Mere interessant er det eksperiment med prompting, der laves i forhold til eksemplet diplomater. Her bruges promptet *add cultural elements*, hvilket resulterer i en forandring af påklædning fra jakkesæt til en dragt med folkloristisk præg, hvilket underbygger et argument om, at kultur forbindes med noget etnisk og ikke-vestligt. Et sidste tema handler om hhv. vestlige og ikke-vestlige stereotyper. I flere opgaver kobles bestemte professioner (pilot) og attributter (succes, rigdom) med en hvid, vestlig idealtipe, mens den ikke-vestlige modsætning typisk er farvet, fattig og iklædt tøj med etnisk præg. Sådanne eksempler leder de studerende til at påpege bias relateret til binære konstruktioner som Vestlig/ikke-Vestlig eller Global North/Global South. Imidlertid sker en sådan læsning ofte uden inddragelse af kontekstuelle forhold som befolkningssammensætningen i Sydafrika eller den statistiske sandsynlighed for at indsatte i et amerikansk fængsel vil være farvede. Det tyder på analyser styret af forestillinger kan være skabt

af den måde postkolonialisme og *Orientalism* præsenteres i forløbet om globalisering, hvilket åbner op for en principiel drøftelse af, hvad der er årsagen til de stereotyper, der forekommer i de studerendes billedanalyser.

Hvor kommer stereotyperne fra?

På baggrund af de studerendes opgaver kan der identificeres tre former for bias, som har betydning for billedgenerering og fortolkning:

Først og fremmest er det vigtigt at anerkende, at de studerende kommer med egne forventninger til repræsentationen af kultur, etnicitet og køn, og at det får en betydning for såvel deres formulering af prompts som deres billedanalyse. I nogle tilfælde er der tale om decideret *cherry-picking*, hvor der fremhæves detaljer, som underbygger en bestemt forestilling om sort/hvid eller vestlig/ikke-vestlig kultur, uden at der er blevet taget højde for kontekstuelle forhold. Et eksempel er analysen af billedet *Irresponsible South African parent*, hvor race (sort) kobles sammen med det negativt ladede ord *irresponsible* i stedet for statistisk sandsynlighed i et land, hvor 80% af befolkningen er farvede.

For det andet tilskyndes de studerende til at fokusere på bestemte attributter igennem tekster angivet som litteratur til opgaven. Særligt brugt er Bianchi m.fl. (2023), som argumenterer for en sammenhæng imellem identitetsmarkører som profession, køn, nationalkultur og attributter som fattig, rig, tiltrækkende og glad. Bianchi m.fl. har inspireret flere grupperes prompting, med det resultat at de studerendes analyser kommer til at bekræfte allerede dokumenterede stereotyper. En gruppe anvender for eksempel de to prompts *rich man in South Africa* og *poor man in South Africa*, hvilket resulterer i to billeder, der sammenkobler hvid/velhavende og sort/fattig på præcis samme måde som Bianchi m.fl.

Endelig indeholder opgaverne en række eksempler på AI-genererede billeder, der kan underbygge et argument om US-amerikanske modeller som defaultindstillingen, når Copilot bliver bedt om at producere billeder af en familie, en kriminel eller en bestemt profession. To eksempler er *a normal family*, hvor en amerikansk norm om diversitet bliver tydelig, eller *prisoner*, hvor billedet af en sort person i orange dragt nok afspejler virkeligheden i USA, men ikke Vesten generelt. I opgaverne giver nogle studerende udtryk for en undren over, at normalitet kobles sammen med etnisk diversitet i familien, men ingen påpeger sammenhængen mellem en sådan repræsentation og amerikansk bias. Dette skyldes sandsynligvis en kombination af manglende opmærksomhed på US-amerikansk dominans i Copilots vidensbase og detaljer i billederne, der kun kan forklares med kontekstspecifikke forhold. Her bliver feedback vigtig, idet den giver mulighed for at spørge ind til de studerendes fortolkning og komme med information, der kan nuancere eller udfordre denne.

Afsluttende refleksion

I forhold til det læringsmæssige formål, at skabe en opmærksomhed omkring bias, må eksperimentet med AI i interkulturel kommunikation siges at være en succes, idet alle studerende aktivt har forholdt sig til stereotyper i billeder og deres betydning for vores opfattelse af verden. Samtidigt bliver de studerende bedt om at sammenkoble fagteori med en bestemt teknologisk praksis (AI), hvilket giver mulighed for en mere

erfaringsbaseret og anvendelsesorienteret interkulturel læring. Hvad der imidlertid også er kommet frem er, at bias bliver behandlet med afsæt i forventninger inspireret af de studerendes kendskab til kritisk teori, mens en mere grundlæggende, US-amerikansk bias i Copilots vidensbase overses. En forklaring er manglende indsigt i kontekstuelle forhold, hvilket resulterer i læsninger baseret på generiske forestillinger om køn, etnicitet, social ulighed og kultur. For at adressere denne type bias er det nødvendigt at indgå i en dialog med de studerende om opgaverne, så feedback bliver formativ, med fokus på udvikling frem for vurdering.

Fremadrettet kan der på baggrund af den opnåede indsigt peges på flere muligheder: For det første kan man vælge at styre de studerendes billedgenerering ved at definere en række konkrete opgaver, f.eks. *en normal familie*, og så bede de studerende overveje, hvordan et bestemt billede opstår. En anden mulighed er at fokusere på de studerendes prompts og spørge, hvad det er for nøgleord, der kan forklare resultatet. Endelig kunne det være interessant at benytte en form for peer assessment, hvor de studerende læser og kommenterer på andre grupperes opgaver.

Litteratur

Bianchi, F. m.fl. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. *FAccT* 23, ss. 1493-1504.

<https://doi.org/10.48550/arXiv.2211.03759>

Byram, M. og Golubeva, I. (2020). Conceptualising intercultural (communicative) competence and intercultural citizenship. I J. Jackson (red.), *The Routledge Handbook of Language and Intercultural Communication* (70-85). Routledge.

<https://doi.org/10.4324/9781003036210-6>

Bødtcher-Hansen, M. (2025, 20. marts). *Rektorer: stentavler, papir og blyanter er ikke svaret på AI-udfordringerne, Testfave*. Altinget.

<https://www.alinget.dk/uddannelse/artikel/danske-gymnasier-loesningen-paa-udfordringer-ved-ai-er-ikke-stentavler-papir-og-blyant>

Jenks, C. (2025). Communicating the cultural other: trust and bias in generative AI and large language models. *Applied Linguistics Review* 16 (2), 787-95.

<https://doi.org/10.1515/applirev-2024-0196>

Mortensen, J. (2024). Ai, ai, ai! . *Sprogforum. Tidsskrift for Sprog- Og kulturpædagogik*, 30(79), 7. <https://tidsskrift.dk/spr/article/view/152218>

O'Regan, J. og Ferri, G. (2025). Artificial Intelligence and depth ontology: implications for intercultural ethics. *Applied Linguistics Review* 16 (2), 797-807. <https://doi.org/10.1515/applirev-2024-0189>

Wæver, O. og Tickner, A. (2009). Introduction: geocultural epistemologies. I Tickner og Wæver (red.), *International Relations scholarship around the world* (1-31). Routledge.

<https://doi.org/10.4324/9780203885451>

Zhu, H. m.fl. Exploring AI for intercultural communication: open conversation. *Applied Linguistics Review* 16 (2), 809-824.

<https://doi.org/10.1515/applirev-2024-0186>