AI OG GLOBAL ULIGHED

En undersøgelse af, hvordan store sprogmodeller (LLMs) bidrager til global ulighed og rejser etiske dilemmaer med særligt fokus på Sydafrika.

JOACHIM SMED-PETERSEN & KATRINE FRIBORG HEITMANN

TITELBLAD

SKREVET AF: JOACHIM SMED-PETERSEN (20231820) & KATRINE FRIBORG HEITMANN (20231822)

UDDANNELSE: KOMMUNIKATION, CAND.MAG. (MA), KØBENHAVN

VEJLEDER: DAVID BUDTZ PEDERSEN

UDGIVET: 2. JUNI 2025

ANTAL TEGN: 247.599

ANTAL NORMALSIDER: 109

ANTAL BILAG: 14

AD NYE VEJE

AALBORG UNIVERSITET

1. ABSTRACT

Abstract

Amidst the rapid global expansion of AI, concerns are growing about how technological advancements, primarily driven by actors in the minority world, exacerbate existing inequalities, particularly in regions with limited digital infrastructure. This thesis investigates whether the development and training of Large Language Models (LLMs) contribute to global inequality, and explores the ethical dilemmas in attempts to make AI more inclusive and just.

The literature review shows how dominant LLMs often overlook or misrepresent linguistic and cultural diversity by relying heavily on Western English-language datasets, thereby reinforcing postcolonial patterns of exclusion and dominance.

The theoretical basis of the paper combines a normative approach to global justice with postcolonial and ethical perspectives. This includes Amartya Sen's capability theory, Floridi and Cowls' data ethics framework, Rachel Adams' terminology and concept of the 'AI Empire', and Couldry and Mejias' theory of data colonialism. Critical discourse analysis, as developed by Norman Fairclough, is also employed to examine the intersection of language, knowledge, and power in the context of AI, and the reproduction of historical patterns of inequality.

Methodologically, the study is based on interviews with AI experts from Denmark and South Africa as well as a comparison of the EU's AI Act and the AU's AI Strategy.

The analysis examines how training data and language representation in LLMs reproduce colonial structures, and how participatory initiatives, such as the South African non-profit Masakhane, attempt to counteract this by building datasets for African languages. By focusing on South Africa, the study explores how AI manifests asymmetries in a country shaped by deep-rooted inequality and limited digital access.
Generative AI is further situated within a global power context, analyzing how regulatory approaches and geopolitical interests influence the trajectory of ethical AI.

In conclusion, this study unravels the complexities of AI development risking reinforcing global and local inequalities, unless underrepresented perspectives, languages, and knowledge systems are actively included. Realizing AI's potential for equity requires rethinking technological power structures, aligning digital infrastructure development with social needs, and ensuring genuine inclusion in AI governance. Only if historically marginalized communities have real agency in shaping AI, from data collection to design and regulation, can generative AI help reduce - rather than reinforce - global inequality.

Indholdsfortegnelse

2. INDLEDNING

1. Indledende problemfelt

"AI has the potential to solve all our problems –
or make them infinitely worse."
(Benjamin Rosman: 2, 18)

Globalt set er der voksende bekymring for, om kunstig intelligens vil forstærke og skabe nye former for ulighed. Det skyldes blandt andet bias i algoritmer og manglende repræsentation og inklusion i datasæt, der bruges til at træne AI systemer og modeller (Helm et al. 2024: 7). I Sydafrika for eksempel, et land med en rig og kompleks historie, manifesterer disse teknologiske uligheder sig særligt i sproglige og kulturelle nuancer, der ofte mangler eller misforstås i de dominerende AI-modeller, udviklet i Vesten.
"At its core, the coming wave is a story of the proliferation of power." (Suleyman 2023: 102). Mustafa Suleyman, CEO hos Microsoft AI og forfatter til bogen *The Coming Wave* (2023), påpeger, hvordan teknologi i dag er blevet en af verdens mest magtfulde strategiske ressourcer, og hvordan de store magtkampe i det 21. århundrede i stigende grad centrerer sig om teknologisk suverænitet og kapløb om at kontrollere teknologisk dominans. I denne udvikling opfattes tech-virksomheder og universiteter ikke længere som neutrale interessenter, men som vigtige nationale strategiske aktører (Suleyman 2023: 124). AI-markedet forventes at overstige 240 milliarder dollars i 2025 og nå op over 826 milliarder dollars i 2030, hvilket vidner om den hastige og omfattende vækst på området (AscendixTech 2025). Den teknologiske dominans kommer desuden til udtryk ved, at verdens fem største AI-relaterede virksomheder målt på markedsværdi – Microsoft, NVIDIA, Apple, Google og Meta – alle er amerikanske og hver især har en markedsværdi, der overstiger 1600 trillioner dollars (CompaniesMarketCap, u.å.). Denne koncentration af ejerskab og kontrol illustrerer, hvordan den globale teknologiske magt i stigende grad samles i hænderne på en lille gruppe vestlige aktører, først og fremmest i USA.
AI-forsker og direktør for Global Center on AI Governance, samt forfatter til *The New Empire of AI* (2025), Rachel Adams, pointerer at kampen mellem få lande om at være først med de nyeste AI-teknologier sætter sit aftryk på resten af verden. Talrige lovgivnings- og politiske forslag er blevet indledt, selv i lande med manglende digital infrastruktur, hvor navne på initiativer som 'AI For Good' og 'AI For All' fremstår vage og virkningsløse, når man betragter den større, globale digitale udelukkelse af lande i udviklingen af kunstig intelligens. Hun påtaler hvordan udviklingen af AI forstærker kløften mellem lande: "As the

AI industries boom in the West and in the Far East, the gap between countries grows wider and wider, at a speed that is only accelerating.'' (Adams 2025: 17).

AI er blevet geopolitik

Den internationale AI-udvikling har styrket en etisk egoisme, hvor det i stigende grad anses som legitimt at prioritere nationale interesser frem for at sikre en retfærdig global fordeling af teknologiens muligheder. Den nuværende amerikanske regering udtrykker denne egoisme som protektionisme, og arbejder på at skabe så få begrænsninger som muligt for at opnå en førerposition i AI-kapløbet. Lande, der ikke aktivt indgår i konkurrencen om teknologisk indpas, risikerer dermed at blive yderligere marginaliseret. Selvom FN's AI Advisory Body (United Nations AI Advisory Body, 2024) taler for en global fordeling af AI's muligheder og fordele ud fra idealer om bæredygtig udvikling og respekt for nationernes selvbestemmelse, ser vi også, hvordan stormagter som USA har koblet adgang til AI-teknologi til strategiske hensyn (University of Oxford 2025). I stedet for at fremme global lighed bruges adgangen til AI-teknologi som et geopolitisk værktøj til at fastholde eller udvide magtpositioner, for at styrke egen indflydelse og mindske rivalers adgang til nye markeder.

Den 11. februar 2025 til AI-topmødet i Paris, blev der kastet lys over den spænding, der opstod mellem udviklingen og implementeringen af kunstig intelligens. Mødet blev præget af, at USA og Storbritannien nægtede at underskrive Paris-erklæringen, en beslutning, der ifølge eksperter fra Oxford University både afspejler en strategisk alliance mellem USA og Storbritannien og en grundlæggende forskel i tilgangen til AI-regulering, hvor især Storbritannien foretrækker en mindre forsigtig tilgang end EU (University of Oxford 2025). En erklæring, der lagde vægt på en inkluderende og bæredygtig tilgang til kunstig intelligens, med fokus på menneskerettigheder, bæredygtighed, inklusion og arbejdstagerrettigheder, i modsætning til tidligere topmøders fokus på eksistentielle AI-trusler og sikkerhed. Den manglende tilslutning fra førende AI-nationer som rejser spørgsmål ved deres vilje til at indgå i et globalt samarbejde om etisk og inkluderende teknologisk udvikling, da de ønsker at undgå, at AI-branchen præges af ideologisk bias og regulering.

Suleyman (2023) fremhæver, at måden, hvorpå vi 'tøjler' teknologien, bør være det første grundlæggende skridt for at sikre de nødvendige forudsætninger for menneskehedens overlevelse. Dette indebærer både regulering, øget teknologisk sikkerhed, nye former for styring og ejerskabsmodeller samt nye mekanismer for ansvarlighed og gennemsigtighed. Som Suleyman inddæmmer, kan teknologiske risici ikke stå alene som løsning på

teknologiens udfordringer, men bør snarere betragtes som et nødvendigt første skridt, der skaber grundlaget for fremtidige løsninger (Suleyman 2023: 37).

Hvordan kan vi sikre en retfærdig brug af teknologien, når det globale 'vi' primært defineres af vestlige lande, mens andre dele af verden holdes udenfor? Rachel Adams fremhæver det internationale initiativ, The Global Partnership on AI, som samler regeringer fra hele verden med det formål at udvikle ansvarlig AI. Hun påpeger dog, at initiativet i høj grad mangler deltagelse fra lande i det globale syd. På trods af bestræbelser på at inkludere ikke-vestlige lande, repræsenterer kun fem ud af de nuværende 28 medlemslande uden for den vestlige verden, og kun ét land, Senegal, kommer fra det afrikanske kontinent (Adams 2025: 46).

Ansvarlig AI

Fælles for både EU's AI-forordning og den Afrikanske Unions AI-strategi er et ønske om at fremme retfærdighed og sikre ansvarlig brug af AI. Mens EU har indført bindende krav til ansvarlighed og gennemsigtighed i anvendelsen af AI-systemer, med udgangspunkt i vestlige idealer og europæiske værdier, har den Afrikanske Union valgt en mere risikobaseret tilgang uden en samlet, kontinentalt bindende lovgivning (Global Center on AI Governance 2025: 15). Denne uensartede regulering afspejler ikke kun forskellige politiske prioriteter, men peger på en grundlæggende ubalance i reguleringsevne og institutionel kapacitet. Ydermere, viser det, hvordan lighed i teknologisk udvikling ikke er til stede og måske aldrig har været det.

På trods af bottom-up-initiativer som Masakhane-projektet er teknologien stadig præget af global ulighed. De illustrerer, hvordan græsrodsorganisationer kan styrke inklusion ved at opbygge 'NLP-ressourcer' (natural language processing), for afrikanere og af afrikanere, og dermed give underrepræsenterede sprog og samfund en stemme i AI-udviklingen (Masakhane 2025). Masakhane opstod som reaktion på, at Afrikas 2000 sprog er næsten usynlige i teknologiske systemer. Det er afgørende at give mennesker mulighed for at forske og deltage i videnskabelige diskussioner på deres eget sprog. Uden mulighed for at deltage på egne modersmål risikerer store befolkningsgrupper at blive udelukket fra vidensproduktion (Wild, 2021). Når nationer står uden for deltagelse og samarbejde, og når præcise termer og koncepter ikke forstås fuldt ud, kan det skabe en følelse af manglende indflydelse. Derfor er det centralt at sikre, at sprog, der tales af en stor del af befolkningen, får et digitalt liv, ellers risikerer de simpelthen at forsvinde, afdækker forskningsjournalist fra

Sydafrika, Sarah Wild: “These are languages [people] speak. These are languages they use every day, and they live with and see the reality that in x number of years, their language might be dead because there is no digital footprint.” (Wild 2021).

Ifølge Afrikas AI-tænketank, Global Center on AI Governance, spiller åbne og lokale AI-initiativer som organisationen Masakhane en vigtig rolle i at skabe mere tilgængelig og lokalt relevant teknologi, men de er fortsat udfordret af ressourcemangel og strukturelle barrierer (Global Center on AI Governance 2025: 16). Det globale syd, herunder store dele af Afrika, står med langt færre muligheder for at deltage i den teknologiske udvikling. Historiske forhold som kolonialisme, økonomisk udbytning og systematisk underinvestering i infrastruktur og uddannelse har skabt strukturelle uligheder, som former adgangen til og kontrollen over nye teknologier. Derfor starter lande i majoritetsverden ikke fra samme udgangspunkt som vestlige lande, når det gælder udvikling, udnyttelse og styring af kunstig intelligens.

Som vi uddyber i afsnit 5.2 anvender vi Amartya Sens kapabilitetsteori som normativ ramme for at analysere disse ulighedsforhold. Ifølge Sen bør retfærdighed ikke alene vurderes ud fra formel lighed, men ud fra de faktiske muligheder, mennesker har for at realisere deres livsmål. I dette lys bliver det tydeligt, at teknologisk lighed ikke kan tages for givet. Mange afrikanske lande mangler de nødvendige forudsætninger for at udvikle AI-kapabiliteter, der repræsenterer deres egne sprog, kulturer og verdenssyn. Denne problemstilling fører os i specialet til at udfolde en retfærdighedsbaseret tilgang.
Med afsæt i teorier om global retfærdighed og normative rammer for udvikling og anvendelse af AI, argumenterer vi for, at teknologisk udvikling ikke blot bør vurderes ud fra, hvem der skaber innovation, men også ud fra, hvem der får mulighed for at deltage i og forme den. Dette udspringer af en forståelse af adgang til viden og kulturel selvbestemmelse som grundlæggende goder på linje med eksempelvis uddannelse og adgang til mad og vand – goder, som alle i udgangspunktet bør have adgang til. En sådan tilgang afspejler også Sens (1993) perspektiv på frihed og kapabiliteter.

Nye AI-innovationer kan gøre globale uligheder endnu større, netop fordi uddannelse, viden og kultur i stigende grad digitaliseres, og fordi det kræver enorme investeringer at udvikle og adoptere nye teknologier. Problemet er imidlertid ikke blot, at adgangen til at udvikle teknologien er ulige, men også at de sociale og materielle omkostninger ved AI-modellerne systematisk skjules. Rachel Adams fremhæver, hvordan AI-industrien bevidst skjuler de ressourcemæssige byrder, som ligger bag teknologiens udvikling og anvendelse. I stedet søger industrien at bevare illusionen om effektivitet og fremskridt. Adams skriver:

> *"The work taken to produce AI, including the waste it creates or to which it contributes, is carefully hidden by an industry that prefers to depict AI as a clever add-on to connected data systems already in existence. (...) This fallacy is bound up with an old colonial dream: the colonizer enjoys an easy life, remaining blind to the fact that his comfort is predicated on oppressing the colonized through appropriation of their labour and expropriation of their land and resources."* (Adams 2025: 83).

Denne fortsatte usynliggørelse af ulighed forbinder nutidens teknologiske udvikling med tidligere koloniale mønstre, hvor fremskridt for nogle nationer byggede på udnyttelse og eksklusion af andre. Når store tech-virksomheder fra USA, Kina og til dels Europa dominerer udviklingen af AI, uden at sikre rum for lokale eller alternative perspektiver, risikerer vi en ny form for datakolonialisme (jf. afsnit 5.4). Kunstig intelligens skaber ikke i sig selv global ulighed, men er snarere et spejl, der forstærker de bias, vi ser kodet ind i systemerne. En konsekvens af dette er den kapitalistiske udnyttelse af mindre ressourcestærke regioner til fordel for nogle få globale tech-virksomheder. Som vores respondent Frej Klem Thomsen, seniorkonsulent ved Nationalt Center for Etik, forklarer, bærer Vesten et særligt ansvar: "Hvis man skal snakke global retfærdighed i lidt bredere termer så mener jeg at velstående lande har meget store forpligtelser til at hjælpe mindre velstående lande, også langt større end de løfter lige nu." (Thomsen: 8, 41-43). Spørgsmålet er derfor ikke kun, hvordan vi udvikler teknologien, men hvem der har reel mulighed for at medudvikle den, og på hvilke vilkår.

Inklusion og repræsentation i store sprogmodeller

I dette afsnit har vi valgt at zoome ind på, hvordan store sprogmodeller (LLM) påvirker global ulighed gennem deres inklusion eller eksklusion af sprog. En betydelig del af verdens befolkning, estimeret til en tredjedel, er digitalt uforbundet. Det betyder, at deres erfaringer og verdenssyn i vidt omfang er fraværende i de data, der anvendes til at træne AI-systemer. Særligt afrikanske sprog er i dag kraftigt underrepræsenteret i store sprogmodeller, hvilket ifølge Global Center on AI Governance skaber fejl i sprogteknologi og fastholder det globale syds afhængighed af teknologier udviklet i andre dele af verden (Global Center on AI Governance 2025: 3).

Manglende sproglig og kulturel repræsentation i datasæt betyder, at AI-træning risikerer at skævvride eller overse kulturelle normer, behov og rettigheder.
Vi placerer os inden for forskningsfeltet omkring udviklingen og træningen af generativ AI,

snarere end dets implementering og anvendelse, som ofte er i fokus i mange vestlige kontekster.
Vi undersøger, hvordan magtforhold cementeres allerede i træningsfasen, hvor bestemte verdensbilleder prioriteres, mens andre ekskluderes. Dette gør sprogmodeller til et centralt felt for at forstå, hvordan AI-teknologi muliggør og begrænser, gennem de sprog, perspektiver og værdier, der prioriteres i træningen. Det er netop her, vi har identificeret et hul i den eksisterende forskning - et område, der endnu er relativt uafdækket, og som vi mener er afgørende at belyse i relation til stigende global ulighed.
En nærmere afklaring af forskningsfeltet og dets nuværende fokusområder findes i afsnit 4.

Vi ønsker at afdække en række normative spørgsmål og undersøge, hvordan udviklingen af generativ AI potentielt kan forstærke eller mindske global ulighed. Med udgangspunkt i en sydafrikansk case og et fokus på sprogmodeller analyserer vi ikke blot tekniske forhold som datasæt og bias, men også hvordan politiske styringsinstrumenter i forskellig grad forholder sig til ansvar, inklusion og etisk udvikling af AI. Vores undersøgelse bygger på eksperters perspektiver og viden om AI's komplekse spændingsfelter, og vi er særligt optagede af, hvordan adgang til teknologi og respekt for lokale normer, sprog og verdenssyn er afgørende for en retfærdig digital fremtid.
Det er i krydsfeltet mellem strukturel ulighed og teknologisk udvikling, at vi i det følgende dykker ned i, hvordan sprogmodeller, som en af de mest udbredte og indflydelsesrige AI-teknologier, påvirker globale uligheder. Med udgangspunkt i en retfærdighedsbaseret tilgang og teorier om kapabilitet og kolonialisme undersøger vi, hvordan adgang til sproglig repræsentation i AI ikke blot handler om teknologi, men om retten til at deltage i og forme fremtidens digitaliserede samfund.

2. Problemformulering

På baggrund af det indledende problemfelt, opstiller vi den overordnede problemformulering:

Hvordan bidrager udviklingen og træningen af generativ AI til global ulighed, og hvilke etiske dilemmaer opstår i bestræbelserne på at gøre AI mere retfærdig og inkluderende?

3. BEGREBER

3. Begrebsafklaring

For at besvare specialets problemformulering er det nødvendigt at afklare en række centrale begreber. På den teknologiske side anvender vi betegnelser som 'artificial intelligence' (AI), 'generativ AI' og 'Large Language Models' (LLM'er), som danner det teknologiske grundlag for de fænomener, specialet undersøger. Samtidig beskæftiger vi os med analytiske begreber som bias og global ulighed, der er afgørende for at forstå, hvordan AI påvirker eksisterende magtforhold og skævheder på tværs af globale kontekster. Derfor præciseres og kontekstualiseres disse begreber i det følgende afsnit.

3.1 AI

Indledningsvis præciserer vi, hvordan vi forstår begrebet kunstig intelligens i denne opgave. Vi anvender både begreberne kunstig intelligens og forkortelsen AI, som i specialet bruges synonymt.

Dataetisk råd peger på, der ikke er en fuldstændig enighed om, hvornår et computersystem kan kaldes for kunstig intelligens, men en bred og letforståelig definition af kunstig intelligens er, "(...) at et computersystem er en kunstig intelligens, hvis det løser en kompleks opgave, som det ellers ville kræve menneskelig intelligens at løse." (Dataetisk Råd 2025: 4).

Mustafa Suleyman beskriver AI således: "Artificial intelligence (AI) is the science of teaching machines to learn humanlike capabilities." (Suleyman 2024: 1).

På baggrund af disse definitioner forstår vi i specialet AI som systemer, der er designet til at udføre opgaver, som normalt kræver menneskelig intelligens, såsom sprogforståelse, beslutningstagning og problemløsning. Det er en software, der ved hjælp af avancerede teknikker kan efterligne menneskelige evner og generere output såsom beslutninger, anbefalinger eller indhold med betydning for de omgivelser, teknologien indgår i. AI er derfor ikke blot en neutral teknologisk løsning, men en dynamisk kraft, der har potentiale til at indgå i, og påvirke sociale, politiske og globale kontekster.

Det er desuden relevant at præcisere, hvordan AI adskiller sig fra tidligere teknologiske revolutioner. Suleyman betegner AI som den 'kommende bølge' og fremhæver fire grundlæggende egenskaber, der gør det særligt udfordrende at kontrollere teknologien:

1. Stor asymmetrisk indflydelse, hvor nye teknologier skaber nyt pres mod ellers dominerende aktører.
2. Accelereret udvikling, en form for 'hyper-evolution'.

3. Omni-brug, idet teknologierne kan bruges til mange forskellige formål (også kaldet *general purpose technology*, som for eksempel internettet).
4. Øget autonomi, som går langt ud over, hvad tidligere teknologier har kunnet. (Suleyman 2023: 105).

3.2 Generativ AI (GenAI)

Eftersom specialet beskæftiger sig med den type af AI, som omhandler generativ AI, begrebsafklarer vi i dette afsnit hvordan det skal forstås.
Dataetisk Råd definerer generativ AI således:

> *"Generativ AI (GenAI) er en særlig type kunstig intelligens, som er kendetegnet ved, at den kan skabe originalt indhold af høj kvalitet, for eksempel tekst, billeder eller lyd. Typisk kræver det blot et simpelt input fra brugeren, et såkaldt 'prompt', at få en GenAI til at skabe indhold."* (Dataetisk Råd 2025: 2).

Eksempler på generativ AI er blandt andet Microsofts Copilot og OpenAIs ChatGPT (Dataetisk Råd 2025: 2). Generativ AI anvender maskinlæring til at generere nyt indhold. Den trænes på store datasæt, som kan bestå af eksempelvis litteratur, nyhedsartikler og andet onlineindhold, og lærer mønstre og strukturer i sproget for at kunne producere tekst, billeder eller lyd. Maskinlæring muliggør udviklingen af meget avanceret kunstig intelligens, men det kan samtidig gøre teknologien vanskelig at gennemskue for mennesker. Modellens kvalitet afhænger i høj grad af både mængden og kvaliteten af de data, den trænes på (Dataetisk Råd 2025: 5). De mest avancerede GenAI-modeller trænes på en væsentlig del af internettet, og OpenAIs GPT-4.0 er angiveligt trænet på tekster svarende til cirka 100 millioner bøger (Dataetisk Råd 2025: 6).

3.3 Large Language Model (LLM)

En central del af mange former for GenAI er en såkaldt 'stor sprogmodel' som på engelsk kaldes 'large language model' (LLM'er). Dataetisk Råd definerer LLM således:

> *"En LLM er en matematisk model, som viser hvordan sproglige udtryk statistisk hænger sammen. Modellen arbejder med et ordforråd, der ikke er hverken hele ord eller individuelle tegn, men sproglige brudstykker, der kan minde om stavelser. De kaldes for "tokens". For at generere tekst vurderer en LLM, hvad den mest*

> *velfungerende fortsættelse af en given tekst er, og tilføjer en token ad gangen, indtil den har løst den overordnede opgave."* (Dataetisk Råd 2025: 7).

Teknologi- og konsulentvirksomheden IBM (International Business Machines Corporation) definerer LLM'er som en type modeller, der er trænet på enorme datamængder, hvilket gør dem i stand til at forstå og generere naturligt sprog samt andet indhold med henblik på at løse en bred vifte af opgaver, som f.eks. oversætte, opsummere tekst og besvarer spørgsmål (IBM 2023). Derudover påpeger de, at eftersom LLM'er kontinuerligt udvikles og forbedres, er de ved at blive en central del af den digitale infrastruktur, idet de grundlæggende ændrer vores måde at interagere med teknologi og tilgå information på (Holdsworth 2023).

LLM'er er trænet med selv-superviseret maskinlæring, hvor modellen lærer at forudsige næste ord baseret på sandsynligheder i store tekstmængder. På den måde tilegner den sig sproglige og konceptuelle mønstre uden direkte menneskelig indblanding (IBM 2023). Modeller kan dog finjusteres med menneskelig feedback for at mindske fx bias og faktuelle fejl, en praksis, IBM ser som afgørende for sikker anvendelse. I dette speciale læner vi os op ad denne forståelse af LLM'ers virkemåde og forudsætninger for ansvarlig brug, der samtidig rummer store potentialer og problematiske implikationer.

3.4 Bias

Bias er et bredt begreb, der kan forstås forskelligt afhængigt af konteksten. I dette speciale beskæftiger vi os med bias i relation til sprogmodeller og ulighed, og det er derfor nødvendigt at præcisere, hvordan begrebet anvendes i denne sammenhæng.

Ifølge Den Danske Ordbog opstår bias, når der sker en skævhed i undersøgelser eller vurderinger, typisk fordi metoden er fejlbehæftet, eller fordi der sniger sig ubevidste antagelser og fordomme ind (Den Danske Ordbog 2025). Denne definition understreger, at bias kan opstå både i dataindsamling og i fortolkning som følge af metodiske fejl eller ubevidste forestillinger, hvilket gør det til et relevant begreb i arbejdet med AI og sprogmodeller.

3.5 AI bias

Vi anvender begrebet bias specifikt i relation til kunstig intelligens, hvor 'AI' bias henviser til de indlejrede data, modellerne trænes på, eller opstår som følge af udviklernes valg og antagelser. IBM definerer bias i AI-sammenhæng som: "AI bias, also called machine learning bias or algorithm bias, refers to the occurrence of biased results due to human biases that

skew the original training data or AI algorithm—leading to distorted outputs and potentially harmful outcomes." (Holdsworth 2023).
Denne definition fremhæver, hvordan menneskelige forforståelser kan videreføres i de datasæt, der anvendes til at træne AI-modeller, hvilket kan føre til forvrængede og potentielt skadelige resultater. Dette kan begrænse individers mulighed for at deltage fuldt ud i samfunds- og økonomiske aktiviteter, hvis AI-bias ikke identificeres og adresseres (Holdsworth 2023).

Selvom AI bias ikke kan elimineres fuldstændigt, fremhæver de vigtigheden af bevidst at identificere og arbejde med de bias, der har betydelige menneskelige konsekvenser. Et centralt redskab i denne proces er inddragelsen af 'human-in-the-loop' - altså menneskelig evaluering i udviklings- og beslutningsprocesser, for at øge gennemsigtigheden og sikre mere unbiased og retfærdige outputs (Holdsworth 2023).

3.6 Bias i sprogmodeller

Derudover læner vi os i specialet opad begrebet 'langauge modeling bias' fra en artikel fra *Ethics and Information Technology* (Helm et al. 2024). Forfatterne udtrykker, hvordan bias er allestedsnærværende, og ikke behøves at være skadeligt, men blot at alt viden og data er situeret og altid indlejret i og præget af den kulturelle, historiske, politiske og økonomiske kontekst. Dette er sprogmodeller naturligvis også påvirket af, og ikke kun i selve sprogbrugen, men også i selve designet af teknologiske redskaber. Forfatterne peger derfor på, at stræben efter fuldstændig objektivitet i teknologier, er en vildledende målsætning, som i stedet kan bidrage til at reproducere forestillingen om teknologisk neutralitet.
Det er derfor nødvendigt at tydeliggøre, hvornår og hvorfor en given bias er problematisk, og hvordan den kan afhjælpes, ikke ved at eliminere bias helt, men ved at erstatte den med en mere retfærdig form for bias (Helm et al. 2024: 11).

Et eksempel på problematisk vi undersøger, er når sproglig bias i LLM'er viderefører uretfærdigheder ved systematisk at favorisere vestlige sprog. Dermed overses andre kulturelle karakteristika i sprog, der tales af mange, men som er underrepræsenterede eller truet af udryddelse. For at kunne afhjælpe dette strukturelt er det nødvendigt at forstå, hvordan sproglig og algoritmisk bias samvirker og danner en ny form for ulighed, som forfatterne betegner som 'language modeling bias'. Det definerer de ved, når sprogteknologi design behandler og fortolker indhold mindre præcist på visse sprog end på andre, derved tvinges mindre sprog, at forenkle deres kommunikation i tilpasningen til den standard, der er

udarbejdet på baggrund af privilegerede sprog (Helm et al. 2024: 7).

Bias i sprogmodeller handler ikke kun om mangelfulde data, men om hvordan teknologiske designvalg og infrastrukturer systematisk viderefører sproglig ulighed og repræsentationsskævhed (Helm et al. 2024: 8).

3.7 Global ulighed

I dette speciale anvendes begrebet global ulighed med særligt fokus på, hvordan udviklingen og anvendelsen af kunstig intelligens, og i særdeleshed sprogmodeller, som kan forstærke eksisterende socioøkonomisk ulighed mellem minoritets- og majoritetsverden. Med udgangspunkt i den sydafrikanske organisation, Masakhane, der arbejder for større inklusion af afrikanske sprog i sprogmodeller, undersøger vi, hvordan ulighed kommer til udtryk via teknologi.

FN definerer ulighed således:

> *"Inequality has multiple dimensions and varies depending on the context. It can include economic inequalities such as inequalities in income, wealth, wages and social protection, as well as social and legal inequalities where different groups are discriminated, excluded or otherwise denied full equality. Inequality can also refer to inequality within a country as well as inequality between different countries."* (UN System Chief Executives Board for Coordination, u.å.)

Ud fra denne definition beskæftiger vi os i specialet med ulighed mellem lande ved at sætte fokus på den ulighed, der opstår, når forskellige grupper er diskrimineret, ekskluderet eller uden for adgang til ressourcer, repræsentation og indflydelse.

International Monetary Fond beskriver at hovedårsagerne til ulighed bl.a. er globale faktorer så som globalisering og teknologisk udvikling, eksempelvis at teknologisk udvikling har bidraget til en stigende lønforskel baseret på kvalifikationer fordi folk med højere uddannelse, ofte har en fordel i at anvende nye teknologier (International Monetary Fund, u.å.).

Center for Global Development beskæftiger sig i en artikel om AI og global ulighed med de risici, der er forbundet med AI's indvirkning på ulighed. De understreger, at AI's effekt på global ulighed afhænger af de politiske valg, der træffes i dag. AI risikerer at forstærke både den interne ulighed i lande og den globale ulighed, da højtuddannede individer og regioner

vil drage uforholdsmæssige fordele, mens lavt kvalificerede arbejdstagere og ressourcefattige områder risikerer at blive efterladt. Lande med avanceret AI-teknologi vil sandsynligvis komme længere foran, hvilket forstærker den globale kløft mellem mennesker.

På den anden side har AI potentiale til at reducere uligheden, hvis teknologien anvendes til sociale formål, såsom at forbedre sundhedsvæsenet, uddannelse og landbrug i udviklingslande. For at modvirke de potentielt negative virkninger af AI er det dog nødvendigt med hurtig handling fra regeringer og internationale organisationer, herunder investeringer i uddannelse, omskoling og sociale sikkerhedsnet, samt progressive skatter og redistribuering af velstand. “Ultimately, ensuring that AI development is inclusive and benefits all requires global collaboration, ethical AI practices, and a commitment to leveraging AI for public services and underserved communities.” (Schellekens & Skilling 2024).

4. **FORSKNINGSFELT**

4. Forskningsfelt

I det følgende kapitel præsenteres eksisterende litteratur, der giver indsigt i det specifikke genstandsfelt, vi beskæftiger os med. Generativ AI og sprogteknologi i relation til global ulighed er et komplekst og hastigt voksende forskningsfelt, som aktuelt diskuteres bredt blandt forskere, politikere og i tidsskrifter. Det er et område med mange dimensioner, som vi i dette afsnit forsøger at belyse med et åbent og flerfacetteret perspektiv.
Litteraturen er struktureret ud fra centrale tematikker, som opstod under søgeprocessen, og bidrager til at belyse, hvordan generativ AI påvirker globale magtforhold og ulighed.

Med udgangspunkt i Cronin et al.'s (2008) tilgang og beskrivelse af forskningsfelt, præsenteres i følgende søgeprocessen i indsamlingen af udvalgt litteratur og eksisterende forskning, der kan give indsigt i genstandsfeltet i specialet.
For at kunne besvare hvordan udviklingen og træningen af generativ AI påvirker global ulighed, og hvordan man kan medvirke til at gøre AI mere retfærdig og inkluderende, strukturer vi forskningsfeltet således:
Indledningsvist afdækkes fokus på etiske principper og lovgivning vedrørende AI og de forskellige AI Act-initiativer, der forsøger at imødegå denne ulighed gennem regulering. Herefter zoomes der ind på sprogmodeller for at illustrere, hvordan AI's træningsdata systematisk favoriserer vestlige normer og sprog, hvilket medfører strukturel eksklusion af ikke-vestlige sprog og lande i selve udviklingen af AI. Dernæst undersøges, hvordan udviklingen af generativ AI kan forstærke både materielle og symbolske uligheder mellem majoritets- og minoritetsverden. Med dette afsæt belyses betydningen af sproglig inklusion, træningsdatas bias og de konsekvenser, det har for selve udviklingen af AI og anvendelse i globale sammenhænge. Afslutningsvis placeres specialet i relation til den eksisterende forskning, hvorefter de teoretiske perspektiver præsenteres som grundlag for den videre analyse.

4.1 Formål og søgestrategi

Forskningsfeltet fungerer i specialet som en gennemgang og kritisk analyse af relevant litteratur inden for et givent emne. Det er et felt, der skal give læseren en grundig forståelse af den eksisterende viden og samtidig fremhæve, hvorfor ny forskning er vigtig (Cronin et al. 2008: 38).

I det følgende fremlægger vi det, som kan kategoriseres som det 'traditionelle' eller

‘narrative’ forskningsfelt. Endvidere har denne type forskningsfelt til formål at inspirere til nye forskningsspørgsmål samt at afgrænse og præcisere det specifikke emne og den undren, der danner afsæt for undersøgelsen (Cronin et al. 2008: 38).

I tilegnelsen af viden trækker det traditionelle forskningsfelt på teknikker fra det systematiske for at sikre en struktureret tilgang, der mindsker risikoen for at overse grundlæggende og relevant forskning. Som en del af denne proces, udvælges først databaser, hvorfra litteraturen indsamles (Cronin et al. 2008: 39-40). Vi har brugt Aalborgs Universitetsbiloteks databaser og søgeværktøjer via Primo, samt Stanford Encyclopedia of Philosophy og Det Kongelige Bibliotek. I søgningen har vi filtreret søgeresultaterne således, at litteraturen er peer-reviewed. I søgningen har vi også begrænset litteratur til at være så aktuel forskning som muligt, eftersom vores genstandsfelt om AI, ulighed, sprogmodeller og dataetik er områder som konstant udvikles og forskes i. Derudover bruges nøgleord i søgningen, for at identificere relevant litteratur, samt kombinationen af nøgleord, som kaldes 'Boolean operators’ som ‘AND’, ‘OR’ og ‘NOT’ (Cronin et al. 2008: 40). På baggrund heraf udarbejdes søgestrenge, der kredser om regulering af kunstig intelligens, sprogmodeller samt globale uligheder, specifikt ift. det afrikanske kontinent, og bias i tilknytning til disse områder.

En del af den identificerede litteratur udvides yderligere gennem anvendelsen af ‘citation pearl growing’, en metode hvor ny relevant litteratur inddrages ved at følge referencerne i allerede fundne kilder (Rowley og Slack 2004: 35).

Vi har afgrænset os fra litteratur, der behandler andre ulighedsperspektiver i relation til AI, eksempelvis klimamæssige konsekvenser eller den skæve arbejdsdeling i træningen af AI, herunder ‘data labellers’, som typisk outsources til det globale syd. Vores undersøgelse koncentrerer sig om, hvordan AI og sprogmodeller og adgangen til disse teknologier påvirker ulighed, med særlig vægt på en konkret case, som inkluderer flere afrikanske sprog i sprogmodeller. Desuden udelader vi forskning, der behandler AI som et teknologisk neutralt værktøj uden at inddrage de sociale, kulturelle og politiske konsekvenser af dets anvendelse. Vores fokus er i stedet rettet mod de etiske og politiske implikationer af AI, med særlig opmærksomhed på global ulighed, sproglig inklusion og de geopolitiske dynamikker mellem minoritets- og majoritetsverden.

4.2 EU's regulering af AI og Afrikas AI-strategi

I 2024 har Den Europæiske Union (EU) vedtaget en omfattende regulering i form af AI Act, mens Den Afrikanske Union (AU) har udformet en overordnet strategi for kunstig intelligens. Selvom der er tale om to forskellige typer af politiske redskaber – henholdsvis lovgivning og strategisk rammesætning – er det relevant at analysere og sammenligne dem for at forstå, hvordan globale institutioner med forskellige udgangspunkter og magtpositioner forsøger at forme AI-udviklingen. Vi inddrager 'EU's AI Act' for at undersøge, hvordan en magtfuld aktør søger at regulere AI med fokus på risikominimering og etiske standarder, med særligt blik for relationen til global ulighed. Vi inddrager 'AU's AI-Strategy' for at belyse hvordan de forholder sig til de strukturelle udfordringer og prioriteter, der gør sig gældende på det afrikanske kontinent, herunder ønsket om teknologisk suverænitet, inklusion og kapacitetsopbygning. Sammenligningen giver os mulighed for at analysere, hvordan forskellige politiske tilgange adresserer eller overser spørgsmål om ulighed og retfærdighed i AI's implementering.

4.2.1 EU's AI-Act

En af de mest omfattende reguleringer på AI-området er EU's AI Act: "The AI Act is the first-ever legal framework on AI, which addresses the risks of AI and positions Europe to play a leading role globally." (European Union 2025). Denne lovgivning på tværs af medlemslandene i EU, som blev vedtaget i 2024, skal sikre, at AI anvendes på en sikker, etisk og ansvarlig måde, samtidig med at den har til formål at fremme innovation og beskytter grundlæggende rettigheder (European Union 2025). Loven kræver blandt andet risikovurdering, menneskelig overvågning og datasikkerhed for højrisiko-AI. Desuden skal generative AI-modeller som ChatGPT overholde gennemsigtigheds- og sikkerhedskrav.
EU's AI Act er en risikobaseret regulering, der opdeler AI-systemer i fire kategorier:

1. Uacceptabel risiko (forbudte systemer): AI, der kan skade borgere, som social scoring og manipulerende adfærd.
2. Høj risiko (strengt reguleret): AI til kritisk infrastruktur, rekruttering og biometrisk identifikation.
3. Begrænset risiko: Krav om gennemsigtighed, f.eks. at chatbots oplyser, at de er AI-drevne.
4. Minimal risiko: Ingen regulering nødvendig.

Loven omfatter også regler for generativ AI og træder fuldt i kraft i august 2026 (European Union 2025).

> *"Diversity, non-discrimination and fairness means that AI systems are developed and used in a way that includes diverse actors and promotes equal access, gender equality and cultural diversity, while avoiding discriminatory impacts and unfair biases that are prohibited by Union or national law."* (Den Europæiske Union 2024: art. 27).

EU har dog i flere år arbejdet på at etablere etiske rammer for AI, sideløbende med den accelererende udvikling. I 2019 offentliggjorde Kommissionen *Ethics Guidelines for Trustworthy AI*, som blev udviklet af en ekspertgruppe nedsat af Kommissionen, som den nye AI-Act bygger på (Den Europæiske Union 2024: art. 27). De syv nøglekrav i den guideline omfatter blandt andet menneskeligt tilsyn, teknisk robusthed, gennemsigtighed, samt mangfoldighed og ikke-diskrimination. Disse retningslinjer søger at sikre, at AI-udviklingen respekterer fundamentale rettigheder og etiske principper (European Commission 2019: 14). Med hensyn til nøglekravene, står der således:

> "*Data sets used by AI systems (both for training and operation) may suffer from the inclusion of inadvertent historic bias, incompleteness and bad governance models. The continuation of such biases could lead to unintended (in)direct prejudice and discrimination against certain groups or people, potentially exacerbating prejudice and marginalisation.* (...) *The way in which AI systems are developed (e.g. algorithms' programming) may also suffer from unfair bias. This could be counteracted by putting in place oversight processes to analyse and address the system's purpose, constraints, requirements and decisions in a clear and transparent manner. Moreover, hiring from diverse backgrounds, cultures and disciplines can ensure diversity of opinions and should be encouraged.*" (European Commission 2019: 18).

EU's AI Act blev dog forsinket i december 2022 ved fremkomsten af ChatGPT af virksomheden OpenAI, som medførte ændringer i udkastet, som var målrettet specifik regulering af generativ AI (Fernhout & Duquin, 2024). Dette illustrerer den hastige udvikling inden for kunstig intelligens og understreger behovet for lovgivning, der kan følge med og regulere nye teknologier.

I EU's AI Act, står også at AI-systemer utilsigtet kan videreføre og forstærke

eksisterende diskrimination, især når de trænes på datasæt, der afspejler tidligere uligheder såsom historiske data, der indeholder en biased beskrivelse af bestemte befolkningsgrupper. Dette kan føre til, at AI udformer anbefalinger eller fremstiller information, der uforvarende diskriminerer eller skader sårbare grupper, som etniske minoriteter, og fastholder og forværrer sociale uligheder.

EU-Kommissionen omtaler selv reguleringen som en 'human-centric' AI-tilgang (Samoili et al. 2021: 10), hvilket understreger, at mennesket skal være i centrum, og at AI bør forstås og udvikles med udgangspunkt i menneskelige værdier og behov. Vi stiller dog spørgsmål til hvilke mennesker der er i centrum her. Det mener vi er et centralt spørgsmål at stille, for at undgå, at AI i praksis bliver 'vest-centrisk', er det afgørende at sikre, at fordelene ved teknologien ikke primært tilfalder Vesten, men også bidrager til at mindske, frem for at forstærke, globale uligheder.

Formålet med EU's AI Act er at regulere udviklingen og anvendelsen af AI-teknologier for at beskytte borgernes rettigheder og sikkerhed. Loven indeholder specifikke forbud mod visse applikationer, og pålægger virksomheder strenge forpligtelser, især for højrisiko AI-systemer (European Union 2025). Denne tilgang er designet til at sikre en etisk og ansvarlig anvendelse af AI inden for EU, som i sidste ende potentielt også vil påvirke partnere uden for EU, som gerne vil sælge AI-produkter til EU medlemslande. EU's lovgivning og regulering fokuserer i højere grad på brugen og anvendelsen og implementering af AI end på selve udviklingen af teknologien.

4.2.2 AU's AI-Strategy

Den Afrikanske Union har i 2024 udgivet deres første AI-strategi, som bygger på tidligere styringsparadigmer samt nogle få nationale AI-strategier samt med støtte fra UNESCO (Alayande & Adams 2024: 1).

I forordet skrevet af H.E. Dr. Amani Abou-Zeid, kommissær for infrastruktur og energi i Den Afrikanske Union, er der et tydeligt fokus på, at denne strategi skal komme kontinentet og den afrikanske befolkning til gode. Hun fremhæver den internationale AI-kløft og understreger, hvordan principper, frameworks og lovgivning er essentielle for etisk design, træning og brug af AI-systemer, der respekterer den afrikanske befolknings rettigheder, kultur og værdier: "This strategy puts forward an Africa-centric, development-oriented and inclusive approach (...)" (African Union 2024: 1).

Strategien fremhæver, at en stor del af det internationale samarbejde og forhandlinger

om AI finder sted i fora som FN, G7, G20, ITU og OECD – alle etableret af udviklingslande – hvor Afrika hidtil har haft begrænset repræsentation og indflydelse. Konkret betyder det, at høje rejseomkostninger til disse møder gør det vanskeligt for afrikanske aktører at deltage, og at visse tekniske diskussioner forbliver utilgængelige på grund af manglende kapacitet inden for AI-teknologi. En yderligere udfordring er 'brain drain' – tendensen til, at højtuddannede afrikanske eksperter forlader kontinentet – hvilket betyder, at de ofte deltager i internationale AI-diskussioner uden at arbejde fra eller for afrikanske institutioner. Den Afrikanske Union har derfor en ambition om at sikre større repræsentation og en stærkere stemme for afrikanske institutioner i alle former for AI-relaterede diskussioner, der påvirker befolkningen på kontinentet (African Union 2024: 58).

Strategien har fem fokusområder:

1. Maximising AI's Benefits
2. Building Capabilities for AI
3. Minimising AI Risks
4. African Public and Private Sector Investment in AI
5. Regional and International Cooperation and Partnerships (African Union 2024: 1).

En af de største udfordringer for AI-økosystemet på det afrikanske kontinent er, at udviklingen og udbredelsen af teknologien stadig befinder sig på et relativt tidligt og eksperimenterende stadie. Meget af den anvendte AI, herunder træningsdata, er i høj grad udviklet uden for kontinentet. Derfor har det været afgørende for AU-strategien at prioritere tilpasningen af eksisterende AI-systemer samt udviklingen af nye, der er skræddersyet til den afrikanske kontekst (Alayande & Adams 2024: 3).

I strategien adresseres den systemiske skævvridning i AI-systemer, hvor algoritmer ofte trænes på datasæt, der ikke er repræsentative for den afrikanske befolkning. Dette skyldes blandt andet, at udviklingen af AI i høj grad foregår i globale tech-centre som Silicon Valley, hvor udviklere og beslutningstagere ofte deler en vestlig teknologiforståelse og prioriterer innovation og profit over repræsentation og inklusion. Denne tilgang adskiller sig fra behovene i mange andre lande - også afrikanske lande, hvor teknologiske løsninger i højere grad må tage udgangspunkt i strukturelle uligheder, i præcise kulturelle kontekster og mange forskellige sprog. Dette gør det afgørende at udvikle mere inkluderende og tilpassede AI-løsninger, der tager højde for Afrikas unikke udfordringer og behov.

> *"AI systems may perpetuate or amplify biases contained in datasets that they are trained on because, more often, data are not equitably sourced - data are usually sourced from developed countries and from non-diverse and inclusive developers' teams."* (African Union 2024: 3).

Selvom risici og negative konsekvenser behandles som et selvstændigt element, indgår disse overvejelser i alle dele af strategien. Strategien understreger behovet for etiske, ansvarlige og retfærdige AI-praksisser, der er tilpasset Afrikas unikke socioøkonomiske kontekst. Den omfatter en femårig implementeringsplan (2025-2030) med fokus på områder som kapacitetsopbygning, infrastrukturudvikling og fremme af investeringer i AI (Alayande & Adams 2024: 4).

4.2.3 EU og AU

I dette afsnit sammenlignes de forskellige prioriteter og dagsordener, der gør sig gældende i henholdsvis EU's regulering og AU's strategi for kunstig intelligens. Selvom der er tale om to forskellige typer dokumenter, giver en sådan sammenligning alligevel indsigt i de forskellige fokusområder og tilgange. At EU i 2024 vedtager bindende lovgivning, mens AU samme år lancerer en overordnet strategi, vidner i sig selv om de to regioners forskellige udgangspunkter i forhold til udviklingsniveauet inden for AI-området.

EU har implementeret bindende lovgivning med klare regler og forbud for AI-anvendelse, hvilket afspejler unionens større ressourcer og etablerede teknologiske infrastruktur. EU's tilgang er præget af stram risikobaseret regulering med fokus på at beskytte borgernes rettigheder og undgå potentielle negative konsekvenser af AI.
På den anden side har AU udviklet en strategisk ramme, der vejleder medlemslandene i at formulere egne politikker uden at pålægge specifikke juridiske forpligtelser. Dette afspejler Afrikas prioritering, hvor fokus er på at udnytte AI-teknologiens potentiale til økonomisk vækst og udvikling, på baggrund af etiske standarder. Denne forskel i tilgang afspejler de meget forskellige udgangspunkter, de to regioner har. EU har allerede etableret et solidt teknologisk grundlag og står overfor udfordringen med at regulere og kontrollere AI-udviklingen til deres fordele. Afrika står overfor et behov for at bygge kapacitet og tilpasse teknologien til kontinentets særlige forhold, hvilket gør det nødvendigt at balancere muligheden for lokal teknologisk fremdrift og samarbejde mellem medlemslandene og internationale partnere. Implementeringsplanerne afspejler også disse forskelle: EU har en detaljeret plan for håndhævelse af AI-lovgivningen på tværs af medlemslandene, mens AU

sigter mod en trinvis implementering, og arbejder på at udnytte AI's muligheder, som stadig befinder sig på et tidligt stadie.

4.3 Bias i udviklingen af store sprogmodeller

4.3.1 Dominerende sprog i træningsdata

Da halvdelen af internettets indhold er på engelsk, er det ikke overraskende, at størstedelen af de data, der bruges til at træne AI-baserede sprogmodeller, også er på engelsk (W3Techs, n.d.). Dette har gjort det muligt for forskere at opnå imponerende resultater inden for sprogteknologi, især i Nordamerika og Europa. Ved at udnytte den sproglige ensartethed i tilgængelige data har virksomheder udviklet succesfulde chatbots og maskineoversættelses-apps.

Ada Lovelace Institute, er en uafhængig britisk forskningsinstitution, der arbejder for at sikre, at data og AI udvikles og anvendes til samfundets fordel på en retfærdig måde. En forsker fra instituttet, Hannah Claus, påstår at, " (...) mainstream AI platforms amplify pre-existing epistemic injustices." (Claus 2024). Hun understreger, hvordan denne udvikling ikke kun ekskluderer ikke-engelsktalende, men også, at LLM'er ikke formår at støtte andre sprog end engelsk. Hermed bidrager de til at underminere og udvande kulturel historie og måder at tænke på:

> *"When LLMs fail to support languages other than English, they do more than excluding non-English speakers. They erase cultural histories and ways of thinking. This erasure is all the more concerning in a world where AI is becoming integrated into our daily lives and interacts with essential services, from education to healthcare, and high-stake functions of government, like border control."* (Claus 2024).

Et eksempel på, hvad sproglige og kulturelle unøjagtigheder kan føre til, er en AI-baseret app, der blev anvendt i asylsagsbehandling, men leverede en upræcis oversættelse. Det resulterede i, at asylansøgere blev tilbageholdt i flere måneder. Sådanne fejl skyldes ofte underliggende bias relateret til kultur, race og køn. Dette er særligt problematisk i kritiske systemer, hvor præcise oversættelser kan være afgørende – i værste fald med menneskeliv på spil (Claus 2024).

Når et dominerende system udbredes som standard af majoriteten, men ikke formår at tjene minoritetsgrupper ligeværdigt, opstår der ofte en modreaktion, hvor alternative systemer

udvikles. Dette kan imidlertid føre til nye udfordringer. Claus peger på to overordnede problemer, som underrepræsenterede grupper typisk møder:

1. For det første bliver disse alternative systemer ofte opfattet som reaktioner på mangler ved de dominerende systemer, snarere end som selvstændige og legitime strukturer med egne perspektiver og værdier.
2. For det andet indgår de ofte i fragmenterede økosystemer, hvor indsatser gentages, og fællesskaber i høj grad mangler koordinering og samarbejde (Claus 2024).

Et thailandsk studie, udgivet i 2024 af Association for Computational Linguistics, *Do Llamas Work in English? On the Latent Language of Multilingual Transformers*, undersøger, om flersprogede sprogmodeller, der er trænet på overvejende engelske data, bruger engelsk som et mellemled, når de behandler andre sprog. Ved at teste Llama-2-modellen med specifikke ikke-engelsksprogede sætninger opdagede forskerne, at modellen i midten af sin bearbejdning ofte først oversætter betydningen til engelsk, før den vender tilbage til det oprindelige sprog. Det tyder på, at modellen er mere tilpasset engelsk som meningskontekst med de muligheder og begrænsninger, det engelske sprog indeholder, hvilket kan skabe skævheder i, hvordan den håndterer andre sprog og oversættelser. Forfatterne understreger, hvordan de fleste mainstream generative AI-platforme er genereret ud fra en anglo-centreret standard, som gør det vanskeligt for en stor del af verdens befolkning at adopterer en teknologi, som ikke er bygget til dem, hvis teknologien ikke understøtter ens sprog eller kultur (Wendler et al. 2024: 8).

Et studie af Lane Schwartz fra Department of Linguistics, University of Illinois, opfordrer NLP-forskere til at anerkende kolonialismens indflydelse på arbejdet med oprindelige sprog. Schwartz argumenterer for, at forskning i disse sprog bør følge etiske retningslinjer, der centrerer oprindelige samfunds mål og viden. Artiklen foreslår udviklingen af en etisk ramme for ansvarlig forskning i relation til oprindelige sprog. I de sidste årtier er engelsk blevet det dominerende sprog i AI-forskning, især inden for 'natural language processing' (NLP), hvilket har konsekvenser for minoriteter, der ikke er repræsenteret i forskningen (Schwartz 2022: 724). Han argumenterer for, at forskning ofte hviler på filosofiske antagelser om, at videnskaben er værdineutral og sjældent overvejer, hvordan den er styret af en bestemt ontologi, epistemologi, metodologi og aksiologi, som sjældent anerkendes eller reflekteres over. Ifølge Schwartz er det første skridt mod en afkoloniseret tilgang at blive bevidst om disse grundantagelser og anerkende alternative videnskabelige

traditioner, såsom oprindelige folks filosofier, der bygger på relationer og metafysisk helhedsforståelse (Schwartz 2022: 726). Uden denne refleksion risikerer forskere ubevidst at placere deres egne metoder over oprindelige samfunds perspektiver. Schwartz påpeger videre, at den sproglige homogenitet i forskning fra Association for Computational Linguistics (ACL) er et symptom på et langt større problem: Forskningens paradigmer er dybt forankret i en vestlig videnskabelig tradition, der er uadskilleligt forbundet med kolonialisme (Schwartz 2022: 726).

4.3.2 Underrepræsenterede sprog i træningsdata

Hannah Claus nævner Sydafrika som eksempel på, hvordan apartheidtiden var præget af massiv undertrykkelse og diskrimination baseret på både race og folks modersmål. Da engelsk var hovedsproget under apartheid, og desuden var det videnskabelige hovedsprog, fik mange oprindelige samfund aldrig muligheden for at kommunikere videnskabeligt på deres eget sprog. Det har resulteret i manglende videnskabelige termer på mange indfødte sprog, og dermed en barriere for brugerne af disse sprog i forhold til at tage del i læring, videnskab og deltagelse. En måde at afkolonisere videnskab på er gennem organisationer som Masakhane, der arbejder på at oversætte videnskabelige tekster til 'low-resource'-sprog, hvilket skaber nye terminologier og udvider datasæt til videre træning af LLM'er (Claus 2024).

I 2008 udgav professor Anil Kumar Singh hos det indiske Language Technologies Research Centre en forskningsartikel kaldet *Natural Language Processing for Less Privileged Languages: Where do we come from? Where are we going?*. Den undersøger de barrierer, der gør det svært at forske i mindre privilegerede sprog, og hvordan de kan overvindes (Singh 2008: 7).

Singh påpeger hvordan de fleste NLP-værktøjer og teknologier er tilpasset engelske og europæiske sprog, som har skabt et: "(...) linguistic, computational and computational linguistics gap between the 'more privileged' and 'less privileged' languages." (Singh 2008: 7). Han uddyber hvad 'privilegeret' vil sige i denne sammenhæng således:

> *"By 'privileges' we mean the availability of finance, equipment, human resources, and even political and social support for reducing the lack of computing and language processing support. The lack of such 'privileges' may be the single biggest reason which is holding back the progress towards providing computing and language processing support for these languages."* (Singh 2008: 9).

Et sprog betegnes som 'low-resource', når der mangler tilstrækkelige mængder tekstdata af høj kvalitet, samt dokumentation, teknologisk støtte og uddannelsesressourcer. Dette kan skyldes kolonialisme, hvor dominerende sprog har fortrængt eller marginaliseret lokale sprog og udnyttet deres ressourcer, utilstrækkelig teknologisk infrastruktur eller mangel på eksperter i programmering, der både behersker sproget og etisk kan analysere de manglende data. Som følge heraf trænes LLM'er primært på sprog fra ressourcestærke nationer, hvilket efterlader minoritetssprog underrepræsenterede og forstærker skævheder i datasættene (Singh 2008: 8).

Dog er selve begrebet 'low-resource' om sprog debatteret. Dette kommer blandt andet til udtryk i artiklen *Endangered Languages are not Low-Resourced!* af forsker fra University of Helsinki, Mika Hämäläinen. Artiklen undersøger forholdet mellem truede sprog og 'low-resourced' sprog indenfor natural language processing (NLP). Hämäläinen kritiserer tendensen til at betragte alle ikke-engelske sprog som 'low-resourced' og påpeger, at dette ikke altid er præcist, især for truede sprog, hvor de lingvistiske ressourcer kan variere betydeligt (Hämäläinen 2021: 1). Nogle truede sprog kan have betydelige ressourcer til NLP-opgaver, som for eksempel Komi-Zyrian, et sprog med kun 217.000 talere, der har opbygget en imponerende samling af digitale ressourcer, herunder grammatiske værktøjer og oversættelser af russiske lovtekster (Hämäläinen 2021: 3).

Hämäläinen understreger, at det derfor er misvisende at sammenligne truede sprog med andre 'low-resourced' sprog, på grund af de store forskelle i tilgængelige ressourcer (Hämäläinen 2021: 7).

Et eksempel, der kan illustrere forskellen i ressourcer og repræsentation mellem sprog, omhandler Basque, et isoleret sprog med cirka 800.000 talere, men som har en langt bedre online-repræsentation end Kiswahili, der tales af 60–150 millioner mennesker. Basque tales i et økonomisk velstående område, mens Kiswahili tales i regioner, der historisk har været ressourcemæssigt udnyttet. Denne forskel understøtter argumentet om, at 'low-resourced' betegnelsen bør anvendes med hensyn til historiske og socioøkonomiske faktorer (Claus 2024).

En sammenligning mellem det danske sprog og sydafrikanske sprog som Tswana og Xhosa viser, at de begge er underrepræsenterede i datatræningssæt til sprogmodeller. Hvor der allerede findes en betydelig mængde videnskabelig data på dansk, som blot skal indsamles og bearbejdes, er situationen anderledes for Tswana og Xhosa. Selvom der er flere talere af disse sprog, er der ikke nødvendigvis samme mængde tilgængelige data. Derfor er det vigtigt at forholde sig til sprog i deres specifikke kontekster, når termer som 'low-

resourced' anvendes, da det ikke er de samme problemer som er gældendene.

I artiklen *Toward More Meaningful Resources for Lower-resourced Languages* argumenterer forfatterne for, hvordan meningsfulde sprogressourcer for 'low-resource' sprog bør udvikles i tæt samarbejde med sprogets modersmålstalende. Artiklen fremhæver, hvordan manglende etiske og grundige annoterings- og udviklingsprocesser kan føre til fejl og begrænset anvendelighed i sprogteknologi (Lignos et al. 2022: 523). Her præsenteres et perspektiv på, hvordan udviklingen af sprogmodeller kan bidrage til øget ulighed, hvis de begrænser anvendeligheden for store befolkningsgrupper gennem lav inklusion og repræsentation i udviklingsprocessen. Forskerne peger desuden på at manglende bevidsthed i indsamlingen af data fra 'low-resource' sprog, kan resultere i lavkvalitets datasæt, der potentielt gør mere skade end gavn:

> *"Failing to exercise caution may result in creating low-quality derived datasets that may do more harm than good. For example, if a multilingual dataset contains a large number of incorrect copies of English in other languages, it may make tasks appear easier than they are because of trivial transfer from English."* (Lignos et al. 2022: 525).

Artiklen argumenterer for, at udviklingen af sprogmodeller for 'low-resource' sprog kræver mere end blot automatiserede datasæt, som ofte fører til lav kvalitet i output, som ikke tager højde for sproglige og kulturelle nuancer. Forfatterne understreger vigtigheden af at involvere modersmålstalere i udviklingen af træningsdata på en ansvarlig og etisk måde, som det ses i Masakhane-projektet. Her engageres talere af afrikanske sprog aktivt i processen, hvilket sikrer mere præcise og kulturelt passende resultater. Artiklen fremhæver, at for at arbejde etisk og ansvarlig med udviklingen af sprogmodeller bør forskere og udviklere samarbejde med sprogets modermålstalere og inkludere deres perspektiver og viden i udviklingsprocessen af sprogmodeller (Lignos et al. 2022: 528).

Forfatterne peger på vigtige retningslinjer for fremtidig forskning, herunder nødvendigheden af at inddrage modersmålstalende i kvalitetskontrol og at undgå at udgive lavkvalitetsdatasæt, der kan forstyrre evalueringsprocesser. De anbefaler også, at menneskelig annotering prioriteres, og at alle beslutninger truffet af de, der annoterer data, offentliggøres for at sikre kvalitet, transparens og ansvarlighed i processen, frem for at benytte automatiserede processer (Lignos et al. 2022: 529).

4.3.3 Udfordringer ved bias i store sprogmodeller

En anden artikel kritiserer, hvordan AI-baseret sprogressourcer primært udvikles til de mest talte og magtfulde sprog, mens forsøg på at inkludere, noget som de betegner som 'underserved languages' ofte lider af en indbygget bias (language modeling bias). Denne bias gør, at AI-systemer favoriserer visse sprog og kulturer, mens de marginaliserer andre. Forfatterne påpeger, at problemet skyldes en forsimplet forståelse af sproglig diversitet blandt udviklere, hvilket fører til 'epistemisk uretfærdighed' – en systematisk udelukkelse af sprogfællesskabers viden og behov. De argumenterer for en ny socio-teknisk tilgang, der bedre kan adressere disse skævvridninger (Helm et al. 2024: 1).

Når det gælder sproglig bias i NLP, identificeres fem kilder til bias: (1) data, (2) annoteringsprocessen, (3) modellerne, (4) inputrepræsentationsprocessen og (5) forskningsdesignet til at undersøge bias (Helm et al. 2024: 2).

De peger på, hvordan forskere har forsøgt at udvikle teknologier som maskinoversættelse, naturlig sprogbehandling og stemmegenkendelse til et stadig større antal sprog. Mange af disse initiativer bygger dog på en forestilling om, at AI kan overføre og anvende eksisterende metoder og systemer, som er skabt ud fra en anglocentrisk teknologikultur, direkte til andre sproglige og kulturelle kontekster. Selvom det har til formål at mindske den digitale kløft, vil en manglende forståelse for sproglig og kulturel diversitet ofte medfører, at væsentlige kvalitetsproblemer overses, som tidligere litteratur også har peget på. Endelig kan dette resultere i en omfattende vestliggørelse af kulturelle udtryk og en forstærkning af epistemisk uretfærdighed (Helm et al. 2024: 2).

Eftersom forfatterne beskæftiger sig med et normativt syn på at beskytte, promovere og preservere diversitet kommer de med en konceptualisering: de anerkender, at diversitet både har en moralsk og en erkendelsesmæssig dimension. Forfatterne skelner derfor mellem to måder at forstå diversitet på. Både som et 'deskriptivt' begreb, der handler om de faktiske sproglige forskelle, der bør indgå i udviklingen af sprogmodeller, og som et 'normativt' begreb, der afspejler de værdier og mål, der ligger til grund for arbejdet med sprogteknologi: "(...) the term "epistemic injustice" not only describes well the homogenization that can result from loss of diversity, but also situates that loss and the attached harms within a broader context of global inequalities." (Helm et al. 2024: 3-4).

Helm et al. refererer til Spivak, som taler om epistemisk kolonisering, som er de processer, hvor én kulturs videnssystemer, tro og erkendelsesformer påtvinges en anden kultur eller et andet samfund – ofte som en direkte, indirekte eller sen konsekvens af kolonialisme eller

imperialisme. Dette indebærer dominansen af en bestemt vidensteori (i dette tilfælde en tro på den universelle anvendelighed af AI-systemer udviklet i Vesten) over andre, hvilket ofte resulterer i marginalisering eller undertrykkelse af lokale videnssystemer og måder at forstå verden på (Helm et al. 2024: 5).

> *"Linguistic bias, for example, is harmful to diversity when it perpetuates existing or produces new forms of hermeneutic injustices related to already vulnerable, and/ or marginalized language communities. Such bias calls for counteraction. When such linguistic bias is then reproduced through LLMs that are geared toward the correct representation of languages of colonial powers, but disregard the particularities of other languages that are also spoken by many people or are at risk of extinction, this demands change."* (Helm et al. 2024: 6).

For at skabe en bæredygtig forandring er det afgørende at undersøge, hvordan sproglig bias og algoritmisk bias påvirker hinanden og tilsammen danner en ny form for bias, som forfatterne betegner som 'language modeling bias' (Helm et al. 2024: 6).

Et eksempel på dette er, når en oversættelse kun kan oversætte 'ris' til én form for 'ris', selvom både swahili og japansk har præcise ord for forskellige former for 'ris'. Når oversættelsen udelukkende foretages på engelsk, går disse sproglige nuancer tabt, hvilket resulterer i både et tab af detaljer og fejlagtige oversættelser, der gør teknologien upræcis og potentielt ubrugelig i den kulturelle anvendelseskontekst (Helm et al. 2024: 8).

Når standardiserede oversættelsesteknologier bliver normen i flersproget kommunikation, kan det føre til, at én person kan udtrykke sig klart på sit eget sprog, mens en anden er begrænset i at udtrykke sine sociale realiteter. Dette er et eksempel på hermeneutisk uretfærdighed. F.eks. ændrer engelsk indflydelse på afrikansk kultur folks selvopfattelse og sociale praksisser, men problemet ligger ikke i kulturel blanding, men i dominansen af visse kulturer, som opretholdes af AI-sprogteknologi og bør overvindes for at fremme epistemisk retfærdighed (Helm et al. 2024: 11).

Forskelle findes ikke kun i grammatiske strukturer, men også i folks sociale praksisser, verdensbilleder og de vidensperspektiver, der er indlejret i sproglig udtryksform (Helm et al. 2024: 12). Forfatterne argumenterer for at problemerne i en etnocentrisk udvikling af sprogteknologi har rødder i en kolonial fortid, der stadig præger nutidens praksis. Der er derfor et presserende behov for at være opmærksom på både de velkendte former for sproglig bias i AI-systemer og den bias, der opstår fra selve designet af

sprogmodellerne.

Forfatterne argumenterer for, at udviklingen af teknologi baseret på bias i værktøjer ikke kun fører til utilstrækkelige sprogbeskrivelser, men skaber en form for uretfærdighed, som de kalder epistemisk eller hermeneutisk uretfærdighed. Denne uretfærdighed handler ikke om fordelingen af viden, men om manglende anerkendelse af bestemte former for viden, udtryksmåder og sociale realiteter. Forfatterne konkluderer, at enhver indsats for at udvide eksisterende sprogteknologier, som ikke bygger på et grundigt samarbejde med de relevante sprogsamfund, bør genovervejes. Dette samarbejde bør baseres på en kritisk tilgang til de privilegier, der følger med hvidhed, og ses som en mulighed for gensidig læring, hvor ingen part er overlegen. Denne tilgang vil sikre en meningsfuld mangfoldighed i stedet for blot kosmetisk mangfoldighed (Helm et al. 2024: 12).

En anden problematisering af bias præsenteres af Kate Crawford i bogen *Atlas of AI* (2021), hvor hun retter fokus mod forskning i AI-genererede værktøjer og de etiske implikationer, der kan opstå, når generativ AI anvendes i stedet for menneskelig interaktion. Ifølge Crawford ligger problemet ikke alene i tekniske fejl som bias i datasæt eller algoritmer, men også i et dybereliggende strukturelt niveau, hvor forskere og udviklere, ofte ubevidst, viderefører idéer, der skader sårbare grupper og forstærker eksisterende uretfærdigheder. Hun argumenterer således for, at det ikke er tilstrækkeligt at fokusere på de tekniske aspekter af AI men der må også tages højde for de sociale og etiske konsekvenser: "The problem here isn't only one of biased datasets or unfair algorithms and of unintended consequences. It's also indicative of a more persistent problem of researchers actively reproducing ideas that damage vulnerable communities and reinforce current injustices." (Crawford 2021: 117).

Crawford fremhæver desuden, at fokus ofte ligger på at rette eller erstatte modeller, når bias opdages, snarere end at engagere sig i en offentlig debat om de underliggende årsager til, at disse former for skævhed og diskrimination gentagne gange opstår. Hun argumenterer for, at bias i mange tilfælde bør ses som et symptom på dybere strukturelle problemer og ikke blot som tekniske fejl, der kan løses isoleret (Crawford 2021: 129).

Derudover fører udviklingen af ny teknologi til en øget koncentration af magt og velstand, hvilket forstærker den globale ulighed. Eksempelvis har tech-giganter som Apple og Google større aktiver end hele nationer. Disse techgiganter tilbyder værktøjer, der spænder fra organisering af fødselsdage til drift af multimillion-dollar virksomheder – en række funktioner og dyb indvirkning på en befolkning, som kun nationale stater har kunne tilbyde på så bredt et plan. 'The coming wave' vil accelerere tendensen yderligere og koncentrere

endnu mere magt og rigdom hos dem, der skaber og kontrollerer teknologien, påpeger Suleyman (Suleyman 2023: 188).

4.3.4 Nye lokale forskningsfællesskaber

I bogen *Responsible AI in Africa* (2023) diskuterer forfatterne AI-etik fra både et teoretisk og praktisk afrikansk perspektiv med fokus på tech-virksomheder indflydelse, globalisering og kolonisering. De undersøger desuden de sociokulturelle kontekster for ansvarlig AI og hvordan den kan udvikles med sensitivitet over for afrikanske kulturer og samfund (Eke et al. 2023: 5). Forfatterne påpeger, at udviklingen og implementeringen af AI primært har været centreret i det globale nord, og at denne magtkoncentration hænger direkte sammen med kolonihistorien. Som følge heraf har store dele af Afrika ikke haft de samme ressourcer til at industrialisere, hvilket har begrænset mulighederne for at udvikle og anvende AI-værktøjer optimalt. Effektiv AI-adoption og implementering afhænger blandt andet af en række faktorer som:

> *"(...) having a local workforce with the required training to develop these solutions, sufficient infrastructural capacity to handle the computationally-heavy training of algorithms, representative datasets, governmental support and regulation to govern the appropriate fair use of these technologies, independent and civil institutions and policymakers that safeguard from harmful applications and reinforce responsibility and accountability."* (Eke et al. 2023: 35)

Artiklen nævner, hvordan nye forskningsfællesskaber som Masakhane ofte operere med begrænsede ressourcer og kapacitet, hvor det kan være vanskeligt at opnå de nødvendige færdigheder, blandt andet fordi de have en utilstrækkelig digital infrastruktur, dårlig internetforbindelse og være markant bagud i en stærkt konkurrencepræget international industri, som ikke tager hensyn til deres behov. Eksempelvis kan udviklingen hurtigt gå i stå, hvis man ikke har adgang til kvalificerede medarbejdere med tilstrækkelige kompetencer til at kunne operere LLM'er og generativ AI (Eke et al. 2022: 80). Dette dilemma kan forstås gennem Spivaks begreb 'epistemologisk vold', som hun introducerer i værket *Can the Subaltern Speak?* (1988) Med begrebet beskriver Spivak, hvordan bestemte grupper – de såkaldt subalterne – ikke alene marginaliseres politisk og økonomisk, men også oplever at deres viden og perspektiver bliver tilsidesat eller forvrænget i dominerende systemer for forståelse og beslutningstagning (Spivak 1988: 24). Det handler ikke bare om, at de ikke

bliver hørt, men at de fratages muligheden for overhovedet at definere deres egne erfaringer på egne præmisser (Spivak 1988: 25). I dag ser vi paralleller i teknologiudviklingen, hvor AI og sprogmodeller i høj grad formes af aktører med bestemte kulturelle og økonomiske udgangspunkter. Det kan betyde, at minoriteters erfaringer og sprog enten udelades eller oversættes gennem standarder, som ikke tager deres virkelighed alvorligt. I forlængelse af Spivaks pointe kan man se på, hvordan man kan udvikle teknologier, der ikke blot taler om marginaliserede grupper, men skaber rum for, at de kan tale 'på egne vilkår'. Spørgsmålet om hvordan man gør AI mere demokratisk, er også blevet til termen 'demokratiseringen af AI'. Et term, som institutioner, organisationer, virksomheder og politikere bruger i samtalen om at sørge for at AI-udviklingen kommer flest muligt til gode. Dette term undersøger Elizabeth Seger, forsker ved Centre for the Governance of AI i Oxford, UK i artiklen *Democratising AI: Multiple Meanings, Goals, and Methods* (Seger 2023: 1). Artiklen undersøger de forskellige betydninger af begrebet 'demokratisering af AI' og identificerer fire centrale dimensioner:

1. Demokratisering af AI-brug
2. Demokratisering af AI-udvikling
3. Demokratisering af AI-profitter og
4. Demokratisering af AI-styring.

Artiklen argumenterer for, at disse aspekter ofte forveksles eller står i konflikt med hinanden og at demokratisering af AI ikke blot handler om øget adgang til teknologien, men også om at håndtere de afvejninger og risici, der opstår på tværs af brug, udvikling og økonomiske gevinster (Seger 2023: 1). For eksempel kan ønsket om at gøre AI-modeller frit tilgængelige (demokratisering af udvikling) kollidere med behovet for at regulere adgangen for at forhindre misbrug (demokratisering af styring). Således kan en ubalanceret tilgang til én dimension utilsigtet underminere en anden, hvilket komplicerer implementeringen af en helhedsorienteret demokratisk tilgang til AI (Seger 2023: 8):

> *"We should therefore be wary of using the term "democratisation" too loosely or, as is often the case, as a stand-in for "all things good". The democratisation of AI use, AI development, and even AI derived profits are not inherently good. Their value is derived from alignment with interests and values of those who will be impacted. As such, where the democratically aligned decision would be to limit accessibility, the*

> *democratisation of AI governance takes precedence over the others as the source from which the moral and political value of the "democratisation" terminology is derived."* (Seger 2023: 8).

Claus (2024) peger også på hvordan organisationer som Masakhane står overfor visse udfordringer i ønsket om at demokratisere og i skabelsen af et oprigtigt inkluderende system. De kan også komme til at reproducere diskrimination ved kun at fokusere og selektere de mest talte afrikanske sprog (Claus 2024).

Selvom det i sig selv kan være vanskeligt at sørge for reel inklusion for alle, kan AI-systemer til minoritetssprog have flere fordele i modsætning til en samlet model, en slags universel ChatGPT, hvor udfordringen blandt andet ligger i at 'open-source AI' kan skabe nye magtstrukturer, hvor virksomheder fralægger sig ansvar ved at give fri adgang til AI-systemer uden at tage stilling til deres etiske eller sociale konsekvenser. I kontrast kan udviklingen af minoritetsbaserede systemer, som Masakhane, tilpasses specifikke sproglige behov, hvilket forbedrer oversættelser og kvaliteten af indhold. Samtidig kan de styrke lokalsamfund ved at respektere deres sprog og kultur, mindske afhængigheden af dominerende systemer og sikre mere etisk behandling af samarbejdspartnere – især med hensyn til betaling og arbejdsvilkår i udviklingen af LLM'er (Claus 2024).

4.4 Når AI tilskrives menneskelige egenskaber

Spørgsmålet om, hvem der bærer ansvaret for en retfærdig og inkluderende udvikling af AI, er centralt, når man undersøger, hvordan teknologier som generativ AI eksempelvis via sprogmodeller påvirker global ulighed. Uanset om det drejer sig om aktører i industrien, nationale eller internationale lovgivere, indkøbere eller brugere, kan ansvaret for teknologiens konsekvenser hurtigt blive udvandet, ikke mindst gennem det sprog, vi anvender til at beskrive AI.

Sproget, vi bruger om AI, har stor betydning for, hvordan vi forstår teknologien og dens konsekvenser. Når vi taler om, hvordan AI kan øge global ulighed, er det afgørende ikke at fjerne menneskelig agens fra ligningen.

Som Kate Crawford udfolder, er AI et system skabt og styret af mennesker, ofte af aktører med betydelig magt og ressourcer. AI har ikke autonomi. Og tilskriver man teknologien agens, risikerer man at sløre de sociale, politiske og økonomiske beslutninger, der ligger bag teknologiens udvikling og anvendelse (Crawford 2021: 8).

Emily M. Bender og hendes kolleger har i deres arbejde kritiseret den måde, hvorpå store sprogmodeller, som ChatGPT-3, bliver fremstillet og opfattet. I artiklen *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* advarer de om, at disse modeller genererer tekst baseret på statistiske mønstre uden nogen egentlig forståelse af indholdet. Dette kan føre til, at AI-systemer opfattes som intelligente eller forstående, hvilket kan sløre det menneskelige ansvar bag teknologien (Bender et al. 2021: 615).
Desuden påpeger de, at træningsdata ofte er skævt fordelt, hvilket kan forstærke eksisterende sociale og kulturelle skævheder, eksempelvis ved at filtrere diskurser om marginaliserede grupper ud i træningssæt samt generelle bias, inklusiv stereotypiske associationer, i sprogmodeller (Bender et al. 2021: 614). De påpeger således på at i filtreringen og udvælgelsen af data, sprogmodeller trænes på, kan være med til udelukkende repræsentere et hegemonisk synspunkt og ultimativt forstærke ulighed: "In accepting large amounts of web text as 'representative' of 'all' of humanity we risk perpetuating dominant viewpoints, increasing power imbalances, and further reifying inequality." (Bender et al. 2021: 614).

I artiklen *Resisting Dehumanization in the Age of AI f*remhæver Bender behovet for at opretholde en klar skelnen mellem menneskelig og maskinel kognition. Hun advarer mod at tilskrive AI-systemer menneskelige egenskaber, da dette kan føre til dehumanisering og en reduktion af menneskelig ansvarlighed (Bender 2024: 4). Bender definerer dehumanisering således:

> *1. Cognitive state of failing to perceive another human as fully human.*
>
> *2. Acts that express that cognitive state or otherwise entail the assertion that another human is not fully human.*
>
> *3. Experience of being subjected to acts that express lack of perception of one's humanity and/or deny human experience or human rights.* (Bender 2024: 4).

Bender peger på, at denne dehumanisering ofte sker i kraft af sproget omkring AI, hvor teknologien personificeres, mens de mennesker, der skaber og anvender systemerne, træder i baggrunden. Hun kritiserer, at AI i høj grad fungerer som en ideologi, der skjuler menneskeligt ansvar og opfordrer til, at vi stiller kritiske spørgsmål til dem, der står bag teknologien: Hvorfor er dette sikkert? Hvem gavner det? Og hvilke interesser kommer hvem til gode? (Bender 2024: 10).

Bender fremhæver vigtigheden af at decentrere hvidhed og det engelske sprog som de

usagte standarder i AI-forskning og udvikling. Når man undlader eksplicit at navngive disse som særlige og lokaliserede kategorier, risikerer man at fremstille erfaringer og data fra hvide, engelsktalende og vestlige grupper som universelle, mens andre grupper fremstår som undtagelser eller afvigelser. Det fører ikke alene til svagere forskning, men fastholder også en hvid, vestlig referenceramme som norm. Ifølge Bender kan et AI-system ikke kaldes præcist eller velfungerende, hvis det systematisk fejler over for marginaliserede grupper – for hvis det ikke virker for alle, virker det ikke. Hun kritiserer desuden, hvordan begrebet 'intelligens' i AI ofte forbindes med vestlige idealer om rationalitet, hvilket risikerer at overse og nedvurdere andre former for viden og kognition. At tage et opgør med dette er ifølge Bender afgørende for at modarbejde dehumaniserende tendenser i udviklingen af AI (Bender 2024: 10).

For at bevare en kritisk og præcis forståelse af AI er det derfor nødvendigt at være opmærksom på, hvordan vi taler om teknologien. AI er ikke en neutral eller selvstyrende kraft, men er et produkt af menneskelige valg, og det er fortsat os, der har ansvaret for, hvordan den anvendes.

4.5 AI's påvirkning på global økonomi

Ifølge et indlæg af den israelsk-venezuelanske økonom og direktør for migrationsprogrammet ved Center for Global Development, Dany Bahar, transformerer AI-udviklingen industrier og fagområder og har uundgåeligt også konsekvenser for forholdet mellem rig og fattig: "With AI already impacting certain occupations and industries, one can't help but ask: how will this transformative technology affect closing the gap between rich and poor countries–one of the most important challenges of our time?" (Bahar & Michael 2024).

I en 'Staff Discussion Note' fra International Monetary Fund (IMF) i 2024 med titlen *Gen-AI: Artificial Intelligence and the Future of Work* diskuteres AI's potentiale til at omstrukturere den globale økonomi. Notatet peger blandt andet på, at udviklede økonomier bør prioritere udviklingen af digital infrastruktur og digitale evner. Den peger også på, at flere faktorer afgør, i hvilket omfang lavindkomstlande kan drage nytte af AI og de medfølgende teknologiske fremskridt. Blandt disse faktorer er digital infrastruktur og andelen af arbejdsstyrken som har digitale færdigheder:

> "*On the other hand, low-income countries, although relatively less exposed, are underprepared across all dimensions to harness the benefits of AI. Notably, weak*

> *digital infrastructure and a less digitally skilled labor force are a concern.*" (Cazzaniga et al. 2024: 20).

I et policy brief fra tænketanken The Policy Center for the New South i 2024 kaldet *Artificial Intelligence in Africa: Challenges and Opportunities*. Rapporten diskuterer AI's potentiale og udfordringer i Afrika, hvor digitale kløfter og uligheder hæmmer udviklingen. Den fremhæver behovet for investeringer i infrastruktur, uddannelse og forskning samt strategiske partnerskaber for at sikre en inkluderende og bæredygtig AI-udvikling. Det understreges hvordan ingen afrikanske lande rangerer i top-50 af lande som er klar til AI og får svært ved at følge med resten af verden. Dette understreger, at den udvikling, mange vestlige lande har været i stand til at gennemføre, kun forstærker kløften mellem Afrika og resten af verden i forhold til adoptionen og udnyttelsen af AI's teknologiske fremskridt.

> "*Africa lags behind other regions of the world in AI adoption and readiness. AI startups in Africa emerged nearly a decade after the beginning of the Fourth Industrial Revolution in 2000, and no African nation ranks among the top 50 countries in terms of government readiness for AI (Arakpogun et al., 2021). This lag highlights the challenges Africa faces in catching up and fully adopting AI technologies.*" (Azaroual 2024: 12).

4.6 Opsamling og placering af undersøgelse

Gennem den skitserede forskning opnår vi indsigt i et lille hjørne af feltet, der belyser hvordan AI's udvikling gennem sprogmodeller bidrager til global ulighed. Datasæt, der trænes af vestlige tech-virksomheder, bærer præg af bias og udelukker ofte 'low-resource languages', hvilket har afgørende betydning for LLM'ers funktion og anvendelighed.

Inddraget forskning har primært fokuseret på de potentielle negative konsekvenser af mangelfulde AI-sprogmodeller, herunder hvordan de påvirker eksisterende uligheder. For tech-virksomheder og teknologisk privilegerede stater ligger opmærksomheden primært på implementering og sikkerhed, snarere end på de tidlige faser af AI-udviklingen. Flere forskere har dog fremhævet behovet for at rette blikket mod selve træningen af sprogmodeller, hvor bias og eksklusion allerede begynder at tage form. Det er i dette spændingsfelt, vores undersøgelse placerer sig. Dette speciale bidrager med en tilgang til, hvordan inklusion og diversitet i udviklingsprocessen af AI, kan reducere uligheder forbundet med sprogteknologi.

Med udgangspunkt i et land som Sydafrika, retter vi blikket mod hvilke konsekvenser det har for lande i majoritetsverden, som ikke har samme teknologiparathed som lande i minoritetsverden. Vi fokuserer på, hvordan strukturel skævvridning kan adresseres allerede i den indledende fase af AI-udviklingen, og dermed udvider forståelsen af, hvordan teknologiens opbygning kan påvirke global ulighed.

5. TEORI

5. Teoretisk udgangspunkt

5.1 Stakeholdermap

Følgende afsnit identificerer de aktører, der er involveret i den sydafrikanske non-profit organisation Masakhane og organisationens arbejde med at fremme afrikanske sprogmodeller. I afsnit 6. laver vi metodisk en uddybende præsentation af Masakhane som vores case.

Med afsæt i stakeholderteori søger vi at undersøge, hvordan forskellige interesser, magtforhold og ressourcer både påvirker og påvirkes af græsrodsorganisationens indsats (Freeman 2015: 52). Masakhane repræsenterer et lokalt forankret alternativ i et teknologisk landskab, der i høj grad er domineret af globale tech-virksomheder.

Med udgangspunkt i Edward Freemans stakeholderteori, der anvender en bred og inkluderende forståelse af interessenter, søger vi at identificere de grupper og aktører, der enten påvirker eller påvirkes af organisationens formål og virke: ”The very definition of ‘stakeholder’ as "any group or individual who can affect or is affected by the achievement of an organization's purpose." (Freeman 2015: 52).

Ud fra Freemans stakeholdermap identificeres de centrale stakeholders i relation til Masakhanes arbejde, med udgangspunkt i information fra organisationens egen kommunikation samt relevante aktører, der indirekte influerer organisationens virke (Freeman 2015: 55).

Figur 1: Stakeholdermap af Masakhane

Dette stakeholdermap identificerer en række aktører, der enten har indflydelse på eller selv påvirkes af Masakhanes arbejde.
Med afsæt i specialets problemformulering fokuseres der i det følgende på udvalgte interessenter, hvis roller, perspektiver og relationer til Masakhane vurderes at have særlig analytisk relevans i forhold til projektets overordnede målsætninger og den magt- og vidensdynamik, der udspiller sig i feltet.

Fællesskab

I Masakhanes femårige strategi beskrives det, hvordan organisationen strukturerer sig omkring et bredt forankret fællesskab: “At the core of this mission is our community. Our community has grown to over 2500 participants from more than 35 African countries, and more globally (our members are based across Africa and in the Diaspora).” (Masakhane Research Foundation 2024: 6). Det er dette fællesskab, der har faciliteret forskning på over 40 forskellige afrikanske sprog. Masakhanes forskningspraksis er karakteriseret ved en åben og deltagelsesbaseret metodologi, hvor input, ekspertise, konsultation og reviews er lavet på tværs af fællesskabet. Denne tilgang, som de selv betegner som 'the multi-stakeholder nature

of our methodology', bidrager til at gøre forskningsresultaterne både mere repræsentative og af højere kvalitet (Masakhane Research Foundation 2024: 6).

Forskere og akademiske institutioner

Masakhane indgår samarbejder med universiteter og forskere med specialisering i kunstig intelligens, maskinlæring og lingvistik med henblik på at udvikle sprogmodeller, der understøtter og integrerer afrikanske sprog. En af disse forskere er Benjamin Rosman, Professor ved University of the Witwatersrand og en af respondenterne vi har interviewet, som i Masakhanes opstartsperiode bidrog med faglig sparring. Som det også fremgår af forskningsfeltet, er det dog ofte en udfordring at tiltrække og fastholde de nødvendige kompetencer, idet mange højt kvalificerede forskere forlader kontinentet – problematiseret ved begrebet 'brain drain', der bidrager til en strukturel ulige fordeling i den globale forskningskapacitet.

Fonde og investorer

For at sikre bæredygtigheden af deres projekter søger Masakhane finansiering fra fonde og økonomiske institutioner, der prioriterer teknologisk innovation og udvikling i det afrikanske kontinent. Ifølge organisationens egne oplysninger modtog de eksempelvis 150.000 USD i støtte i deres første år som registreret enhed (Masakhane Research Foundation 2024: 8). Denne form for finansiering understreger betydningen af globalt samarbejde og støtte til initiativer, der arbejder for øget repræsentation og inklusion i udviklingen af AI. Samtidig må det dog anerkendes, at sådanne ressourcer og støtteordninger i stigende grad påvirkes af den aktuelle geopolitiske kontekst, hvor prioriteringer og magtforhold kan udfordre kontinuiteten og omfanget af denne type engagement.

Tech-virksomheder

Teknologivirksomheder med fokus på kunstig intelligens og maskinlæring spiller en central rolle i udviklingen af den nødvendige infrastruktur og de tekniske værktøjer til opbygningen af sprogmodeller. På den ene side fungerer sådanne aktører som støttende elementer i Masakhanes arbejde, blandt andet ved at stille platforme og ressourcer til rådighed, hvor Masakhanes modeller kan integreres og anvendes. På den anden side optræder de dominerende tech-giganter samtidig som konkurrenter, særligt i bestræbelserne på at inkludere flere sprog i kommercielle sprogmodeller. Masakhane positionerer sig eksplicit som et alternativ til disse aktører, hvilket tydeliggøres i deres strategi, hvor de fremhæver

deres ikke-kommercielle og fællesskabsdrevne tilgang: "(...) we present an alternative to the only-for-profit or corporate models that are driving technological advances globally." (Masakhane Research Foundation, 2024: 15).

5.2 Kapabilitetsteori

I dette afsnit præsenterer vi den normative ramme, der danner grundlag for vores stillingtagen til retfærdighed, ulighed og lighed i forbindelse med vores undersøgelse. Formålet er at tydeliggøre, hvilke idealer vi vurderer teknologiudviklingen og dens konsekvenser ud fra, og hvordan vi forstår retfærdighed i en global teknologisk kontekst. Vi vælger at tage udgangspunkt i Amartya Sens kapabilitetsteori, der udgør et alternativ til traditionelle mål for udvikling og velfærd baseret på økonomisk vækst, præferencer eller formel adgang til ressourcer.

Sen definerer retfærdighed ud fra menneskers reelle muligheder, deres 'kapabiliteter', for at leve det liv, de har grund til at værdsætte (Sen, 1993: 32). Det centrale er ikke blot, hvilke ressourcer eller muligheder folk har til rådighed, men hvad de faktisk kan gøre og opnå med dem, altså hvordan de kan handle og udfolde sig i praksis, givet de betingelser de lever under. Et teknologisk redskab som en stor sprogmodel er for eksempel først en kapabilitet, når den kan omsættes til reel adgang til viden, ytringsfrihed, deltagelse eller indkomst. Sen skelner mellem kapabiliteter, altså de reelle muligheder og friheder, og 'functionings', som kan oversættes til 'funktioner', de aktiviteter og tilstande, som mennesker faktisk kan realisere i deres liv. Ifølge Sen vil der være en forskel mellem en person, der vælger ikke at bruge en teknologi, og en, der ikke har mulighed for det. Det er afgørende for, om vi kan tale om ulighed, at frihed forstås som både mål og middel for retfærdighed, og som socialt og kulturelt indlejret. Sen insisterer dermed på at flytte analysen væk fra eksempelvis indkomst eller nytte, og mod det konkrete liv mennesker kan efterleve (Sen, 1993, s. 33).

Teorien er relevant i vores analyse af generativ AI og global ulighed, hvor systemer som store sprogmodeller ofte er trænet på vestligt forankrede datasæt og fungerer bedst for lande i Vesten. Sens begreb 'conversion factor', som vi oversætter til 'konversations faktorer', tydeliggør, hvordan teknologiens formelle tilgængelighed ikke nødvendigvis fører til reelle muligheder, hvis sociale, kulturelle eller infrastrukturelle forhold står i vejen, som sprogbarrierer, marginalisering eller mangel på lokal adgang. Her viser Sens tilgang sin styrke i at fremhæve, hvordan samme teknologi kan føre til forskellige grader af frihed alt efter den enkeltes kontekst: "The conversion of income into basic capabilities may vary

greatly between individuals and also between different societies." (Sen 1993, s. 41). Eksempelvis hvis to brugere har adgang til den samme AI-teknologi, en i Danmark og en i Sydafrika, men kun den dansktalende kan bruge systemet effektivt på sit modersmål. Den afrikaans-talende støder derimod på sproglige barrierer. Ifølge Sen er det ikke reel lighed, fordi det ikke er adgangen i sig selv, men muligheden for at handle og opnå mål der er afgørende (Sen 1993: 46). Teknologien giver kun frihed, når den er brugbar på brugerens egne vilkår.

Sen advarer desuden imod at reducere muligheder til blot at være en form for ressource eller middel, sådan som man ser det i andre retfærdighedsteorier - som John Rawls' ide om 'primære goder' (Sen 1993: 82). For Sen er det ikke nok at have adgang til teknologiske eller økonomiske ressourcer, hvis man ikke reelt har mulighed for at bruge dem på en måde, der giver mening i ens liv. I vores sammenhæng betyder det, at en stor sprogmodel kun udgør en reel mulighed, hvis man kan anvende den på sit eget sprog og i sin egen kontekst, og ikke bare hvis man formelt har adgang til den. Der kan være andre behov i lavindkomstlande, som skal opfyldes, før digital infrastruktur prioriteres. Samtidig understreger Sen, at spørgsmålet om retfærdighed ikke bør begrænses af, hvad der er teknisk eller politisk muligt her og nu (Sen 2015: 80). Det åbner for at se sproglig inklusion og retfærdighed i AI som noget, vi har grund til at stræbe efter, også selvom det kræver langsigtet forandring i teknologi og magtstrukturer. Her bliver Sens teori ikke kun en ramme for vores vurdering af eksisterende ulighed, men også en invitation til at tænke fremad mod, hvordan mere lige adgang til teknologi kan gøres mulig gennem social og politisk handling.

Normative afvejninger

Martha Nussbaum (2003) videreudvikler Sens teori, men adskiller sig på flere vigtige punkter. Hvor Sen afholder sig fra at fastsætte en universel liste over relevante kapabiliteter, vælger Nussbaum eksplicit at opstille en liste over ti grundlæggende kapabiliteter. Listen er tænkt som et minimum, der bør garanteres alle mennesker, og som ifølge hende bør skrives ind i nationale forfatninger (Nussbaum 2003: 44-46). Hun lægger dermed en langt stærkere normativ ramme ned over kapabilitetstilgangen og trækker tydeligt på en etik baseret på menneskelig værdighed. Hvor Sen vægter procedurer og lokal demokratisk deltagelse som grundlag for udvælgelsen af kapabiliteter, vægter Nussbaum mere faste etiske principper.

Dette har givet anledning til en skelnen i litteraturen mellem to spor: en filosofisk tilgang af Nussbaum, hvor kapabiliteter udledes normativt, og en procedural tilgang med Sen, hvor udvælgelsen bør ske kontekstuelt og i dialog med de berørte grupper (Nussbaum 2003: 34). Deres forskelle er tydelige i spørgsmålet om paternalisme, som vi blandt andet undersøger i hvilket udstræk tech-industrien har magt over stat, virksomhed og individ.

Retfærdighed og ansvar

Endelig gør Sens tilgang det muligt at koble spørgsmål om retfærdighed til ansvar. Ved at analysere, hvem der faktisk får adgang til teknologiens potentialer, og hvem der systematisk ekskluderes, bliver det muligt at stille spørgsmål ved, hvordan teknologisk udvikling former globale magtforhold (Sen 2006: 215). I stedet for at vurdere AI ud fra præcision eller nytte, insisterer vi på, at teknologiske konsekvenser bør vurderes ud fra, om de udvider eller begrænser menneskers handlemuligheder, og om de skaber reel frihed for de grupper der hidtil har været teknologisk marginaliseret.

Et centralt aspekt i Sens teori er, at retfærdighed altid må forstås i den kontekst, den udfolder sig i, det vil sige, at hvad der er retfærdigt i en situation, ikke nødvendigvis er det i en anden. Spørgsmål om retfærdighed må tage højde for, at de relevante problemstillinger kan variere afhængigt af, om man ser på et rigt land i Europa i dag eller på det postkoloniale samfund i Afrika, Asien eller Latinamerika. Det er vigtigt at understrege, at disse postkoloniale samfund på ingen måde er ens, men hver især har deres egne unikke historiske og sociale udfordringer (Sen 2015: 86).

Nussbaums teori er universel og dækker alle mennesker, uanset om de lever i et liberalt demokrati eller i en undertrykkende stat, og uanset om de er i stand til at handle autonomt. Hun kritiserer Sen for ikke at tage tilstrækkeligt normativt stilling og for at undlade at specificere, hvilke kapabiliteter der bør prioriteres, og hvilket minimumsniveau et retfærdigt samfund bør sikre "(...) because of Sen's reluctance to make commitments about substance (which capabilities a society ought most centrally to pursue), even that guidance remains but an outline.'' (Nussbaum 2003: 35). Hendes fokus er på et niveau hvor retfærdighed ikke er opfyldt. Når først dette minimumsniveau er nået, tager hendes teori dog ikke eksplicit stilling til, hvordan retfærdighed skal vurderes derefter. Med Sens tilgang, der afviser en idealteoretisk model, kan vi anvende teorien som en praktisk tilgang, hvor formålet er at kunne sammenligne grader af uretfærdighed og arbejde hen imod et mindre uretfærdigt samfund (Sen 2006: 217). For Sen handler retfærdighed om at skabe betingelser for konkret forbedring og offentlig debat. I den forstand er hans tilgang åben og fleksibel og tillader

forskellige samfund selv at fastlægge, hvilke kapabiliteter der bør prioriteres.

En vigtig forskel mellem de to ligger også i deres behandling af ansvar og pligter. Nussbaum insisterer på, at alle mennesker skal have adgang til visse fundamentale kapabiliteter, men det er uklart, hvem der præcist bærer ansvaret for at sikre disse rettigheder, ud over en generel henvisning til statens forpligtelser (Nussbaum 2003: 37). Sen forholder sig i højere grad til spørgsmålet om fordeling af retfærdighed og ansvar, 'comparative justice', men uden at fastlægge en bestemt regel. Ifølge Sen bør en teori om retfærdighed ikke blot identificere, hvad der er ideelt, men støtte individer og samfund i at tage ansvar for at mindske åbenlyse uretfærdigheder, hvor de forekommer (Sen 2006: 218-219). Han advarer imod at fokusere for meget på moralsk perfektion og opfordrer i stedet til at identificere konkrete sociale uretfærdigheder og handle på dem.

Kapabilitetsteoriens anvendelighed i vores undersøgelse

Selvom kapabilitetstilgangen hovedsageligt er udviklet i vestlige filosofiske traditioner, rummer især Sens version en åbenhed over for kontekstspecifikke værdier og livsformer. Det gør tilgangen velegnet til at analysere retfærdighed i ikke-vestlige samfund, herunder i en sydafrikansk kontekst, hvor kapabiliteter som kollektivt selvstyre, sproglig deltagelse og teknologisk inklusion kan anskues som kulturelt betingede og relationelle. Sen anerkender netop, at kapabiliteters værdi ikke er universel, men afhænger af sociale og politiske betingelse (Sen 2006: 229). På den måde kan kapabilitetstilgangen også understøtte anerkendelsen af kollektive rettigheder og ikke-vestlige idealer om velfærd og værdigt liv.

Kapabilitetsteorien har traditionelt været anvendt inden for udviklingsetik, sundhedsretfærdighed og uddannelsespolitik, hvor fokus har været på menneskelig trivsel, agens og frihed i bred forstand (Sen 2015: 77-78). I vores undersøgelse aktiverer vi teorien i en anden sammenhæng, nemlig som normativ ramme for at vurdere ulighed i teknologisk sprogudvikling og adgang til generativ AI. Vi undersøger, hvordan teknologier som store sprogmodeller enten udvider eller begrænser menneskers reelle handlemuligheder, og hvordan sociale, sproglige og infrastrukturelle faktorer former adgangen til disse teknologier i en sydafrikansk kontekst.

Ved at anvende kapabilitetstilgangen på dette felt, bliver det muligt at stille spørgsmål ved, om teknologisk udvikling bidrager til menneskelig værdighed og frihed, eller om den snarere fastholder eksisterende globale magtstrukturer, som vi også holder koloniale teorier op mod (jf. afsnit 5.4, 5.5). Det gælder særligt i vores case, hvor vi bruger teorien til at vurdere, hvilke kapabiliteter der er nødvendige for at deltage i og påvirke AI-udviklingen, og

hvordan manglende sproglig inklusion, ressourcer eller politisk indflydelse kan udgøre en kapabilitetsmangel. Dermed bliver teorien ikke blot et redskab til at måle udvikling, men en kritisk ramme for at identificere, hvor teknologisk ulighed opstår, og hvem den rammer.

Rawls' retfærdighedsbegreb

John Rawls' forståelse af retfærdighed komplementerer således kapabilitetsteoriens fokus på at fremme menneskelig frihed, værdighed og muligheder for alle, særligt i relation til globale uligheder i AI-udviklingen (Tsou & Walsh 2023: 8). Rawls' teori om retfærdighed kan bidrage med en relevant forståelse af 'fairness' og fordeling i udviklingen af AI og dataetik. Som præsenteret af Tsou og Walsh bygger Rawls' teori på ideen om en oprindelig position bag et 'slør af uvidenhed', hvor retfærdige principper fastlægges uden kendskab til egen position i samfundet. Rawls' lighedsprincip og differensprincip understreger, at alle har lige grundlæggende rettigheder, og at uligheder kun er acceptabel, hvis de er til fordel for de mest dårligt stillede (Tsou & Walsh 2023: 11). Selvom Rawls' teori ikke er det primære teoretiske fundament for specialet, kan hans begreber understøtte vores normative ramme ved at pege på, at udvikling og implementering af AI ikke blot er tekniske valg, men etiske beslutninger med konsekvenser for, hvem der får adgang til teknologiens fordele – og hvem der risikerer at blive ekskluderet.

Et konkret eksempel er spørgsmålet om kryptering og privatlivsbeskyttelse, hvor der er en konflikt mellem individers ret til at beskytte deres privatliv og statens ret til at regulere teknologi for at sikre offentlig sikkerhed. Her kan Rawls' teori om retfærdighed anvendes til at argumentere for, at individernes rettigheder bør prioriteres, da de er den mest sårbare gruppe. En vigtig begrænsning ved disse teorier er, at Rawls' tilgang primært fokuserer på social retfærdighed og statens rolle i liberale demokratier.

5.3 Dataetik

Eftersom specialet undersøger, hvordan udviklingen af AI og sprogmodeller forstærker globale ulighedsstrukturer, er det nødvendigt at placere et dataetisk perspektiv. Dataetik er et relativt nyt begreb, der endnu ikke er klart defineret eller konsensusbaseret (Thomsen & Skou Olsen 2021: 4). I dette speciale anvendes dataetik som del af vores normative rammer, der bygger på værdier og principper for, hvordan data bør behandles, uanset hvor teknologien udvikles eller anvendes (Thomsen & Skou Olsen 2021: 11). Dataetik handler ikke kun om de etiske implikationer for mennesker, men også om, hvordan vi forholder os til de data og systemer, som AI og sprogteknologi bygger på.

Vi inddrager Luciano Floridis teoretiske framework, til at forstå hvordan dataetik placeres i praksis. Floridi og Cowls beskriver AI som en del af en større 'infosphere', et begreb, der dækker over det samlede informationsmiljø, hvor mennesker, data, teknologier og systemer er forbundne i en kompleks helhed. Etiske spørgsmål handler derfor ikke kun om individers rettigheder og handlinger, men også om de strukturer og relationer, der opstår. I denne forståelse bliver dataetik et spørgsmål om, hvordan vi forvalter dataansvar og sikrer retfærdighed i den måde, AI udvikles, distribueres og anvendes på tværs af globale uligheder (Floridi & Cowls 2019: 2). Floridi og Cowls er optaget af, at AI ikke blot bør minimere skade, men aktivt skabe værdi for samfundet. De peger på, at der lige nu er en bred række af foreslåede initiativer fra forskellige organisationer, som kan blive overvældende og uforståelige. De peger på, at det kan lede til to problemer: Enten er de forskellige sæt etiske principper for AI ens, hvilket fører til unødvendig gentagelse og redundans, eller, hvis de adskiller sig markant, vil der opstå forvirring og uklarhed. Det værste resultat ville være et "marked for principper", hvor interessenter kan blive fristet til at "shoppe" efter de mest tiltalende principper (Floridi & Cowls 2019: 2).

Floridi og Cowls præsenterer fem principper, som er udformet for at skabe et samlet etisk rammeværk for AI i samfundet. Disse principper blev senere inkorporeret i EU-Kommissionens Ethics Guidelines for Trustworthy AI. Principperne nedenfor giver derfor en struktureret tilgang, som kan anvendes til at udforme politikker, bedste praksis og andre anbefalinger relateret til udvikling og anvendelse af AI (Floridi & Cowls 2019: 8).

De fem principper for etisk teknologivurdering:

1. **Beneficence**: 'Velgørenhed' eller 'gavnlighed'. AI skal kunne fremme velfærd, handle i deres bedste interesse og sikre, at AI har en positiv indvirkning på mennesker og planeten (Floridi & Cowls 2019: 6).
2. **Non-maleficence**: 'Ikke-skadelighed' eller 'skadeforhindring'. Det refererer til at vi skal undgå at teknologien vil påføre skade eller lidelse på andre ved misbrug eller overbrug af den (Floridi & Cowls 2019: 6).
3. **Autonomy**: 'Autonomi'. Individets ret til at træffe egne beslutninger og handle uafhængigt indebærer, at personen altid skal have mulighed for at vælge, hvornår AI skal anvendes, og hvornår menneskelig autonomi skal genvindes (Floridi & Cowls 2019: 7).

4. **Justice**: ‘Retfærdighed’. Refererer til behovet for at sikre lige adgang til AI's fordele og bekæmpe diskrimination. Det indebærer også at undgå skævheder i datasæt, der træner AI, og at beskytte mod trusler mod social solidaritet og fælles velstand (Floridi & Cowls 2019: 7).
5. **Explicability** – ‘Forklarlighed’. Beslutninger og handlinger, især i forbindelse med teknologi, bør være forståelige og kunne forklares på en gennemsigtig måde. Det handler om både intelligibilitet (hvordan virker AI?) og ansvarlighed (hvem er ansvarlig for, hvordan det virker?). Det er essentielt for AI-etik, da det muliggør forståelse af både de gode og dårlige konsekvenser af AI, samt ansvarlighed i tilfælde af negative udfald.(Floridi & Cowls 2019: 8).

Princippet om ‘beneficence’ lægger vægt på, at AI skal gavne alle, herunder de mest marginaliserede grupper (Floridi & Cowls 2019: 2). Dette er særligt relevant for vores analyse af, hvordan AI påvirker globale magtstrukturer og risikerer at forstærke eksisterende uligheder, hvis ikke udviklingen sker ansvarligt. Det er også tæt knyttet til vores case om Masakhane, som netop arbejder for at sikre, at folk i majoritetsverden, særligt dem, der taler afrikanske sprog, bliver repræsenteret i udviklingen af sprogmodeller og får adgang til teknologiens muligheder. Floridi et al. (2018) fremhæver desuden, at mangfoldighed, herunder køn, klasse, etnicitet, disciplin og geografisk repræsentation, bør integreres i AI’s design- og udviklingsprocesser for at øge inklusivitet og sikre retfærdighed i teknologiens anvendelse (Floridi et al. 2018: 20).
Dog påpeger Floridi og Cowls, at de fem principper primært er udviklet ud fra vestlige og internationale perspektiver, og derfor ikke nødvendigvis afspejler synspunkter fra regioner og kulturer, der i dag er underrepræsenterede i AI-udviklingen (Floridi & Cowls 2019: 8). De fremhæver vigtigheden af, at etik ikke må være forankret i én kultur eller kontekst, men bør afspejle geografisk, kulturel og social mangfoldighed. Derfor understreger de, at alle aktører – herunder virksomheder, offentlige institutioner og forskningsmiljøer – har et ansvar for at følge etisk rammesætning, der også tager højde for disse forskelle.

Floridi og Cowls argumenterer desuden for, at etableringen af en socialt ønskværdig kurs for AI ikke alene afhænger af regulering og fælles standarder, men i høj grad af brugen af etiske principper som konkret rettesnor (Floridi & Cowls 2019: 10). Rammeværket, de præsenterer, skal fungere som en værdifuld struktur for at sikre positive sociale resultater af AI-teknologi og for at skabe overgangen fra abstrakte principper til praktisk anvendelse.

Floridi et al. (2018) understreger samtidig, at arbejdet med etiske principper ikke kan stå alene, men bør suppleres af politiske processer og lovgivning, som understøtter inkluderende og retfærdige beslutninger om AI's udvikling og anvendelse (Floridi et al. 2018: 22).

5.4 Datakolonialisme

Nick Couldry og Ulises A. Mejias (2019a, 2019b) introducerer begrebet datakolonialisme som en måde at forstå kapitalismens koloniale videreførelse, hvordan dataficeringen af menneskers liv i dag udgør en ny form for kolonial udvinding og dominans. Ligesom traditionelle koloniale europæiske magter udnyttede ressourcer i majoritetsverden, så udnytter AI og tech-giganter fra især USA og Kina profit for egen vinding, i en totalitær vision: "(…) the bipolar world of U.S.–China data domination, colonialism starts to operate well beyond the confines of European traditions" (Couldry & Mejias 2019a: 346).

Ved at se dataudvinding i et 500-årigt perspektiv, bliver det tydeligt, at udviklingen ikke er et afvigende fænomen, men en ny fase i kapitalismens historie, direkte forbundet til de økonomiske og sociale transformationer, som kolonialismen muliggjorde (Couldry & Mejias 2019b: 3). Datakolonialisme placeres dermed i forlængelse af koloniseringen af land og kroppe i det 15. og 16. århundrede. Couldry og Mejias forsøger at udforske parallellerne til den historiske kolonialistiske udvikling af globale økonomier, samt dens normalisering af at rive ressourcer til sig og dens omdefinering af sociale relationer, så alt sammen kom til at virke naturligt, at blive frataget sine rettigheder (Couldry & Mejias 2019a: 339).

Datarelationer og vidensdomians

Et centralt begreb i datakolonialismens logik er det Couldry og Mejias kalder for 'datarelationer', der muliggør socialt liv på globalt plan og bliver en åben kilde som er frit tilgængelig for kapital. Menneskers hverdagsliv, bevægelser, følelser og sociale relationer bliver bearbejdet, rangeret og solgt for profit: "(...) a privileged 'window' onto the world of social relations, and a privileged 'handle' on the levers of social differentiation." (Couldry & Mejias 2019b: 4). Derfor handler datakolonialisme ikke kun om sociale medier eller tech giganter, det er et bredere skifte, hvor menneskers liv og sociale relationer omdannes til råmateriale for kapital (Couldry & Mejias 2019b: 5). Udviklingen gøres mulig, fordi data fremstilles som noget universelt og ejerskabsløst, der kan høstes bæredygtigt til fordel for menneskeheden. Denne ideologi legitimerer den globale dataudvinding og forskyder fokus væk fra de etiske og sociale konsekvenser. Det legitimeres dermed gennem en ide om

Vestens videndominans, hvor 'Big Data' har fremstået som en sandhedskilde i teknologiudviklingen (Couldry & Mejias 2019b: 12).

Det centrale her er, at denne subjektiveringsproces af individer og samfund reduceres til 'data subjekter', ofte foregår gennem en kombination af afhængighed, vaner og infrastruktur og at de sociale konsekvenser ikke rammer ens. Denne forståelse af viden, og magt som nogen har over andre, marginaliserer alternative perspektiver og erfaringer, især fra majoritetsverden, hvor datamangel ofte misforstås som mangel på viden, snarere end som resultat af ulige infrastruktur og historieskabt ulighed (Couldry & Mejias 2019a: 345).

Afkolonisering i data- og teknologiforskning

I forlængelse af de to tekster fra 2019 bygger Couldry og Mejias (2023) videre på begrebet datakolonialisme og placerer det i den bredere postkoloniale vending i teknologi- og dataforskning. Her argumenterer de for, hvordan datarelationer ikke kun handler om økonomisk udvinding, men også om race, viden og global ulighed. (Couldry & Mejias 2023: 792). Denne pointe understreger, at AI og datasystemer ikke blot viderefører, men også intensiverer racedifferentiering, fordi de bygger på asymmetriske infrastrukturer og bias i dataindsamling og bearbejdning. Samtidig advarer Couldry og Mejias mod at tro, at dataficering udfolder sig ens overalt og fremhæver, at dataudvinding foregår forskelligt i det minoritets- og majoritetsverden. Tech-giganter skaber ofte infrastruktur i majoritetsverden med henblik på fremtidig dataudnyttelse, for eksempel ved at tilbyde gratis internetadgang, cloud-løsninger eller AI-ekspertise: "This is an important point about the unevenness of how data colonialism may spread globally." (Couldry & Mejias 2023: 793).

Couldry og Mejias (2023) argumenterer også for, at modstanden mod datakolonialisme må være strukturel, ikke blot rettet mod enkelte aktører eller teknologier. De kalder på en afkolonisering af de teknologiske rationaler og epistemiske systemer, som legitimerer dataficeringen: "Solutions to the harm done by data colonialism therefore require profound societal change, and not just the reform of particular corporations." (Couldry & Mejias 2023: 796). Afkolonisering betyder her at gentænke, hvordan data og teknologi bruges, styres og forstås, ikke ud fra vestlige idealer om fremskridt og objektivitet, men ud fra globale og kollektive forestillinger om retfærdighed, frihed og autonomi. Forfatterne understreger vigtigheden af, at modstand mod dataekstraktivisme også er lokal og inkluderer de mest marginaliserede, herunder oprindelige folk og racialiserede grupper (Couldry & Mejias 2023: 798). De afslutter med en fremtidsvision om at udvikle et nyt konceptuelt rum

for teknologisk udvikling, et rum, der bryder med Silicon Valleys profit-motiveret logik og det kinesiske kommunist partis, kontrol-motiveret model (Couldry & Mejias 2023: 798).

5.5 Teknologisk kolonialisme

Kolonialismens arv i digitaliseringen

Rachel Adams introducerer i *The New Empire of AI* (2025) begrebet 'majoritetsverden' som en bevidst modsætning til tidligere betegnelser som 'tredje verden', 'udviklingslande' eller 'industrialiserende lande', også kendt som det globale syd. Med termen majoritetsverden ønsker Adams at anerkende, at størstedelen af verdens befolkning lever uden for det privilegerede Vesten, men samtidig undgå de nedvurderende konnotationer, der har præget tidligere begreber (Adams 2025: 18). 'Minoritetsverden' betegner i Adams' terminologi ikke marginaliserede minoritetsgrupper, men derimod den lille del af verdens befolkning i Vesten, som lever i det mest velstillede lande - typisk kaldet det globale nord (Adams 2025: 19). Vi benytter os af begreberne majoritetsverden og minoritetsverden konsekvent i opgaven for at tydeliggøre de globale uligheder, som Adams' terminologi indfanger.

Ifølge Adams bør AI ikke forstås som en neutral teknologisk udvikling, men som en ny form for magtudøvelse, der skaber en ny verdensorden. Hun beskriver, hvordan AI-industrien koncentrerer magt på en måde, der ikke følger de klassiske mønstre fra europæisk kolonialisme, men som alligevel bygger videre på mange af de samme ulighedsstrukturer. Europas rolle som globalt centrum er aftaget, mens globale techgiganter og stormagter i både Vesten og Østen nu konkurrerer om teknologisk dominans (Adams 2025: 16). Denne udvikling betegner Adams som det 'nye AI-imperium', en modalitet af magt, som ifølge hende er endnu mere gennemgribende end de historiske imperier.

I forhold til majoritetsverdenen advarer Adams om, at AI's effekter på arbejdsmarkedet misforstås. Når diskussioner om AI fokuserer på jobtab, tager de ofte udgangspunkt i formelle økonomier, hvilket er irrelevant for store dele af den globale befolkning, hvor størstedelen arbejder i den uformelle sektor (Adams 2025: 111). I Afrika, hvor analfabetisme fortsat er udbredt, eksempelvis blandt 25% af mænd og 40% af kvinder i lande der ligger syd for Sahara Ørkenen, vil opfordringer til blot at tilegne sig nye færdigheder i en AI-økonomi efterlade mange yderligere marginaliserede (Adams 2025: 107).

Adams knytter også udviklingen af AI tæt til kolonialismens historiske arv. Hun påpeger, at europæisk kolonialisme skabte dybt rodfæstede globale uligheder, som fortsat

præger den teknologiske fordeling i dag (Adams 2025: 14). Kolonialismen var først og fremmest et økonomisk projekt, hvor ressourcer blev udvundet fra kolonierne til fordel for imperiemagterne. Ifølge Adams er det ikke tilfældigt, at mange lande i den globale majoritet stadig mangler ressourcerne til at opbygge egne AI-økosystemer, denne mangel er en direkte arv fra kolonialismen. Hun fremhæver, at mens teknologien accelererer ulighederne i dag, er forbindelsen mellem koloniale historiske strukturer og nutidens AI-kløft stort set ubehandlet i både forskning og politik (Adams 2025: 15).

Teknologisk parathed

Adams peger på, at teknologi historisk set har været anvendt som et redskab til imperial magtudøvelse, hvor kolonimagter brugte teknologiske forspring til at kontrollere og udnytte andre regioner (Adams 2025: 15). I dag videreføres disse uligheder i en ny form gennem kunstig intelligens. AI's hurtige udvikling i Vesten og det rige Østen betyder, at kløften mellem regionerne vokser i hastigt tempo, og at lande i den såkaldte majoritetsverden risikerer at blive efterladt endnu længere bagud (Adams 2024: 17).

Særligt i Afrika udgør manglende adgang til internet, elektricitet og computerkapacitet væsentlige barrierer for at udvikle og anvende AI på egne vilkår. Adams introducerer her begrebet 'absorptive capacity', eller 'optagelseskapacitet', som et udtryk for, hvorvidt en region har de nødvendige infrastrukturelle forudsætninger for overhovedet at kunne absorbere og anvende avanceret teknologi (Adams 2025: 76). I dag har majoritetsverden under 3% af verdens samlede computerkapacitet, og Afrika, med næsten 1,5 milliarder indbyggere, råder blot over 0,2% (Adams 2025: 63).

Det skaber en situation, hvor lande som Sydafrika, trods en relativt stærk digital sektor, stadig kæmper med udbredt fattigdom, manglende internetdækning og energikriser, som f.eks. udbredt elektricitet-tyveri (Adams 2025: 69). Disse strukturelle problemer begrænser ikke alene mulighederne for teknologisk innovation, men fastholder Afrika i en position som afhængig forbruger af teknologi udviklet i andre regioner. Adams understreger, at ansvaret for denne ubalance ikke kan placeres hos de marginaliserede, men må bæres af dem, der i dag høster de største gevinster af AI's udvikling (Adams 2025: 125).

Underrepræsentation af sprog i store sprogmodeller

Rachel Adams fremhæver, hvordan store sprogmodeller ikke blot spejler de eksisterende skævheder i sproget, men aktivt forstærker dem. Fordi LLM'er primært trænes på engelsksprogede data, viderefører de stereotype associationer – eksempelvis koblinger

mellem ‘Afrika’ og ‘primitivitet’ eller mellem ‘kvinde’ og ‘emotionel’ – som allerede findes online (Adams 2025: 142). Når nye tekster genereres af LLM’er og derefter anvendes i fremtidige træningsdata, skabes en selvforstærkende cyklus af bias.

Adams kritiserer, at tech-virksomheder har prioriteret at udvikle større og større modeller frem for at håndtere de fejl, som har været indbygget gennem træning af fejlagtige datasæt. Fokus har ligget på kvantitet frem for kvalitet, og det etiske arbejde med datasæt er i vid udstrækning blevet outsourcet eller nedprioriteret (Adams 2025: 144).

Selvom LLM’er ofte fremstår som politisk neutrale, påpeger Adams, at denne neutralitet dækker over en vestlig, liberal verdensopfattelse (Adams 2025: 145). Med inspiration fra Foucault fremhæver hun, at sproget ikke er neutralt: Sprog former materielle realiteter, og når AI får monopol på sproglig kommunikation, bærer det en kompleks og hidtil forsømt ansvarlighed overfor ikke-vestlige og ikke-engelsktalende kontekster (Adams 2025: 147-148). Sprogets magt i AI er ikke kun teknisk, men kulturel og politisk. Ifølge Adams fortsætter engelsk-dominante LLM’er den imperialistiske udbredelse af vestlige normer og idéer, idet sprog bærer både kultur og verdenssyn (Adams 2025: 149). Skal AI afkoloniseres, må der skabes og trænes sprogmodeller på afrikanske sprog, som projekter som Lelapa AI og Masakahne allerede arbejder for (Adams 2025: 150).

5.6 Kritisk diskursanalyse

For at undersøge, hvordan AI og diskurser formes og reproduceres, inddrager vi Norman Faircloughs (2003) kritiske diskursanalyse som teoretisk og analytisk ramme. Hvor de foregående afsnit har belyst AI som teknologisk fænomen og teoretiserer globale magtforhold og retfærdighedsprincipper, zoomer vi her ind på sproget som en central mekanisme, der både afspejler og opretholder magtrelationer. Vi bruger det til at se på, hvordan diskurser om AI ikke blot beskriver en teknologisk udvikling, men bidrager til at forme opfattelser af ansvar, magt og muligheder i en global kontekst. Med Faircloughs perspektiv kan vi analysere, hvordan bestemte fortællinger og repræsentationer af AI skaber mening og legitimitet, og hvordan disse fortællinger ikke er neutrale, men præget af ideologier og magtstrukturer.

Faircloughs tredimensionelle model kombinerer analyse af tekst, diskursiv praksis og sociokulturel praksis. Modellen består af tre niveauer:

1. Tekstanalyse (beskrivelse): Tekstens sproglige træk, fx ordvalg, grammatik og struktur. Målet er at identificere mønstre, der kan pege på underliggende ideologier eller magtrelationer.
2. Diskursiv praksis (fortolkning): Denne del handler om, hvordan teksten produceres, distribueres og forstås i en konkret kontekst. Vi undersøger fx, hvem der taler, til hvem, og hvilke andre tekster der refereres til.
3. Sociokulturel praksis (forklaring): Til sidst placerer vi diskursen i en bredere samfundsmæssig sammenhæng. Her analyserer vi, hvordan sproget afspejler, opretholder eller udfordrer sociale strukturer, magtforhold og ideologier (Fairclough 2003: 23).

Diskurs, governance og hegemoni

Diskurs er en måde at repræsentere virkeligheden på. Forskellige diskurser kan konkurrere om at definere den samme del af verden fra forskellige perspektiver. Repræsentation er dermed ikke neutral, men bundet til magt og ideologi (Fairclough 2003: 26). Fairclough peger på, at diskursive processer ofte indebærer, at særlige identiteter, interesser eller repræsentationer fremstilles som almene. Brugen af digitalt sprog er ikke kun en neutral afspejling af virkeligheden, men med til at forme sociale relationer og magtbalancer gennem tekstlige strategier og diskursive tendenser.

Fairclough forstår 'governance' i en bred forstand som enhver institutions- eller organisationsbaseret aktivitet, der har til formål at regulere eller styre andre sociale praksisser. I hans optik er governance ikke kun et spørgsmål om hierarkisk statslig styring, men består af et samspil mellem markeder, netværk og bureaukratier (Fairclough 2003: 32). Dette er ifølge ham et spørgsmål om hegemoni, altså om, hvordan bestemte sociale grupper opnår og opretholder dominans ved at fremstille bestemte interesser og identiteter som universelt (Fairclough 2003: 41). Diskurs fungerer som et redskab for ideologisk arbejde ved at fremme bestemte betydninger som almengyldige, for derigennem at opnå og fastholde dominans. Når tekster præsenterer særlige forståelser, f.eks. en global økonomi eller teknologisk fremtid, som en uomgængelig og ubetvivlelig realitet, udfører de ideologisk arbejde ved at skjule, at dette blot er en version af virkeligheden (Fairclough 2003: 58). Dermed kan diskurser vedrørende AI og teknologi udgøre et ideologisk arbejde, når de fremstår som objektive sandheder og dermed skjuler deres partikulære karakter.

Magt og agens

Fairclough betragter sprogbrug som et middel til at udøve magt. Magt handler i hans optik grundlæggende om kapaciteten til at handle, om hvem der har agens. Agens er ikke en abstrakt egenskab, men tæt forbundet med adgang til ressourcer. Nogle aktører har større handlekraft, fordi de har adgang til teknologiske, økonomiske og politiske ressourcer, mens andre er strukturelt begrænsede (Fairclough 2003: 47). I vores speciale bliver dette relevant, fordi AI og sprogmodeller i sig selv er produkter af diskursive og teknologiske praksisser, der er skabt i specifikke kontekster. Når AI udvikles og implementeres globalt, sker det som en del af en kamp om, hvordan teknologi skal forstås, hvem der har adgang til den, og hvilke formål den skal tjene. Faircloughs begreb om magt som differentielt tilgængelige ressourcer synliggør, hvordan AI's udvikling og udbredelse ikke blot er tekniske processer, men også sociale og politiske.

Med Faircloughs perspektiv bliver det tydeligt, at sprog og diskurser ikke blot afspejler virkeligheden, men også former den. Når AI beskrives som effektiviserende, neutral eller objektiv teknologi, udføres der ideologisk arbejde, som skjuler de magtforhold og interesser, der ligger bag.

6. METODE

6. Metode

I det følgende afsnit præsenteres specialets metodiske refleksioner. Det indebærer en beskrivelse af undersøgelsesdesignet, dataindsamlingen, analysestrategien og de overvejelser, der knytter sig til interviewet som metode. Dernæst reflekteres der over kvalitetskriterierne i relation til projektets kvalitative tilgang. Med afsæt i de metodiske valg og det teoretiske grundlag udfoldes afslutningsvist projektets videnskabsteoretiske refleksioner.

6.1 Undersøgelsesdesign

Undersøgelsens metodiske tilgang er inspireret af Brymans (2008) beskrivelse af kvalitativt forskningsdesign, hvor flere datakilder kombineres for at opnå en dybere forståelse af et komplekst fænomen: "(...) we can understand social phenomena better when they are compared in relation to two or more meaningfully contrasting cases or situations." (Bryman 2008: 58). I specialet benytter vi et flerstrenget forskningsdesign, der kombinerer kvalitative ekspertinterviews med aktører i både Danmark og Sydafrika, et casestudie af græsrodsorganisationen Masakhane, samt dokumentanalyse af centrale politiske tekster om AI-regulering. Ekspertinterviewene trækker på elementer fra komparativ kvalitativ metode, idet vi sammenholder perspektiver fra forskellige geopolitiske kontekster uden at gennemføre en klassisk systematisk case sammenligning. Casestudiet følger Yin og Brymans forståelse af casestudiet som en intensiv undersøgelse af en enkelthed, som belyser bredere strukturelle dynamikker. Vi inddrager både EU AI Forordning og AU AI Strategy-dokumenter som adgang til hvem der har governance, og vi bruger dem under vores interviews til at understøtte undersøgelsen, med henblik på AI-regulering (Bryman 2008: 526). I tråd med Brymans pointe om, at politiske og institutionelle tekster kan analyseres for at afdække diskurser og magtforhold.

Det samlede design giver os mulighed for at undersøge, hvordan AI-regulering og udvikling af LLM'er både formes og udfordres i globale magtforhold på tværs af landegrænser, og hvordan aktører i Danmark og Sydafrika forholder sig til disse strukturer og problematikker.

6.2 Ekspertinterviews udvælgelseskriterier

Formålet med interviewene var at engagere forskellige eksperter for at opnå en helhedsforståelse af, hvilke faktorer der er på spil i udviklingen af AI-sprogmodeller og deres

potentielle påvirkning af global ulighed. Vi har udført 11 semistrukturerede interviews med seks danske og fem sydafrikanske aktører. Dette antal for at sikre en bred repræsentation af faglige baggrunde, institutionelle tilknytninger, geografisk forskellighed og et tilstrækkeligt bredt datagrundlag.

Vi skulle også beslutte, hvad en ekspert betød i vores forskning. Det har vi gjort gennem vores udvalgskriterier (Figur 2), for at sikre adgang til kvalificeret viden. Figur 2 viser et overblik over udvælgelseskriterierne for deltagelse i specialet:

Udvalgskriterier	**Beskrivelse**
Interviews	11 interviews
Industri	Bred repræsentation af eksperter, herunder akademikere, ph.d.-studerende, journalister, fagfolk, eller forfattere.
Institutionel tilknytning	Relevante organisationer, initiativer eller universiteter, der beskæftiger sig med AI, dataetik eller teknologisk regulering.
Geografisk placering	Folk baseret i Danmark eller Sydafrika.
Ekspert definition	Graden af teknisk viden og erfaring inden for generativ AI, LLM'er eller etisk teknologiudvikling.
Rolle i teknologidebat	Deres rolle og indflydelse i debatten om teknologisk udvikling og etik, f.eks. offentlige talere, policy makers, eller forskere.

Figur 2: Udvælgelseskriterier til interviews

Vi anvender ekspertinterviews som en central metode, hvor fokus er på den viden, respondenterne bringer med sig i kraft af deres ekspertrolle. Ifølge Brinkmann og Kvale handler denne type interview især om indholdet af det, der siges, frem for måden, det siges på. Samtidig bidrager både interviewer og respondent med faglig indsigt om kunstig

intelligens, hvilket skaber en mere ligeværdig og informeret dialog. Det er netop gennem disse interviews, at vi skaber vores empiri, som udgør fundamentet for analysen og giver undersøgelsen sin aktualitet og originalitet (Brinkmann & Kvale 2015: 201).

6.2.1 Semistrukturerede interviews

Interviewene blev udført af os begge gennem en semistruktureret tilgang, som beskrives af Brinkmann ofte som lighedstegnet til kvalitative interviews (Brinkmann 2013: 21). Dette format gav os mulighed for at udvikle og differentiere interviewguiden, hvilket var særligt relevant, da forskellene mellem eksperterne i Danmark og Sydafrika nødvendiggjorde tilpasninger i vores spørgsmål til respondenternes ekspertise (Bilag: Interviewguide).

For at opretholde fokus og muligheden for at stille opfølgende spørgsmål blev alle interviews optaget. Optagelserne fra Teams blev transskriberet direkte, mens andre blev optaget på memo-applikationer. Når vi begge var til stede under et interview, havde den ene af os til opgave at tidsstyre og introducere, mens den anden agerede som primær interviewer (Bryman 2008: 442)..

Fleksibiliteten i vores tilgang til at interviewe, var også reflekteret i måden vi mødte vores respondenter på, hvilket resulterede i et fysisk møde, et telefonmøde, og resten via digitale Teams møder. Denne tilgang gjorde det muligt for os at møde respondenterne på deres foretrukne sted, hvilket fremmede en mere åben og ærlig dialog. Vi tilpassede også rækkefølgen af spørgsmål baseret på samtaleforløbet, hvilket ofte førte til en mere fleksibel spørgeramme og sammenhæng i respondenternes svar (Bryman 2008: 456).

6.2.2 Præsentation af respondenter

Danske respondenter:

1. **Anette Høyrup**
   Medlem af Dataetisk Råd og seniorkonsulent ved Implement Consulting Group
2. **Frej Klem Thomsen**
   Seniorkonsulent ved Nationalt Center for Etik
3. **Ida Ebbensgaard**
   Tech-journalist og forfatter
4. **Laura Haaber Ihle**
   Vice President of Ethics, Governance & Policy ved Responsible AI Future Foundation

5. **Marie Høst**
   Tech-journalist og medievært
6. **Morten Sodemann**
   Klinisk Professor og forskningsleder ved Center for Sundhedsfilosofi og Etik ved Syddansk Universitet
7. **Sine Nørholm Just**
   Professor ved Roskilde Universitet og projekt- og forskningsleder ved Algoritmer, Data og Demokrati (ADD)

Sydafrikanske respondenter:

8. **Archana Arakkal**
   AI-specialist og medstifter af Synthesis Software Technologies
9. **Ayantola Alayande**
   Forsker ved Global Center on AI Governance
10. **Benjamin Rosman**
    Professor ved University of the Witwatersrand
11. **Hellen Mukiri-Smith**
    Selvstændig advokat inden for kunstig intelligens og dataregulering

6.2.3 Sampling

For at imødekomme den ofte kritiserede mangel på gennemsigtighed i samplingprocesser inden for kvalitativ forskning, har vi anvendt en bevidst og målrettet samplingstrategi. Vi stræber efter en strategisk sammenhæng mellem vores forskningsspørgsmål og den måde, vi har udvalgt vores datakilder på. Vi har derfor anvendt en 'purposive sampling' metode, hvor udvælgelsen af enheder, det være personer, organisationer, dokumenter eller relevante afdelinger, er direkte relateret til de specifikke spørgsmål, vi stiller i vores undersøgelse (Bryman 2008: 458).

I lyset af vores speciales fokus på sammenligninger mellem minoritets- og majoritetsverden, var det nødvendigt at balancere antallet af interviews mellem danske og sydafrikanske eksperter. Det har tilladt os at udføre en sammenlignende analyse med en ligelig repræsentation fra hvert geografisk område, hvilket er kritisk for validiteten af vores datagrundlag (Bryman 2008: 462).

6.2.4 Tematisk kodning af interviews

Til bearbejdelse af data fra vores kvalitative undersøgelse, der består af 11 ekspertinterviews, har vi anvendt en tematisk kodning som analytisk tilgang (Bryman 2008: 554). Denne åbne metode gør det muligt systematisk at identificere, organisere og fortolke mønstre i datamaterialet. Vores analyse bygger på et framework, hvor temaer og undertemaer er udviklet både med udgangspunkt i vores predefinerede temaer og spørgsmål i interviewguiden (bilag), og induktivt gennem grundige nærlæsninger af de transskriberede interviews. Den iterative bevægelse mellem teori og empiri har gjort det muligt at danne hovedtemaer og nuancere vores forståelse af materialet, hvorved vi kunne danne analysens fokus yderligere.

Vi har struktureret vores framework i en Excel-matrix, hvor interviewdata organiseres i et regneark, med respondenter i rækker og overordnede temaer som kolonner. Af de 11 interviews systematiserede vi citaterne i følgende fem kategorier: regulering, diskrimination & bias, etisk teknologiudvikling, ansvar, magt & ulighed (Bilag: Kodningsmatrix). I celler indsatte vi relevante citater, der illustrerer eller problematiserer temaernes indhold. Dette systematiske overblik understøtter vores analytiske processer og sikrer transparens i udvælgelsen af empiri. Tilgangen er inspireret af Bryman, som beskriver, hvordan tematiske analyser kan operationaliseres i praksis gennem kategorisering og sammenstilling af citatmateriale (Bryman 2008: 555).

6.3 Case Study

Indsigterne fra interviews suppleres af en caseanalyse af Masakhane. Casen er valgt for at illustrere, hvordan et lokalt AI-initiativ i Sydafrika kan belyse problematikkerne omkring global ulighed i AI-udvikling (Yin 2014: 16).
Masakhane, som betyder "vi bygger sammen'' på Ndebele, repræsenterer et unikt 'bottom-up' fællesskab dedikeret til udvikling af sprogmodeller for afrikanske sprog.
De er et eksempel på, hvordan teknologisk udvikling kan drives af dem, det direkte berører, i modsætning til de mere udbredte 'top-down' tilgange fra globale tech-giganter. Dette fællesskab har sat sig for at adressere det koloniale efterslæb i AI-udviklingen ved at fremme lokal ejerskab og styrke kulturel og sproglig mangfoldighed.

I en artikel, *African languages to get more bespoke scientific terms,* fra et partnerskab mellem organisationen Masakhane og AfricArXiv, nævner de et af deres projekter 'Decolonise Science', som har til formål at producere frit tilgængelige online ordlister med

videnskabelige termer på seks sydafrikanske sprog, og dernæst bruge dem til at træne maskinlæring-algoritmer til oversættelse. Dette også med formålet om at reducere risikoen for at sprogene forældes og forsvinder, ved at styrke deres online tilstedeværelse (Wild 2021). Masakhanes forskere nævner også hvordan større globale tech virksomheder historisk har ignoreret afrikanske sprog. En talsperson fra Google medgiver også at de er bevidste om at flere tusinde afrikanske sprog er underrepræsenteret i oversættelsessoftware, men at de forsøger at integrere modersmålstalende af sprogene, for at inkludere flere sprog og forbedre kvaliteten af Googles oversættelser.

Casen undersøges gennem en kvalitativ tilgang, hvor data indsamles fra dokumenter relateret til Masakhane, i samtalen med relevante respondenter og relevant sekundærlitteratur. Denne metodetilgang er valgt for at opnå en dybdegående forståelse af, hvordan en bottom-up model for AI-udvikling kan fungere i praksis, og for at illustrere et autentisk verdensperspektiv for udviklingen i teknologisektoren placeret i Sydafrika (Yin 2014: 16). Bryman fremhæver, at casestudier ofte er inddragende og genererer en teoretisk analyse af de indsamlede data, hvilket understøtter vores anvendelse af Masakhane som en kritisk og illustrativ case. (Bryman 2008: 57).

6.4 Kvalitativ generaliserbarhed

Induktiv tilgang

Vi har anvendt en induktiv tilgang til at analysere og fortolke både interview- og dokumentdata. Gennem en grundig tematisk kodning af interviewdata har vi identificeret mønstre og temaer, der belyser, hvordan eksperter i Danmark og Sydafrika opfatter og diskuterer AI's negative og positive effekter. Dette iterative arbejde, hvor teorien gradvist udspringer af dataene, understøttes af den kritiske analyse af reguleringsdokumenter. Her søger vi indsigt i, hvordan magt og ulighed manifesterer sig i de strategier, der benyttes til styring af AI-teknologi (Bryman 2008: 11).

Dette fremhæver den dobbelte natur af vores analyse, hvor vi konstant bevæger os mellem empiri og teori, hvilket også indebærer deduktive elementer i det ellers overvejende induktive arbejde. Vores problemformulering startede med en deduktiv hypotese, men gennem dataindsamling og -analyse har vi justeret og udviklet vores forståelsesramme.

Kvalitetskriterier og generaliserbarhed i kvalitativ forskning

En central udfordring i kvalitativ forskning er at sikre validiteten af de opnåede resultater. I dette projekt fokuserer vi på konstruktionsvaliditet, hvorved vi sikrer, at vores forskning faktisk undersøger det tilsigtede fænomen. I forhold til generaliserbarhed tager vi hensyn til analytisk generalisering, hvor vi sammenligner respondenternes svar med eksisterende teori,

således at det er muligt at generalisere på et teoretisk niveau frem for et statistisk (Bryman 2008: 57). Dette gør, at vi kan drage bredere implikationer af vores resultater og anvende dem til at nuancere forståelsen af lignende fænomener andre steder.

Disse overvejelser understreger, at selvom vores studie ikke skaber ny teori, anvender vi teori som en baggrund for at forstå og fortolke de observerede fænomener. Det er vigtigt at anerkende, at skellet mellem deduktive og induktive tilgange ikke er skarpt adskilt, men snarere bør forstås som en glidende skala af tendenser (Bryman 2008: 57). Vores tilgang afspejler denne fleksibilitet og tilpasningsevne, da vi arbejder induktivt og søger at skabe empiriske generaliseringer snarere end at udvikle nye teorier.

For at sikre en høj kvalitet i vores kvalitative undersøgelse har vi fra begyndelsen haft fokus på gennemsigtighed og metodisk stringens. De første fire interviews blev gennemført i fællesskab, hvilket gjorde det muligt at kalibrere vores spørgeteknik og tilpasse interviewguiden i takt med, at vi blev klogere på feltet. Denne tilgang skabte en fælles forståelse af undersøgelsens retning og styrkede den interne konsistens. Med inspiration fra Guba og Lincolns (1994) kriterier for kvalitativ forskning, som refereret af Bryman, har vi dokumenteret hele forskningsprocessen fra design til analyse for at sikre troværdighed og autenticitet (Bryman 2008: 377). Vi har reflekteret løbende over vores egen rolle og mulige forforståelser, hvorfor vi inkluderer repræsentativ empiri og teori. Vores data og fremgangsmåde er derfor beskrevet detaljeret til en grad hvor andre vil kunne vurdere, hvordan resultaterne er overførbare og pålidelige til lignende kontekster.

Vi har desuden været opmærksomme på undersøgelsens autenticitet, ved at inddrage vores informanter uden en anonymisering og lade deres perspektiver komme til udtryk så retvisende og nuanceret som muligt i vores analytiske konklusioner. Begrebet autenticitet rummer ifølge Guba og Lincoln flere aspekter, der handler om at repræsentere deltagernes perspektiver på en fair og meningsfuld måde (Bryman 2008: 379). Autenticiteten refererer også til, at både os som forskere og eksperterne som deltagere opnår større indsigt i hinandens perspektiver og tilgange til emnet, der skabte dialog og diskussion baseret på vores spørgsmål.

6.5 Videnskabsteoretiske refleksioner

Konstruktivisme og kritisk realisme

Specialet forankrer sig i en kombination af socialkonstruktivistisk- og kritisk realisme-ontologi og epistemologi. Den socialkonstruktivistiske tilgang indebærer en forståelse af virkeligheden som noget, der ikke eksisterer uafhængigt af menneskelig erfaring, men skabes i og gennem sociale processer og fortolkninger. Som Bryman påpeger, opstår sociale fænomener ikke som isolerede størrelser, men som resultatet af menneskers interaktioner og fælles handlinger. De er med andre ord ikke uafhængige størrelser, der eksisterer uden for de sociale processer, der skaber dem (Bryman 2008: 366). Det vil sige, at når vi undersøger fænomener som AI i relation til global ulighed, betragter vi dem som størrelser, der får betydning gennem deres relationer og samspil. Virkeligheden betragtes således ikke som noget givet, men som noget, der hele tiden bliver til af vores konstruktioner af den. Med afsæt i denne forståelse opfatter vi viden som kontekstafhængig og situeret. Den skabes gennem diskurser, interaktioner og de sociale strukturer, der omgiver dem (Bryman 2008: 18). Derfor giver det mening for os at rette blikket mod, hvordan deltagelse i vidensproduktion gennem muligheden for at forske og udtrykke sig på egne sprog i store sprogmodeller, som afgørende for at kunne være med til at forme de forståelser og teknologier, der præger samfundet. Forståelser af f.eks. inklusion og etik i AI opstår i sociale sammenhænge og har samtidig konkrete konsekvenser for, hvordan teknologien designes og anvendes i praksis. Det er denne gensidige påvirkning mellem diskurser og teknologisk udvikling, vi søger at belyse, og deraf er vores analytiske blik forankret i en socialkonstruktivistisk tilgang.

Når vi inddrager kritisk realisme i vores videnskabsteoretiske tilgang, er det for at anerkende, at virkeligheden er kompleks og sammensat, og at sociale fænomener både formes af underliggende strukturer og af aktørers potentialer og handlinger. Kritisk realisme bygger på antagelsen om, at alle objekter – fysiske såvel som sociale – rummer visse potentialer, som kan aktiveres under bestemte betingelser (Buch-Hansen & Nielsen 2012: 183). Det metodiske ideal er derfor at bevæge sig fra det observerbare og konkrete mod en forståelse af de bagvedliggende mekanismer og strukturer, der enten har udløst, hæmmet eller formet det givne fænomen.

I vores speciale anvender et kritisk blik til at anskue generativ AI som et teknologisk og socialt objekt, der rummer forskellige potentialer afhængigt af, hvilke interesser, værdier og former for viden teknologien udvikles ud fra. Særligt fokuserer vi på sprogmodeller som det

observerbare fænomen, vi kan analysere for at afdække, hvilke strukturelle betingelser og magtforhold, der har præget udviklingen, og hvordan disse forstærker eller fastholder global ulighed. Kritisk realisme giver os dermed en ramme for at undersøge, hvordan teknologiske løsninger både er formet af og selv former sociale virkeligheder.
Kritisk realisme forholder sig desuden eksplicit til forholdet mellem viden og magt og adskiller sig her fra postmodernismen. Hvor postmodernismen anser al viden som en konstruktion af magtrelationer, deler kritisk realisme forståelsen af, at viden er socialt produceret, men insisterer samtidig på, at videnskab kan og bør fungere som et kritisk redskab. Videnskab har således ikke kun en erkendelsesmæssig funktion, men også en etisk og politisk forpligtelse til at bidrage til en mere retfærdig verden (Buch-Hansen & Nielsen, 2012: 304–305). Denne forståelse af videnskab flugter med specialets formål: en insisteren på, at en magtfuld teknologi som sprogmodeller ikke blot er et neutralt værktøj, men en videnskabelig praksis, der rummer et ansvar for at modvirke ulighed og fremme retfærdighed.

Opsummerende kombinerer vores videnskabsteoretiske udgangspunkt socialkonstruktivismens blik på magt, mht. hvordan virkelighed og viden skabes gennem sprog, interaktion og diskurser, med den kritiske realismes fokus på de underliggende strukturer, som former og begrænser disse processer. Sammen giver tilgangene os mulighed for både at undersøge, hvordan diskurser om AI og ulighed opstår og får betydning, og samtidig analysere de strukturelle vilkår og magtrelationer, der muliggør og fastholder bestemte teknologiske udviklingsretninger.

7. **ANALYSE**

7. Analyse

Følgende kapitel udgør specialets analytiske del, hvor den indsamlede empiri undersøges gennem det præsenterede teoretiske rammeværk. Analysen tager afsæt i spørgsmålet om, hvordan udviklingen og træningen af generativ kunstig intelligens bidrager til globale uligheder, samt hvilke etiske dilemmaer der opstår i bestræbelserne på at gøre teknologien mere retfærdig og inkluderende.
Analysen er struktureret i to dele på baggrund af vores tematiske kodning (jf. afsnit 6.4):

I første analysedel zoomer vi ind på hvordan træningsdata, sproglig repræsentation og bias i store sprogmodeller reproducerer koloniale strukturer. Vi undersøger også hvordan Masakhane's deltagelsesbaserede tilgang forsøger at forbedre vilkårene for panafrikanske sprog.

I anden analysedel zoomer vi ud, og analyserer generativ AI som geopolitisk teknologi, hvor vi også kigger på regulering og behovet for internationale retningslinjer. Vi ser nærmere på, hvordan globale magtforhold former udviklingen af AI, og hvilke implikationer det har for en mere retfærdig og ansvarlig teknologisk fremtid.

7.1 Analysedel 1

Inklusion, repræsentation samt træningsdata, sproglig diskrimination og bias i store sprogmodeller var centrale temaer i vores interviews med respondenterne. Disse fænomener i AI-teknologi rejser en række dilemmaer, som vi vil udfolde i følgende afsnit.
Som beskrevet i forskningsfeltet kan AI medvirke til at videreføre koloniale strukturer samt bidrage til at reproducere og forstærke eksisterende socioøkonomiske uligheder, som yderligere udfoldes i denne del af analysen. Med særligt fokus på generativ AI og sprogmodeller undersøges det desuden hvordan aktører som organisationen Masakhane arbejder for at udfordre og omforme disse forhold gennem mere inkluderende og retfærdige praksisser.

7.1.1 Bias i kunstig intelligens

I følgende afsnit vil vi dykke ned i magt og sprog, og for at afdække det, trækker vi på interviews fra respondenterne, Rosman, Just og Ihle, da de bidrager med de væsentligste pointer relateret til dette.

For at forstå hvordan ulighed har relation til store sprogmodeller er det afgørende

først og fremmest at forstå relationen mellem sprog, magt og ulighed.
Som beskrevet jf. Fairclough i teoriafsnittet (jf. afsnit 5.6), er sprog og diskurs centrale mekanismer til at forme og opretholde magtstrukturer. Dette sker blandt andet ved, at særlige perspektiver og interesser fremstilles som almengyldige, hvilket giver visse sociale grupper mulighed for at fastholde dominans gennem definitionen af, hvad der betragtes som 'normalt' eller 'universelt'. Schwartz (2022) fremhæver en lignende pointe og beskriver det som en videnskabelig udfordring: Forskning overser ofte, hvordan den selv er indlejret i en kolonial videnstradition, hvilket påvirker både de spørgsmål, der stilles, og de løsninger, der udvikles. Han argumenterer for, at et første skridt mod en mere afkoloniseret tilgang er at blive bevidst om de grundantagelser, der former forskningen, da videnskabelige paradigmer historisk er tæt forbundet med koloniale strukturer. Disse videnskabelige paradigmer gør sig også gældende i udviklingen af teknologier som generativ AI.
Adams (2025) fremfører, hvordan AI ikke blot er et teknologisk redskab, men også et vidensprojekt, hvor definitionen af sprog, normer og værdier i modellerne reproducerer et verdenssyn, der bidrager til en epistemisk ulighed, hvor Vestens måder at forstå og beskrive verden på fortrænger andre. Som Benjamin Rosman, professor ved University of the Witwatersrand, og en af vores respondenter, pointerer: "Everyone is biased, all data is biased, and you want your technology to reflect your values" (Rosman: 7, 27-28).
Viden og sandhed er ikke neutral, men formet af det perspektiv, de anskues fra – hvorfor bias uundgåeligt præger både sprogmodeller og de aktører, der udvikler dem. Dette perspektiv deles også af vores respondent Sine Nørholm Just, Professor ved Roskilde Universitet og projekt- og forskningsleder ved Algoritmer, Data og Demokrati (ADD), som understreger:

> *"Sprogmodeller har altid en ideologi – ofte en amerikansk eller vestlig ideologi – og selvom der findes kinesisk-drejede modeller som DeepSeek, kan vi konstatere, at man ikke kan lave en sprogmodel uden nogen form for bias eller drejning.*
> *Det betyder, at visse tanker, holdninger, måder at formulere sig på, og helt konkrete sprog favoriseres over andre. Vi har en debat om dansk og behovet for en dansk sprogmodel, men i det bredere globale perspektiv ser vi, at nogen får gevinsterne, og andre i bedste fald bliver overset, i værste fald udnyttet og får det værre."* (Just: 2, l. 16-22).

En central pointe i citatet er, at AI ikke er en neutral teknologi. En pointe, der kan relateres til Faircloughs forståelse af 'governance', hvor teknologiske aktiviteter, i dette tilfælde

sprogmodeller, fungerer som styringsmekanisme for andre sociale praksisser. Fairclough anerkender samtidig 'agens' – altså individers handlekraft og mulighed for at gribe ind i sociale praksisser. Men denne handlekraft kan være svær at udøve i mødet med generativ AI og sprogmodeller, der ofte er præget af lav transparens. Brugerne får sjældent indsigt i, hvilke data og vidensgrundlag modellerne bygger deres svar på, hvilket gør det vanskeligt at gennemskue og udfordre de underliggende ideologier og magtstrukturer, som modellerne potentielt viderefører. Dette taler Rosman om:

> *"The problem is, when you look at a neural network, you can't just read off if it's biased, if it hates people, or how good it is at addition. You can't force it to behave a certain way by hardcoding rules. It's all hidden in these parameters"* (Rosman: 4, 24-26).

Rosman pointerer desuden, at store sprogmodellers adfærd, som ChatGPT, er uforudsigelig og indlejret i komplekse teknologiske strukturer, hvilket gør det næsten umuligt at forudsige eller identificere skjulte bias uden omfattende testning af hver en mulig funktion (Rosman: 4, 27-28). Det viser, hvordan magt og ideologi ikke nødvendigvis udøves eksplicit, men kan være skjult i modellernes design og i deres træningsdata. Det, mener vi, understreger behovet for kritisk opmærksomhed og governance i arbejdet med sprogmodeller.

Når store, kommercielle sprogmodeller bliver allestedsnærværende og ikke udfordres af alternativer, bliver det vanskeligt at bryde ud af deres dominans. Bias i sprogmodeller, indebærer, at modeller, der primært understøtter engelsk, ikke blot ekskluderer ikke-engelsktalende grupper, men også bidrager til at udviske kulturel historie og alternative måder at tænke og kommunikere på. Som Adams (2025) formulerer det, er automatiserede sprog noget, som former verden, favoriserer bestemte udtryk, og reproducerer de værdier og normer, de er skabt af. Dette er særligt bekymrende, når sprogmodeller fra tech-virksomheder integreres i samfundskritiske funktioner som uddannelse og sundhedsvæsen, hvor præcise oversættelser og kulturelt funderede forståelser kan være afgørende.

En central kritik, som vores respondent Just fremsætter, angår den ulige fordeling af adgang, indflydelse og repræsentation i udviklingen af AI. Hun påpeger, at visse grupper overses, udnyttes eller får det værre, og hvordan sprogmodeller og deres underliggende datasæt kan fastholde og forstærke eksisterende strukturelle uligheder (Just: 2, 20-25). Det gælder særligt i forhold til sproglige og kulturelle fællesskaber, som enten ikke er tilstrækkeligt repræsenteret i datasættene eller ikke har adgang til den nødvendige

teknologiske infrastruktur til at præge udviklingen. Dette er hvad Claus (2024) betegner som en form for epistemisk skævhed, hvor bestemte former for viden, sprog og formuleringer fremstår som objektive eller universelle, mens andre marginaliseres. Det medfører, at nogle aktører og sprogområder opnår uforholdsmæssigt meget indflydelse og gevinst, mens andre risikerer at blive yderligere teknologisk, kulturelt og økonomisk marginaliseret. Hvilket vores respondent Laura Haaber Ihle, Vice President of Ethics, Governance & Policy ved Responsible AI Future Foundation, også peger på:

> *"Hvis systemet genererer output baseret på dårlige eller ufuldstændige data, kan det føre til, at nogle mennesker bliver diskrimineret eller på andre måder behandlet uretfærdigt. Der er også noget, der handler om kulturelle forskelle. Mange af disse modeller har en indbygget teknologisk imperialisme, fordi de er trænet på data, der repræsenterer bestemte værdier og normer, der måske ikke er universelt gældende."* (Ihle: 4, 4-11).

Som Floridi og Cowls (2019) fremhæver i deres dataetiske princip om 'retfærdighed', forudsætter en retfærdig udvikling af AI, at der sikres reel inklusion og lige adgang til teknologiens muligheder. Dette kræver bevidst handling og regulering - ikke mindst fordi kommercielle aktører sjældent prioriterer dette hensyn af sig selv.
Retfærdig udvikling og adgang til AI, er i nogen grad afspejlet i EU's AI-forordning, som lægger vægt på gennemsigtighed og beskyttelse af brugeres rettigheder (Europa-Parlamentet og Rådet, 2024). Dog tager EU-lovgivningen i mindre grad højde for de globale uligheder i selve udviklingsprocesserne. Her adskiller Den Afrikanske Unions AI-strategi fra 2024 sig, ved at fremhæve betydningen af lokal innovation, sproglig mangfoldighed og opbygning af teknologisk infrastruktur. AU's strategi peger på, at AI-teknologier ofte viderefører eller forstærker eksisterende skævheder i deres udvikling. Samtidig rejser strategien bekymring for manglende gennemsigtighed, trusler mod menneskerettigheder, sikkerhedsrisici i civile og militære sammenhænge, herunder spredning af misinformation samt politisk manipulation og 'deepfakes' (uigennemsigtig billedmanipulation) drevet af generativ AI (African Union, 2024, s. 3).
Som både Ihle peger på og som der står i AU-strategien, fører snævre datasæt og manglende mangfoldighed til vidensdominans, hvor minoritetsverdenens normer gøres til globale standarder. Bias i datasæt risikerer dermed ikke blot at marginalisere store befolkningsgrupper, men også at forstærke eksisterende uligheder i adgang til viden og

rettigheder. Det bekræfter Floridis og Cowls (2019) pointe om retfærdig AI kræver aktiv inklusion af forskellige perspektiver, også i selve data- og udviklingsprocesserne.

7.1.2 Koloniale strukturer videreføres

I dette afsnit trækkes der hovedsageligt på respondenterne Arakkal og Alayande, hvor vi undersøger hvordan udviklingen af generativ AI og sprogmodeller, særligt i USA og Kina, risikerer at ekskludere både afrikanske, europæiske samt andre kontekster og dele af verden. Denne eksklusion kan overføres til det, som kritikere omtaler som en form for 'algoritmisk apartheid'. Begrebet trækker på referencen til apartheidregimet i Sydafrika, hvor samfundet blev opdelt efter hudfarve, og hvor sorte borgere systematisk blev frataget rettigheder og adgang til f.eks. uddannelse, arbejde og muligheden for at blive undervist på deres egne sprog. I denne sammenhæng peger begrebet på en gentagelse af strukturel ulighed, hvor nogle har adgang til at udvikle, definere og drage nytte af teknologien, mens andre reduceres til passive brugere eller utilsigtede ofre for teknologiens indbyggede bias.
Claus (2024) påpeger at initiativer som Masakhane arbejder på at udvikle nye terminologier og udvider og datasæt for såkaldte 'ressourcefattige' sprog. Det gør de for at genoprette nogle af de privilegier, de historisk set ikke har haft.
Singh (2008) peger tilsvarende på, at manglen på økonomiske ressourcer, teknisk udstyr, menneskelig kapacitet og politisk opbakning er blandt de væsentligste barrierer, der hæmmer fremskridt og innovation inden for sprogteknologisk understøttelse af underrepræsenterede sprog.
Som vores respondent Archana Arakkal, AI-specialist og medstifter af Synthesis Software Technologies, nævner i interviewet: "The whole purpose of scaling these large language models to ensure that people who don't have access to this information finally do have access, but instead you reversing all of that." (Arakkal: 6, 33-35). Hun italesætter denne 'omvendte udvikling' ved at pege på, hvordan AI egentlig rummer potentialet til at give flere mennesker adgang til information, de tidligere ikke havde, men at dette mål undergraves, når der eksisterer bias og diskrimination i store sprogmodeller. Derfor understreger hun nødvendigheden af, at man ved træning af modellerne inddrager specifikke og nuancerede datasæt, der er relevante for bestemte lande. Dette er i tråd med Adams (2025) synspunkter om at afkolonisering i Afrika, med henblik på AI, forudsætter, at der arbejdes aktivt med afrikanske sprog som fundament for lokale modeller.
Arakkal fremhæver Masakhane som et positivt eksempel på en tilgang, der bevidst inddrager

sprog, der ellers risikerer at blive marginaliseret eller ekskluderet (Arakkal: 7, l. 31-37). Der er således en enighed i at den nuværende dominerende teknologi ikke kommer majoritetsverden og underrepræsenterede 'ressourcefattige' sprog til gode, men at initiativer som Masakhane som bevidst arbejder på inklusion af flere sprog, er et skridt i den rigtige retning (Lignos et al. 2022).

Når det gælder underrepræsenterede sprog i store sprogmodeller, herunder også dansk, viser der sig forskellige konsekvenser og former for ulighed sammenlignet med afrikanske sprog, selvom de kan være lige så underrepræsenteret i de store sprogmodeller. Nussbaum (2003) argumenterer for, at alle mennesker bør sikres adgang til et sæt fundamentale kapabiliteter, uanset kulturel eller politisk kontekst. Set i dette lys er det problematisk, at AI-principper ofte baseres på vestlige normer eller teknologisk privilegerede sprog som engelsk. Når store sprogmodeller ignorerer underrepræsenterede sprog, overses det etiske ansvar for at sikre lige adgang til teknologisk infrastruktur. Ifølge Nussbaum er det netop uetisk at bygge globale systemer uden at sikre et minimumsniveau af retfærdighed, hvilket her også indebærer sproglig inklusion.

Ifølge Claus (2024) kan små sprog, som eksempelvis dansk, trods et relativt lavt antal talere, være langt bedre digitalt repræsenteret og dermed lettere anvendeligt i træningen af sprogmodeller. Som beskrevet i forskningsfeltet, står dette i kontrast til eksempelvis Kiswahili, der tales i mindre ressourcestærke regioner og har langt mindre tilgængeligt digitalt sprogligt materiale, hvilket begrænser dets tilstedeværelse i AI-udviklingen, selvom der er 150 mio. som taler sproget.

Som respondent Ayantola Alayande, forsker ved Global Center on AI Governance, også påpeger om afrikanske sprogs digitale fodspor: "Less than 1% of African languages are represented on the internet. (…) We need to start a conversation around first digitalising a lot of these languages." (Alayande: 15, 23-25). Vi kan dermed udlede, at digital repræsentation, inklusion og muligheder i AI ikke afhænger af antallet af sprogbrugere, men af økonomiske ressourcer, digital infrastruktur og adgang til teknologisk dannelse.

Adams (2025) peger på at problemer vedrørende infrastruktur, såsom ustabil elektricitet og internet, skaber en lav kapacitet i store dele af Afrika og dermed begrænser mulighederne for teknologisk udvikling og inklusion. I tilfældet med adgang og brug af sprogmodeller handler det ikke blot om teknisk tilgængelighed, men om en række underliggende forhold, herunder social og kulturel kapital, teknologisk infrastruktur og økonomisk kapacitet, som i praksis fungerer som barrierer. Hvis en sprogmodel for eksempel ikke understøtter ens sprog fyldestgørende, eller hvis teknologien ikke tilfører værdi i ens dagligdag, bliver den i praksis

ubrugelig - uanset om den formelt set er tilgængelig. Der er således tale om ulighed i 'teknologiske kapabiliteter' og adgang til viden, hvilket begrænser muligheden for at anvende og drage nytte af AI på lige vilkår. Dette står i kontrast til lande som Danmark, hvor både strukturel og individuel kapacitet i langt højere grad muliggør meningsfuld og produktiv brug af teknologien sammenlignet med lande som f.eks. Sydafrika.
Ifølge Sen er der forskel på at have friheden til at fravælge muligheden, hvad Sen kalder 'functionings', til slet ikke have adgang eller kapabiliteter til at bruge teknologien; altså er det noget andet, at vælge ikke at bruge sprogmodeller og generativ AI (Sen 1993).
Vores respondent, Arakkal, påpeger i forbindelse med et kvalitetsmonitoreringssystem i callcentre, at der var et tidspunkt, hvor engelsk var den eneste løsning, i et land som Sydafrika med 11 officielle sprog (Arakkal: 4, 35-36). Når teknologier kun virker på engelsk, ekskluderes en stor del af befolkningen, og dermed også hele brancher og markedssegmenter:

> *"We're excluding them by default, which means there's a large portion of South Africa that isn't being taken into account when trying to create models. So it's quite a big problem because we're not really tapping into that value proposition that we should be."* (Arakkal: 5, 7-10)

Arakkal understreger, hvordan fraværet af flersproglig understøttelse ikke blot er en teknisk mangel, men en strukturel udelukkelse med direkte konsekvenser for hele befolkningsgrupper og brancher. Når teknologien ikke tilpasses konteksten, forbliver den et værktøj, som kun kommer til gavn for de få. Set i lyset af Freemans (2015) stakeholderteori bliver det tydeligt, hvordan ulig adgang til AI-teknologi ikke blot handler om tekniske barrierer, men om en manglende inddragelse af væsentlige aktører i det globale AI-økosystem. Freeman definerer en stakeholder som enhver gruppe eller person, der kan påvirke eller påvirkes af en organisations formål, og han betoner vigtigheden af at identificere og engagere alle relevante parter - ikke kun dem med økonomisk eller teknologisk magt. Når lande og lokalsamfund med begrænsede teknologiske kapabiliteter ikke bliver hørt eller inddraget i udviklingen af sprogmodeller, illustrerer det en mangel på inklusion af stakeholders. Der sker med andre ord en systematisk eksklusion af dem, der er stærkt påvirket af AI's udbredelse, men som har begrænset mulighed for at påvirke den. Det forstærker den asymmetri, der allerede findes mellem globale tech-virksomheder og lokalt forankrede organisationer som Masakhane.
Ifølge Rawls' (2023) idé om, at samfundets institutioner kun må tillade ulighed, hvis det kommer de mest udsatte til gode, bliver dette problematisk, da store dele af

majoritetsverdenen udelukkes fra både udviklingen og udbyttet af AI.
Floridi og Cowls (2019) principper om 'retfærdighed' og 'gavnlighed' forklarer, at AI bør fremme velfærd og fordele udbytte retfærdigt. Dette skaber en forpligtelse for de mere velstillede nationer, der besidder udviklingskapacitet og ressourcer.
Som tidligere påpeget af respondent Thomsen, har disse lande et omfattende ansvar for at støtte mindre velstående lande i langt højere grad, end det hidtil er blevet gjort. Samtidig understreger han, at det også er et prioriteringsspørgsmål, og at det handler om at finde de mest effektive måder at yde hjælpen på (Thomsen: 12, 13-14). Dette etiske krav understøttes af Rawls' forståelse af 'fairness', som hævder, at ingen bør nægtes beskyttelse eller goder baseret på vilkårlige kriterier som geografi eller kultur. Det bliver derfor uacceptabelt, at teknologisk adgang er forbeholdt dem, der allerede er magtfulde – for havde det været omvendt ville de aldrig have accepteret denne verdensorden.
Det normative krav er derfor dobbelt: Ikke kun at handle retfærdigt, men at definere retfærdighed i fællesskab med de grupper, der i dag er ekskluderede. Som Adams (2025) pointerer, bærer aktører af AI-udviklingen et globalt ansvar, som de i dag ikke lever op til, særligt ikke i forhold til ikke-vestlige og ikke-engelsktalende kontekster, hvor eksklusionen er mest udbredt.
Vi konkluderer, at der er en reel risiko for, at vi er ved at etablere en teknologisk verdensorden, hvor algoritmer, udviklet og kontrolleret af et begrænset antal nationer og tech-virksomheder, aktivt ekskluderer dele af verdens befolkning. Dette er ikke blot beklagelsesværdigt, men ud fra flere dominerende normative teorier etisk uacceptabelt.

7.1.3 Lokale alternativer

Dette afsnit undersøger den sydafrikanske non-profit organisation Masakhane som et eksempel på lokal teknologisk modmagt til de store tech-virksomheder, der arbejder for at fremme inklusionen af panafrikanske sprog i udviklingen af AI.
I afsnittet trækker vi på interviews med Alayande, Thomsen, Mukuri-Smith og Ihle, som taler om repræsentation og inklusion i sprogmodeller.

Som tidligere fremført, er AI-sprogteknologier som chatbots ikke neutrale, men bygger på data og sprog, der ofte afspejler bestemte verdenssyn. Den sydafrikanske organisation Masakhane søger at modvirke dette ved at skabe og træne data til sprogmodeller, der er repræsenteret gennem panafrikanske sprog og videnssyn.

Som nævnt i forskningsfeltet påpeger Lignos et al. (2022) det er vigtigt at inddrage

modersmålstalende i udviklingen af træningsdata, hvilket skaber mere præcise og kulturelt relevante resultater, som det eksempelvis ses i Masakhane-projektet. De fremhæver, at menneskelig inddragelse, i modsætning til automatiserede processer, er central for at opnå den mest retvisende kvalitet i sprogmodeller. Dog peger vores respondent Alayande på, at det er helt centralt at behandle selve roden af problemet, og at det kan ses som symptombehandling, hvis man alene arbejder med inklusion og lingvistisk repræsentation i AI:

> *"I think the point on local language AI and inclusion is very important. But I think it's just basically addressing the wound rather than treating the disease. (...) For me, this comes back to a broader problem in our approach to development practice; you know, just top-down treatment of the symptoms rather than addressing the structural issues."* (Alayande: 3, 11-14; 5, 3-4).

Vores respondent Frej Klem Thomsen, seniorkonsulent ved Nationalt Center for Etik, deler denne kritik og understreger, at selvom det kan være god symbolik at udvikle sprogmodeller ved inddragelsen af forskellige grupper i udviklingen, er det ikke nødvendigvis en løsning på de dybereliggende problemer. Han fremhæver, at det for mange mennesker vil være svært at deltage i udviklingen af sådanne teknologier på grund af kompleksiteten og mangel på ressourcer:

> *"Jeg tror, det har risiko for at være en form for symptombehandling. Det kan godt blive noget, der ser godt ud, uden at have den effekt, man håber på. Det er vanskeligt for mange almindelige mennesker at indgå i de her processer, og det løser ikke nødvendigvis de problemer, man håber at løse ved at inkludere."* (Thomsen: 9-10, 32-9).

Alayande deler denne kritik og påpeger, at selv om udviklingen af store sprogmodeller på afrikanske sprog er et vigtigt skridt for at bevare kulturen, så løser det ikke fundamentalt spørgsmålet om inklusion. Han anerkender betydningen af at udvikle sprogmodeller på afrikanske sprog for at bevare kultur, men advarer om, at det ikke i sig selv skaber reel inklusion. Han ser det som symbolpolitik, fordi det egentlige problem handler om, hvem der ejer teknologien, og hvordan den kulturelle suverænitet sikres (Alayande: 3, 15-17).
Ifølge både Thomsen og Alayande kræver reel forandring, dybdegående strukturelle ændringer og en reel omfordeling af ejerskabet over AI-teknologier.

Når Alayande taler om kulturel suverænitet, trækker han på samme logik som Spivak (1988), nemlig at det ikke er tilstrækkeligt blot at give stemme til de marginaliserede, men handler om magt, ejerskab og retten til at definere teknologiske rammer på egne præmisser. Det er et godt eksempel på hvordan 'demokratisering af AI' (Seger 2023) rummer flere nuancer og dilemmaer. F.eks. kan demokratisering af brugen af AI, ved at gøre chatbot-teknologi åbent til alle, kollidere med spørgsmålet om demokratisering af styring og profitdeling - hvis udbredelsen af disse chatbots sker gennem kommercielt styrede modeller med koncentreret ejerskab.

Set i et postkolonialt perspektiv rejser vi spørgsmålet, om det overhovedet er muligt at udvikle lokalt forankret teknologi i en postkolonial virkelighed. Med udgangspunkt i Spivaks (1988) kritik spørger vi, om de 'subalterne' - i dette tilfælde panafrikanske sprogfællesskaber - overhovedet kan 'tale' eller blive hørt i udviklingen af AI, når teknologiske standarder fastsættes af vestlige aktører. Samtidig står en aktør som Masakhane over for strukturelle og ressource-prægede udfordringer. Denne strukturelle ubalance forstærkes yderligere af en observation fra vores respondent Hellen Mukuri-Smiths, selvstændig advokat inden for kunstig intelligens og dataregulering, at de globale magtcentre, særligt Europa, USA og Kina, har opnået fordelen ved at være først, i kraft af teknologisk forspring. Mukiri-Smith advarer om, at uden effektiv regulering og modvægt til de dominerende aktører, vil disse magtcentre fortsætte med at definere rammerne for AI-udviklingen, hvilket blot forstærker ulighed og yderligere marginaliserer alternative aktører, som eksempelvis Masakhane (Mukiri-Smith: 3, 46).

Det er netop dette spændingsfelt Mukiri-Smith påpeger, og som både Alayande og Thomsen fremhæver: At reel inklusion forudsætter, at ejerskabet over teknologien flyttes, og at der samtidig etableres stærke governance-strukturer, hvor eksperter kan sikre ansvarlighed og retfærdighed i AI-modellernes udvikling og anvendelse, som vi vil analysere i afsnit 7.2.3. Eke et al. (2023) understreger yderligere, at det kan være vanskeligt for alternative aktører som Masakhane at operere med begrænsede ressourcer, såsom dårlig internetforbindelse og manglende kompetencer til at kode og operere LLM'er. Dette perspektiv er også væsentligt i diskussionen om ulighed, som vores respondent Alayande pointerer (Alayande: 8, 17-20). Ulighed og inklusion i forhold til AI kan nemlig forstås både som kontinentale uligheder mellem Afrika og resten af verden, samt som interne uligheder i de enkelte lande. Alayande påpeger, at det ikke kun handler om adgangen til AI, der repræsenterer deres sprog, men også om, hvorvidt individer har de nødvendige midler til at bruge teknologien, eksempelvis via en computer og en stabil internetforbindelse:

> *"So there are two ways you can think about inclusivity or equality with AI. You can think about Africa catching up with the rest of the world and you can also think about within-country equality in AI access/skilfulness, both are very important." (Alayande: 8, 11-12).*

Respondent Ihle fremhæver ligeledes, at det er nødvendigt at afveje investeringer i AI op mod mere presserende behov i mange lande, hvor ulighed internt i lande også er en central problematik. Hun påpeger, at der kan være en skævhed i, hvad der vurderes som værdifulde investeringer, både lokalt og af internationale aktører: "(...) nu skal alle med på den der AI-bølge. (...) Og så er der jo også prioriteringsspørgsmål: Skal man investere milliarder i det her, når man ikke har veje?" (Ihle: 12-13, 32-2).
Derudover peger hun på de praktiske udfordringer, som begrænset adgang til strøm, stabile internetforbindelser og digital infrastruktur udgør, og understreger, at uden basale teknologiske forudsætninger, såsom wifi, giver det ikke mening at implementere avancerede AI-systemer (Ihle: 13-14, 32-1).
Flere af specialets respondenter og kilder i litteraturen peger på lignende udfordringer, som rejser spørgsmålet om, hvorvidt det er muligt, og under hvilke betingelser, at opbygge lokale og inkluderende AI-systemer under de nuværende globale strukturer.
Claus (2024) påpeger derudover, at selv projekter, der arbejder for inklusion, som Masakhane, kan komme til at forstærke eksisterende uligheder. Det sker, når de især støtter de største afrikanske sprog, mens mindre sprog bliver overset. Hvis et sprog ikke får teknologisk støtte, bliver det svært for dem, der taler det, at få adgang til og deltage i den digitale udvikling.
Tilsammen viser det, hvor svært det kan være at sikre reel teknologisk inklusion i en postkolonial sammenhæng. Alligevel mener vi, at det er vigtigt at diskutere, hvilke former for AI-udvikling der er mest etiske og ansvarlige, hvilket vi vil udfolde i analysedel 2 (jf. afsnit 7.2).

7.1.4 Fremtidens sprogteknologier

I dette afsnit af analysen zoomer vi ind på de paradokser og dilemmaer, der præger nutidens AI-teknologi i sprogmodeller, og undersøger, hvordan den nuværende udvikling former fremtidige scenarier samt hvilke forudsætninger og tilgange der er nødvendige for at modvirke de ulige strukturer, teknologien risikerer at forstærke. Afsnittet bygger på perspektiver fra Høst, Rosman, Just, Arakkal, Ebbensgaard, Ihle og Mukuri-Smith, som alle

bidrager med centrale indsigter.

Vores respondent Marie Høst, tech-journalist og medievært, understreger, at selv i et velstillet land som Danmark, hvor man kunne forvente gode forudsætninger, er det både økonomisk og teknologisk krævende at udvikle og vedligeholde en sprogmodel på et lille sprog. Dette perspektiv udvider hun ved at sammenligne med lavindkomstlande i Afrika, som har endnu færre ressourcer og langt svagere digital infrastruktur. Pointen er, at hvis det er så svært for os, så er det næsten uoverkommeligt for dem, og dermed risikerer vi, at AI-udviklingen forstærker de eksisterende uligheder mellem lande og regioner. Som Høst pointerer:

> *"Vi cementerer endnu en gang de globale forskelle med de løsninger, vi indtil videre præsenterer, for det kræver jo en enorm økonomi. Bare det at bygge en dansk sprogmodel er en kæmpe investering. (...) Og hvis det er dyrt for os, nogle af de rigeste i verden, så kan man jo godt forestille sig, hvordan det ser ud for mindre afrikanske lande med langt færre ressourcer. Jeg tror desværre, at det her kan være med til at skabe en endnu større kløft."* (Høst: 6, 9–18).

Høst beskriver udviklingen af sprogmodeller således, at algoritmer forstærker den eksisterende ulighed (Høst: 1, 21-24). Det kan knyttes til det fænomen, Couldry og Mejias (2019) betegner som datakolonialisme – en fortsættelse af koloniseringen i det digitale, hvor eksisterende strukturelle uligheder ikke blot reproduceres, men forstørres gennem datadrevne teknologier. Når Høst og flere af vores respondenter, samt forskningslitteratur, peger på denne sandsynlige forstærkning af ulighed, kalder det på hvilken fremtid vi træder ind i. Desuden opstår en risiko for at indgå i det Adams (2025) kalder et 'nyt AI-imperium', hvor teknologisk dominans og ressourcer koncentreres omkring en privilegeret minoritet, mens majoritetens sprogmodel-initiativer holdes udenfor udviklingen og værdiskabelsen.

Rosman peger på de indlejrede paradokser fra et effektiviseringssynspunkt, som findes i den nuværende udvikling af AI-teknologi, hvor man ville se positivt på, hvis hele planeten ender med at tale samme teknologiske sprog. Ulempen ved dette er, at det sletter en masse kultur og historie, fra de steder hvor den slags arv udelukkende sker oralt og ikke er nedskrevet nogen steder – og hvis du mister evnen til at engagere med det, mister du måske adgangen til lokal viden (Rosman: 6, 26-30).

Just påpeger, at det er et privilegium at kunne bruge sit modersmål i mange sammenhænge, og det ikke at kunne bruge det, skaber kulturel ulighed. Lighed handler i dette tilfælde om

adgang til teknologi på sit eget sprog, og at kunne bevare sin kultur (Just: 7, 9-12). Dette viser, hvordan AI-imperialisme jf. Adams (2025), manifesterer sig som et epistemistisk tab for de befolkningsgrupper, hvis viden ikke eksisterer digitalt. Justs påpegning af privilegiet ved at kunne bruge sit modersmål i globale sammenhænge tydeliggør hvordan sprog er en magtfaktor. Derfor peger flere respondent Arakkel også på, at den mest retfærdige og ansvarlige udvikling af AI og særligt sprogmodeller, kommer i dialogen og samarbejdet mellem de aktører som udvikler teknologien og de fællesskaber hvis sprog er underrepræsenteret digitalt (Arakkel: 4, l. 7). Dog er eksempelvis respondent Ida Ebbensgaard, tech-journalist og forfatter, kritisk over for de større tech-giganters reelle interesse i overhovedet at træne på et lille sprogområde, fordi det ud fra et forretningsmæssigt og kommercielt synspunkt ikke giver profit nok: "De har ikke nødvendigvis stor interesse i at træne på et lille sprogområde eller et lille kulturområde, hvor de ikke skal sælge så meget" (Ebbensgaard: 6, 19-20).
Just påpeger også hvordan vi nødvendigvis heller ikke kan tvinge virksomheder til at lave sine produkter på en særlig og bestemt måde, hvis man ikke decideret gør det ulovligt. Hun forudser også hvordan udviklingen af sprogmodeller kommer til at accelerere, at de vil bruges til mange typer af arbejde, og jo derfor får en systematisk påvirkning af vores verdenssyn, som kan være svært at lovgive om (Just: 8, 14-23).
Flere påpeger også vanskelighederne i et fælles reguleringsapparat, som Rosman:

> *"I actually think the best solutions for how to regulate AI might not come from the US. Maybe the right answers come from Denmark, or Senegal, or somewhere else — from places with different perspectives and values."* (Rosman: 6, 6-8).

Sammenlagt peger vores respondenter på konturerne af en fremtid, hvor flere lokalt forankrede og alternativer til de store tech-giganter og dominerende magter, vokser frem, fordi der er en reel interesse og vilje f.eks. på det afrikanske kontinent til ikke længere at være hægtet af og som en part, der er afhængig af hjælp.
Rosman beskriver, hvordan et panafrikansk initiativ og årlig konference, *the Deep Learning Indaba*, har katalyseret hele AI-fællesskabet på tværs af Afrika og hvordan dette har affødt subfællesskaber som Masakhane og startups som hans egen, Lelapa AI: "We had 330 participants at our first Indaba – it has since grown to 700–800 participants each year." (Rosman: 1, 28-29). Det vidner om en accelererende teknologisk mobilisering, der både skaber viden og kapacitet.

Vores respondent Mukiri-Smith deler en forsigtig optimisme og beskriver, hvordan der er en modbevægelse i gang – særligt i afrikanske lande – hvor lokale aktører udvikler egne løsninger: "Local people are very clever, very creative, doing a lot of things with tech and developing the models." (Mukiri-Smith: 6, 43-44). Det viser, at der er spirende potentialer for at forme teknologien nedefra, også i mødet med dominerende globale kræfter, som Meta og Google, som nu møder modvind idet, accelerationen af teknologien har taget fart på en måde, hvor det regulatoriske behov samt uhensigtsmæssige fejl i de store modeller opdages samtidig som teknologien slippes løs. Dette er det, Couldry og Mejias (2023) kalder for modstandsformer mod datakolonialismen, hvor lokal viden og kollektive teknologiske mobiliseringer bliver modstrategier.
Respondent Ihle peger på et lignende alternativ i Mellemøsten, hvor en tech-virksomhed ønskede at udviklede en arabisk sprogmodel, som ifølge Ihle skyldtes et ønske om 'ikke at begå de samme fejl igen' og 'gøre tingene ordentligt' – altså at udvikle AI på etisk forsvarlig vis fra begyndelsen. Hun tilføjer: "De tager virkelig det der seriøst, og det tror jeg, mange andre lande også gør – altså nu vil vi være dem, der laver det her etisk forsvarligt." (Ihle: 12, 21–22). Dette fremhæver, hvordan ressourcestærke aktører uden for den vestlige sfære i stigende grad søger at definere egne etiske og politiske standarder for AI. Desuden fremhæver Ihle, at korrekt trænede modeller, som netop tager højde for bias, at teknologien kan bruges til at minimere ulighed, ved at være mindre diskriminerende end mennesker, f.eks. være: "(...) mindre diskriminerende end en sagsbehandler" (Ihle: 10, 4-5).
Generativ AI rummer potentiale til at mindske ulighed og fremme mere retfærdig og etisk behandling. Men dette afhænger i høj grad af, hvem der designer teknologien, hvordan den udvikles, og hvilket input den bygger på. Teknologi er ikke neutral, den formes af menneskelige valg, værdier og magtforhold.

Nedenfor pointerer respondent Just om måden hvorpå AI bliver anskuet som løsningen på mange af verdens problemer, også ift. global ulighed:

> *"Hvis det er forestillingen – at AI en dag skal redde os – risikerer vi at skabe større ulighed undervejs. Ligesom når man siger, at teknologi skal løse klimakrisen, men overser, at teknologien i sig selv forurener og bruger enorme ressourcer."* (Just: 6, 31- 33).

Dette bekræfter Couldry & Mejias' (2019) pointe om, at datakolonialisme netop fungerer gennem normaliseringen af dataudvinding som 'fremskridt' uden hensyn til de sociale

omkostninger. Det vidner om en kritisk tilgang til, at teknologien i sig selv i dag primært er styret af markedskræfter, som for nuværende ikke fører til mindre socioøkonomisk ulighed.

Sammenlagt viser ovenstående, hvordan AI og sprogmodeller i dag risikerer at forstærke eksisterende uligheder, særligt for sprogfællesskaber og regioner, der i forvejen er dårligt repræsenteret digitalt. Samtidig ser vi, at der vokser nye initiativer frem, både i Sydafrika, som Masakhane, og andre steder, hvor lokale aktører forsøger at tage kontrollen over udviklingen og skabe løsninger, der bygger på deres egne behov og værdier. Fremtiden for AI bliver formet i spændet mellem store teknologivirksomheders interesser, staters regulering af teknologien samt lokale fællesskabers kamp for at blive hørt og få indflydelse.

7.1.5 Delkonklusion

Vi har analyseret, hvordan udviklingen og brugen af store sprogmodeller er tæt forbundet med globale uligheder og historiske magtstrukturer. Sprogmodeller trænes primært på vestlige datasæt og bærer særlige ideologier og bias og udelukker dermed mange sprog og vidensperspektiver. Dette viderefører et historisk ulige udgangspunkt og en reproduktion af et kolonialt videnshierarki, hvor viden og teknologi i høj grad er blevet produceret og bestemt af Vesten, og dermed også har defineret, hvad der tæller som legitim viden.

Samtidig har vi undersøgt Masakhane, der gennem inddragelse af modersmålstalere og panafrikanske datasæt arbejder for at skabe mere kulturelt relevante og retfærdige input til sprogmodeller. Projektet Masakhane viser dog også, at reel indflydelse kræver både teknisk viden, ressourcer, digital infrastruktur og samarbejdspartnere.

Flere respondenter understreger desuden, at sådanne ’symbolpolitiske’ tiltag alene ikke er tilstrækkelige. Ægte inklusion forudsætter strukturelle ændringer i ejerskab og governance, så lokale aktører kan definere de teknologiske rammer på egne præmisser. Dette rækker ud over tekniske justeringer og handler også om en demokratisk fordeling af magt i forhold til udvikling, profit og styring. Det kræver, at teknologisk udvikling ikke blot tilpasses flere sprog, men omstruktureres, så flere aktører får reel beslutningskraft.

Samlet peger denne analysedel på, at AI og sprogmodeller ikke udvikles i et neutralt vidensvakuum, men formes af eksisterende magtstrukturer. Koloniale mønstre spiller stadig en rolle i, hvordan viden og sprog vurderes og anvendes, og dermed i, hvem der skaber diskurserne angående udviklingen og brugen af AI.

7.2 Analysedel 2

Gennem vores interviews, står det klart, at sprogmodel-teknologiers udvikling og dens indvirkning på global ulighed i høj grad afhænger af (manglende) regulering samt om viljen til at påtage sig et internationalt ansvar, for at fremme en mere etisk teknologiudvikling.
I det følgende afsnit analyserer vi, gennem fire afsnit, hvordan temaerne 'geopolitik', 'etisk teknologiudvikling', 'regulering', og 'ansvar' har betydning for global ulighed i forbindelse med generativ AI.

7.2.1 AI som geopolitisk teknologi

I dette afsnit undersøger vi hvordan kunstig intelligens er blevet en geopolitisk nøglefaktor, og hvordan vores respondenter forholder sig kritisk til magtforhold.
Vi trækker her på Thomsen, Just, Ebbensgaard, Rosman, Alayande og Sodemann, fordi de fra forskellige positioner belyser AI's magtdimensioner og rejser grundlæggende spørgsmål om magtstrukturer i det teknologiske landskab.

Thomsen fremhæver, at vi med styresystemer, søgemaskiner og sociale medier har skabt problemer, ved at tillade tech-monopoler på de her områder: "Og jeg tror, noget af det, som rigtig mange er bekymrede om nu, det er, at vi får også et monopol. Eller måske et oligopol med, lad os sige, tre store spillere, som sidder totalt på kunstig intelligens." (Thomsen: 7, 1-4).
Adams (2025) fremhæver, at adgangen til at forme AI ikke er tilfældig. Tværtimod påpeger hun, at monopol- og oligopol strukturer i tech-sektoren er en videreførelse af historiske magtforhold, hvor majoritetsverdenen ekskluderes fra centrale beslutninger. Det gør AI-udvikling til en ny form for global ulighed, snarere end et globalt fællesprojekt.

Set i et europæisk perspektiv vurderer Just, at den aktuelle udvikling i USA, hvor ideologiske strømninger former den globale AI-retning, både kan ses som en udfordring og en mulighed. Hun peger på, at en krisesituation kan skabe grobund for et mere mangfoldigt og løsningsorienteret AI-landskab, hvis det fører til større konkurrence og flere reelle alternativer. Ifølge Just er det altid problematisk, når magten koncentreres i monopollignende strukturer, og derfor bør pluralisme og åbenhed være centrale mål i udviklingen af fremtidige AI-systemer (Just: 5, 31-34).
Accelerationen af teknologisk udvikling er ikke ny, men hastigheden synes alligevel voldsommere end hidtil, som Suleyman (2023) beskriver som en teknologisk 'hyper-evolution'. Adams (2025) udlægning af det 'nye AI imperium', viser også hvordan skellene

ikke bare er mellem Vesten og Afrika, det sker også mellem EU og USA, hvor hun beskriver Europas rolle som globalt centrum, er aftaget. I forlængelse af det opstår en bekymring for, at USA dominere Europa, forklarer Ida Ebbensgaard:

> *"Ja det er jeg dybt bekymret for, jeg tror virkelig det er vigtigt at vi laver noget, som afspejler vores værdier. Og det tror jeg også er vigtigt, at Sydafrika gør. Ellers får du den vestlige synsvinkel ind ad bagdøren i alt, hvad du gør. Specifikt den amerikanske synsvinkel."* (Ebbensgaard: 7-8, 31-3).

Respondent Rosman beskriver to grundlæggende filosofier for, hvordan AI udvikles. Den ene lægger vægt på at bygge systemerne langsomt og sikkert, mens den anden er præget af en typisk startup-logik, hvor man lancerer hurtigt for at opnå markedsandele og først retter eventuelle fejl på bagkant (Rosman: 5, 9-12). Ifølge Rosman er det særligt store tech-virksomheder, der vælger den sidste tilgang, hvor kommercielle interesser, om at 'være de første' er vægtet højere end fejlfrie løsninger.
Som tidligere nævnt påpeger Adams (2025), at vi står overfor en ny magtudøvelse, med en ny verdensorden, der bygger videre på klassiske mønstre fra koloniale ulighedsstrukturer. Forskellen er nu bare, at det globale magtcentrum defineres ud fra det nye AI-imperium, en modalitet af magt distribueret mellem USA og Kina. Det er en af disse udfordringer Rosman spørger undrende til, om teknologierne udviklet i minoritetsverden kan oversættes og tilpasses til andre dele af verden, hvor de må imødekomme andre sprog, kulturer eller lokale behov (Rosman: 5, 28-31).
Vores to respondenter Rosman og Alayande er enige om, at AI er transformativ, og at der er uligevægt i gevinsterne ved teknologien.
Alayande italesætter at fokus er rykket fra 'AI safety' til 'AI security' (Alayande: 12, 26–27). Det betyder, at opmærksomheden ikke længere i samme grad ligger på ansvar og beskyttelse, men i højere grad på kontrol og magt. Alayande formulerer det sådan:

> *"Rather than focusing on real harms to people, we're back to the doomsday scare about existential risk. These are the kind of conversations that are going to then dominate AI ethics discussion.''* (Alayande: 12, 28-30).

Samtidig fremhæver Rosman, hvordan politiske partier i både USA og Kina fortsat taler om AI som et kapløb, hvor den, der udvikler teknologien først, også får magten til at lede verden (Rosman: 5, 32). Couldry og Mejias (2019) forklarer det som en dataficering af

befolkningsgrupper, hvor AI og tech-giganter udnytter ressourcer fra og i majoritetsverden. Ifølge Fairclough (2003) ser vi hvordan hvordan koloniale strukturer videreføres og hvordan der således opretholdes dominans over interesser og identiteter.
Vores respondent Morten Sodemann, klinisk Professor og forskningsleder ved Center for Sundhedsfilosofi og Etik ved Syddansk Universitet, supplerer teoretiseringen om datakolonialisme, med hvordan folk mister deres naturlige kritiske refleksion som et marginaliseret folk:

> *"Jeg har arbejdet i mange år i Vestafrika med sundhedsforskning. Jeg har set, hvordan de gik fra slet ikke at have papir til fax, og fra fax til computer og internet – de sprang hele den mellemliggende udvikling over.''* (Sodemann: 2, 14-16).

Disse strukturelle begrænsninger forstærker den ulighed, som også Sodemann fremhæver, når lande eller befolkningsgrupper implementerer ny teknologi uden at have gennemgået de tidligere teknologiske udviklingstrin, kan det føre til en ukritisk brug af teknologien (Sodemann: 2, 7-23). I sådanne tilfælde risikerer dele af befolkningen at blive hægtet af, hvilket underminerer demokratisk deltagelse (Seger 2023), hvor både tech-giganter og stater kan udnytte manglende kritisk bevidsthed til deres egen fordel. Der opstår en hegemonisk suverænitet, der udnytter denne naivitet, som er konsekvensen af at springe teknologiske udviklingstrin over.

Vi ser tech-giganter som magtfulde stater i sig selv, med en vidensdominans i stand til at kontrollere befolkninger, ved at marginalisere alternative perspektiver og erfaringer i en teknologisk udvikling. Som Fairclough (2003) pointerer, forskydes magtbalancen til en universel diskurs, der gør udemokratiske handlinger svære at være kritiske over for. Sodemann siger afsluttende om det nuværende geopolitiske landskab: "USA vil have, at AI skal udvikle sig fuldstændig ultraliberalt. De siger jo direkte til europæiske ledere, at de begrænser udviklingen ved at snakke om etik. Så det er Wild West." (Sodemann: 4, 29-31). Couldry og Mejias (2019) kalder det for en global dataudvinding, med mangel på gennemsigtighed og mangel på demokrati, der forskyder fokus væk fra de etiske og sociale konsekvenser.

7.2.2 Regulering af AI

Som del af vores elleve interviews, spurgte vi samtlige eksperter, hvordan vi bedst kan regulere udviklingen af AI.

Der hersker forskellige globale interesser og tilgange til udviklingen og anvendelsen af AI. Som tidligere beskrevet adskiller USA, Kina, EU og Afrika sig markant i både digitalt udgangspunkt og i deres syn på teknologisk udvikling, herunder hvilke reguleringsmæssige og etiske hensyn der skal tages.

I dette afsnit inddrager vi pointer fra både Høyrup, Alayande, Rosman og Mukiri-Smith, som er de respondenter, der har mest indgående kendskab til regulering. AI-regulering udvikler sig i mange forskellige tempi på tværs af verdensdele, og med forskellige formål.

For lande i majoritetsverdenen handler meget af den nuværende aktivitet om at tiltrække investeringer og signalere strategisk vilje, snarere end konkret og bindende regulering. Mukiri-Smith ser på regulering som et vigtigt skridt mod mere ligeværdig AI udvikling, men samtidig erkender hun, at kapacitetsopbygning mangler for at kunne regulere. (Mukiri-Smith: 7, 10-13). Som Alayande formulerer det, adskilles strategier fra regulering. AI-strategier i afrikanske lande, er ofte en slags forretningsstrategi, for at vise, at kontinentet er værd at investere i. Samtidig understreger han følgende: ”I think right now African countries are lagging behind in terms of regulation, and just to put it on the record, AI strategies are not regulation, even though we tend to conflate the two’’ (Alayande: 6, 21-22). Fælles for begge er en erkendelse af, at det politiske tempo i reguleringen halter efter både teknologisk udvikling og sociale behov, fordi de kommer fra et helt andet udgangspunkt end f.eks. Europa. I lyset af Sens kapabilitetsteori indebærer den manglende kapacitetsopbygning, at mange aktører ikke har reel mulighed for at omsætte AI-teknologiens potentiale og investeringer.

I en europæisk kontekst er det ikke manglen på strategier, men kompleksiteten i eksisterende eksisterende lovgivning og de geopolitiske hensyn, der bremser nye reguleringstiltag. Vores respondent Anette Høyrup, medlem af Dataetisk Råd og seniorkonsulent ved Implement Consulting Group, peger på, at EU-lovgivningen endnu ikke stiller krav til, i hvilket omfang man skal træne sprogmodeller på tværs af lande, og advarer om, at det kan føre til urimelige eller ulige sprogmodeller (Høyrup: 3, 18-23). Dermed fremstår reguleringen både som et forsinket og potentielt ineffektivt svar på hastigt eskalerende uligheder.

Høyrup uddyber også, at lovgivning altid kommer lidt på bagkant, fordi den tager lang tid at lave, og teknologien løber vanvittigt stærkt: “Men man prøver at lave så teknologineutral lovgivning som muligt, så man kan imødekomme udviklingen - uanset hvilken vej den løber, så er teknologien tænkt ind i lovgivningen.’’ (Høyrup: 10, 3-6).

Fra europæisk side fremhæver Høyrup, at reguleringen muligvis bør flyttes ud over EU og

placeres i internationale fora, fordi AI's globale rækkevidde overstiger, hvad europæisk lovgivning alene kan håndtere. Som Floridi og Cowl (2019) præsenterer 'dataetik', så har alle der arbejder med udvikling eller implementering af AI, en forpligtelse til at følge et etisk rammeværk, der går på tværs af landegrænser, og netop tager højde for kulturel og social mangfoldighed.

Behovet for international koordinering og samarbejde omkring AI-regulering nævnes af alle fire respondenter, men med markant forskellige vurderinger af, hvordan og af hvem det skal udvikles og implementeres. Spørgsmålet bliver også, om regulering eller mangel på, sker til fordel for minoritetsverdenen, på bekostning af majoritetsverden.

Rosman peger på, at forslag om FN-lignende organer for AI-tilsyn ikke er realistiske, hvis teknologien ikke allerede er bredt tilgængeligt. I stedet foreslår han en langt mere koordineret, tværsektoriel global indsats, som afspejler teknologiens hastighed og størrelse (Rosman: 3, 27-29).

Ideen om globalt tilsyn møder modstand, især fra perspektiver i majoritetsverdenen. Mukiri-Smith adresser spørgsmålet kritisk:

> *"We need some sort of global cohesion and maybe a global body that regulates AI, although that again raises questions about do we want another hegemonic power or body ruling over all of us telling us how we should regulate?"* (Mukiri-Smith: 4, 23-25).

Hun advarer om, at en sådan global instans risikerer at marginalisere stemmer fra f.eks. Afrika og Asien, hvis den amerikanske hegemoniske vidensstruktur videreføres (Fairclough 2003). Lignende forbehold findes hos Alayande, som fremhæver governance - at mange afrikanske reguleringstiltag i dag allerede er styret eller finansieret af internationale institutioner (Alayande: 7, 2-10). Udfaldet ved dette er, at en universel global ramme, der igen vil afspejle de stærkeste aktørers interesser, ikke nødvendigvis flertallets behov.

Selvom der er bred enighed om behovet for AI-regulering, er der langt fra konsensus om, hvad den konkret skal rette sig mod.

Rosman fremhæver kompleksiteten i at regulere sprogmodeller, fordi det endnu er teknisk svært at forudsige, hvad modeller vil kunne. Han skitserer et dilemma: "We can't simply say 'no model should do X, Y, or Z,' because it's technically very difficult to predict or control capabilities right now" (Rosman: 3, 41-42). I stedet foreslår han kontrol med infrastrukturen, f.eks. adgang til store mængder 'GPU'er' (graphics processing unit), som et potentielt mere

effektivt værktøj, og ser på hvordan den tekniske uforudsigelighed og manglende kontrol med modellernes kapabiliteter udfordrer muligheden for at sikre reel retfærdighed gennem regulering. Ifølge Sens (1993) retfærdighedsperspektiv risikerer det at føre til, at adgang til AI forbliver en formel mulighed for de få, snarere end en reel handlefrihed for de mange, særligt i kontekster, hvor sociale og infrastrukturelle forhold begrænser konverteringen af teknologi til faktisk livskvalitet.

AU-strategien fremhæver, at AI ikke blot forstærker uligheder mellem minoritets- og majoritetsverden, men også skaber intern ulighed. Derfor lægger strategien vægt på, at regulering skal tage udgangspunkt i afrikanske værdier, rettigheder og sociale kontekster, og ikke ureflekteret overtage reguleringsmodeller udviklet i minoritetsverden (African Union, 2024, s. 3, 32).

Mukiri-Smith bringer her en vigtig pointe i spil, at regulering må ikke kun handle om hvordan AI kontrolleres teknisk, men også om at fastsætte grænser for, hvad teknologien overhovedet bør bruges til, uanset kapabilitet. Det handler om etik og ansvar, ikke bare funktionalitet (Mukiri-Smith: 7, 9-14). Denne tanke spejles hos Couldry og Mejias (2019), der opfordrer til at gentænke hele udgangspunktet for teknologiudvikling ud fra en postkolonial forestilling om retfærdighed. AU-strategiens fokus på lokal kontekst og rettighedsbaseret governance bekræfter dette perspektiv, at en retfærdig AI-fremtid kræver, at majoritetsverden får mulighed for at definere egne rammer for teknologi, autonomi og viden, frem for at underlægges minoritetsverdens idealer og løsninger.

Hvor AU's strategi insisterer på en kontekstnær og rettighedsbaseret tilgang til AI, fremstår EU-forordningen som et eksempel på, hvordan AI-regulering formes af bredere geopolitiske interesser. Høyrup peger på, at EU-lovgivningen ikke kun handler om etik, men i høj grad også om at styrke konkurrenceevnen: "EU-lovgivningen er meget optaget af at sikre EU's konkurrence i forhold til USA og Kina (...) men også af at være til gavn og ikke gøre skade" (Høyrup: 8, 21-23). Dette bliver tydeligt i AI-forordningen, hvor reguleringen af AI forsøges harmoniseret på tværs af medlemsstater: "Formålet med denne forordning er at forbedre det indre markeds funktion ved at fastlægge ensartede retlige rammer, navnlig for udvikling, omsætning, ibrugtagning og anvendelse af kunstig intelligens-systemer'' (Europa-Parlamentet og Rådet, 2024, artikel 1).

Vi ser, hvordan EU søger at sikre lige konkurrencevilkår ved at forhindre, at udbydere af generative AI-modeller opnår fordele gennem lavere ophavsretlige standarder end dem, der gælder i Unionen (Europa-Parlamentet og Rådet, 2024, artikel 106).

Men selv med øget regulering peger Mukiri-Smith på, at spørgsmålet om hvordan vi

udvikler og regulerer teknologi, samt hvor den bør implementeres, ikke kan reduceres til lovgivning alene. For hende er strategisk regulering ét værktøj blandt flere, men der er behov for klare grænser for, hvilke former for anvendelse af teknologien der bør undgås (Mukiri-Smith: 7, 15).
Couldry og Mejias' (2023) vision om et fælles forum, der har flertallets interesser til gode frem for kontrol og profit, stemmer overens med Mukuri-Smith, som udviser bekymring for skellet mellem majoritet- og minoritetsverden. Denne bekymring åbner for næste afsnit, hvor vi undersøger perspektiver om en mere etisk teknologiudvikling.

7.2.3 Etisk teknologiudvikling

Vi stiller i følgende afsnit skarpt på, hvordan AI bliver til normbærende teknologi, hvor udviklingen i høj grad styres af tech-giganter i minoritetsverden.
Vi trækker på perspektiver fra respondenterne Ihle, Thomsen, Just, Høst, Alayande, Rosman og Arakkal, som alle på forskellig vis stiller skarpt på, hvilke værdier der indlejres i AI, og hvem de gavner.
Thomsen rejser en grundlæggende pointe om ansvar for fejl i store sprogmodeller: "Det er etisk set helt oplagt at spørge om, når man ved at man ikke kan eliminere dem [fejlene], men hvor mange skal man så acceptere, og hvordan skal man måle det?" (Thomsen: 1, 20-21).
Thomsen fremhæver hvordan generativ AI, i bedste fald kan fungere som et springbræt, der gør det muligt for majoritetsverden at springe nogle skridt over i en ellers langstrakt udviklingsrejse (Thomsen: 2, 17-18). Sådan kan man se AI som en ressource, der i teorien har potentiale til at udligne forskelle og understøtte lokal innovation. Men som Thomsen understreger, er der vigtige forudsætninger for, at dette potentiale kan realiseres i praksis. Det kræver nemlig både adgang til sprogligt relevante modeller og den nødvendige tekniske infrastruktur, f.eks. kodningsplatforme, software og hardware (Thomsen: 7-8, 20-6).
Her ser vi på baggrund af Sens kapabilitetsteori, at vi kan skelne mellem de som kan fravælge AI, og de som ikke har kapabilitet til at tilvælge den. I mange sammenhænge forbliver teknologien passiv eller ekskluderende, fordi kontekstuelle 'konversationsfaktorer' (Sen 1993), f.eks. sprog, infrastruktur eller digital dannelse, ikke er til stede. Thomsens pointerer betydningen af udviklingen af tekniske standarder, for en mere etisk teknologiudvikling:

> *''Der tror jeg simpelthen at vi er nødt til at udvikle nogle tekniske standarder og de skal selvfølgelig udvikles på baggrund af en forståelse af hvordan de her problemer typisk opstår. Så der er dels en proces med at gøre dyrekøbte erfaringer, hvor vi*

> *kommer til at ramme ved siden af en del gange med den her teknologi - fordi den er ny og det er svært at forudse alle de problemer der kan opstå.*'' (Thomsen: 10, 16-20).

Denne tilgang til teknologiens rolle understøttes af respondent Ihle, der beskriver et skifte i den generelle bevidsthed om etik i AI. For få år siden var etik fraværende i debatten, men i dag er det ikke længere muligt 'at slippe afsted med' ureflekteret udvikling i sprogmodeller, forklarer hun (Ihle: 6, 10-11). Floridi og Cowls (2019) dataetiske udlægning fremhæver dog, at denne udvikling er baseret på initiativer fra vestlige ideologier, hvor etik også har forskellige sondringer, om man er i USA eller i Europa. De peger på, at etik bør implementeres i udviklingen af AI, uafhængigt af hvilket stadie eller profession det formidles fra. Alligevel er det stadig uklart, hvem der bærer ansvaret for at sikre, at teknologier faktisk lever op til etiske standarder, og hvordan disse standarder skal udvikles, for som Ihle pointerer:

> *"Du kan ikke bruge det samme governance værktøjer til at monitorere modeller, der er lavet på andre sprog nødvendigvis. Det vil være godt at selv de små sprog havde deres egne sprogmodeller eller i hvert fald havde nogle partnere til at udvikle nogle modeller med dem."* (Ihle: 5, 12-14).

Spørgsmålet om ansvar i udviklingen af kunstig intelligens behandles i det følgende analyseafsnit. Her undersøger vi, hvordan der på trods af en stigende erkendelse af behovet for mangfoldighed i datasæt, fortsat mangler klare retningslinjer for, hvordan dette omsættes til konkret udviklingspraksis.
Respondent Høst problematiserer den teknologiske AI-udvikling ved metaforen om et bakspejl, som også findes i teoretiseringen af Marshall McLuhan; at vi anskuer nutiden ved at se i bakspejlet og gå baglæns mod fremtiden (Prescott 2016):

> *"Det handler jo alt sammen om data og det, der er puttet ind i de store sprogmodeller, er vores historie. Det fungerer som et bakspejl, som der er mange der beskriver, at det er som at køre mens man kigger i bakspejlet."* (Høst: 2, 2-4).

Dermed beskriver Høst den samme problematik som vi ser i vores koloniale teorier, at historiske data som skrives ind, baseres på det ulige samfund, der har eksisteret, og reproducerer dermed et datakolonialt verdenssyn (Couldry & Mejias 2019). Faircloughs diskursanalyse er nyttig til at forstå, hvordan AI fremstår som objektiv og neutral teknologi,

men i praksis formidler og forstærker eksisterende magtforhold. Når historiske datasæt ureflekteret gøres til grundlag for ny teknologi, bliver sociale uligheder indlejret som naturlige standarder. Høsts metafor om bakspejlet understreger netop, hvordan tidligere magtstrukturer bliver kodet ind i nutidens og fremtidens generative AI-systemer.

Selvom meget af debatten om AI etik centrerer sig om regulering, risiko og strukturel ulighed, peger flere af vores respondenter også på vigtigheden af at synliggøre eksisterende alternativer og handlemuligheder. Just understreger, at fokus på store tech-giganter kan overskygge de mange mindre initiativer, der allerede eksisterer, såsom Masakhane. "Hvorfor har EU ikke en tech-gigant?" er det forkerte spørgsmål ifølge Just. I stedet bør vi spørge: ''Hvordan bliver vi opmærksomme på de alternativer, der allerede eksisterer?'' (Just: 8, 5-8). Problemet er ifølge Just ikke nødvendigvis fraværet af alternativer, men at disse ikke er synlige nok i en offentlighed domineret af få tech-giganter. Samtidig fungerer den universelle tankegang som et 'etisk korrektiv' overfor idéen om, at der kun findes én teknologisk udviklingsvej. Som Helm et al. (2024) peger på, er dette Vestens ideal på anvendelse af teknologi - en tankegang som vi ser anerkendt på tværs af litteraturen, en indirekte eller sen konsekvens af kolonialisme. At anerkende værdien i små og lokale løsninger åbner for en mere kontekstbaseret tilgang til AI, hvor etik ikke nødvendigvis institutionaliseres centralt, men kan opstå i praksis. Det indebærer også at se teknologi som noget, der kan tilpasses lokale behov snarere end noget, der nødvendigvis skal importeres som færdige løsninger.

Flere respondenter peger på betydningen af sproglig mangfoldighed og inklusion som en del af løsningen. Det er væsentligt at nævne, at muligheden for at skabe egne systemer på egne præmisser ofte starter med synliggørelsen af alternativer og en etisk vilje til at investere i andet end det dominerende.

Sen (1993) fremhæver vigtigheden ved kontekstualisering - at retfærdighed og kapabiliteter altid må anskues i forhold til sted og sociale forhold, hvor postkoloniale realiteter også kan variere meget fra hinanden.

Flere af de internationale respondenter fremhæver, hvordan udviklingen af kunstig intelligens globalt set stadig er præget af en top-down tilgang, hvor vestlige aktører definerer retningen, mens et land som Sydafrika må tilpasse sig allerede udviklet systemer. Alayande beskriver, hvordan mange initiativer i Afrika fokuserer på lokal dataindsamling og opbygning af sprogmodeller på lokale sprog: "There's a lot of move towards local language AI, so everyone is kind of doing local data collection and building indigenous LLMs to ensure independence and inclusion.'' (Alayande: 3, 5-6). Men samtidig peger han på, at de

dybereliggende strukturelle udfordringer ofte overses. Problemet er ikke nødvendigvis manglen på sprogdata, men som Alayande beskriver, at afrikanske innovatører ikke har de rigtige ressourcer (Alayande: 16, 12-13). Det handler om fravær af kapital, teknologisk infrastruktur og politisk autonomi, som gør det svært for afrikanske aktører at udvikle egne løsninger på egne præmisser. Med Adams (2025) analyse om LLM'er og postkolonial teknologikritik bliver det tydeligt, hvordan udviklingen af AI i minioritetsverden ikke blot skaber ulighed, men aktivt reproducerer afhængighed gennem teknologiske standarder, der sjældent afspejler lokale behov. Samtidig viser Sens kapabilitetstilgang, at det ikke er nok med adgang til teknologi - friheden til at handle kræver reelle muligheder. Når Alayande formulerer mangel på ressourcer, er problemet ikke blot fraværet af data, men fraværet af evnen til at omsætte teknologisk adgang til faktisk handlekraft og meningsfuldt ejerskab.
I stedet for at adressere disse forhold, bliver udviklingen ofte symptomatisk og kortsigtet, præget af topstyrede partnerskaber og internationale projekter, der sjældent sikrer langsigtet kapacitetsopbygning forklarer Alayande (Alayande: 5, 3-5). Denne bekymring deles af Arakkal, der fremhæver, hvordan ægte inklusion kræver gensidige partnerskaber, hvor teknologiske løsninger udvikles i tæt samspil med dem, de er tiltænkt (Arakkal: 8, 7-13). For hende handler etik ikke blot om at udforme standarder, men om at skabe rum for dialog og reel deltagelse og samarbejde med individer fra det afrikanske kontinent (Arakkal: 3, 34-37). Hun peger samtidig på, at dette kræver både finansiering og adgang til teknologiske kompetencer, og at mange initiativer i dag savner netop disse forudsætninger:

> *"The large sort of US based industries aren't taking into account what Africans need. It's taking into account what the US needs, not by default or of any means. And the only way to fix that is to ensure that we have data that is equally on the same level - So from a scale perspective and a coverage perspective''* (Arakkal: 7, 7-9).

Alyande og Arrakal peger således på et centralt paradoks, der tales meget om inklusion fra stater, men de strukturelle betingelser for at realisere den mangler ofte i løsningerne.
Ifølge Couldry og Mejias (2019) bygger dette paradoks på forestillingen om, at data er en ejerskabsløs ressource, der frit kan udvindes globalt og raffineres i minoritetsverden. Det skaber det, de betegner som *datarelationer*, hvor menneskers sociale og kulturelle liv i majoritetsverdenen udvindes som råstof til udviklingen af teknologier, som de sjældent har reel indflydelse på. Dette hænger tæt sammen med Benders (2021) kritik af, hvordan sprogmodeller og andre AI-systemer ofte præsenteres som neutrale og autonome, hvilket

bidrager til en dehumanisering og en tilsløring af de menneskelige valg, interesser og magtstrukturer, der former teknologien. Ved at tilskrive teknologien menneskelige egenskaber risikerer man ifølge Bender at ignorere det ansvar og de aktører, der reelt træffer beslutningerne bag systemerne.
Arakkals udsagn viser desuden, hvordan dette fortsat udspiller sig i praksis. Det gør de etiske overvejelser både mere presserende og langt mere komplekse, at AI fungerer som geopolitisk magt bestemt af tech-giganter.
Rosman pointerer, at AI er forskellig fra tidligere teknologier, fordi hvis vi tager fejl, er det ikke sikkert, at vi kan rette op på det (Rosman: 5, 14-15). Han kritiserer princippet om 'move fast and fix later' og erkender samtidig, at sådanne samtaler er svære at realisere i en situation, hvor teknologien allerede er bredt distribueret og svær at kontrollere, særligt når 'open source-modeller' muliggør, at millioner af aktører udvikler AI uafhængigt.
Historisk set peger Adams (2025) på, at teknologi har været et magtredskab anvendt til at øge egen stats styrke, alt imens andre stater holdes tilbage, men at i dag har teknologisk konkurrenceevne givet en gennemtrængende øget global status.

Det politiske spændingsfelt bliver særligt tydeligt i diskussionen om internationale fora og topmøder. Alayande beskriver mange AI-summits som symbolske, uden reel effekt (Alayande: 11, 26-27). Alligevel anerkender han, at der er sket fremskridt, særligt ved topmøder som Paris AI Summit, hvor deltagelsen har været mere global og inkluderende end tidligere (Alayande: 10, 23). Spørgsmålet er dog, om denne inklusion rækker ud over det symbolske og bliver til reelt medejerskab over normer og retning, eller, som Couldry og Mejias (2019) beskriver det, marginaliserer alternative perspektiver og erfaringer, på baggrund af mangel på viden set fra minoritetsverden synspunkt.
Rosman og Alayande deler en erkendelse af, at AI-etik ikke kan være ét lands eller én branches ansvar alene. Teknologiens rækkevidde kalder på strukturer, der kan sikre repræsentativitet, fordeling og fælles ansvar. Men samtidig viser deres udsagn, hvor vanskeligt det er at skabe globale mekanismer, som både har reel beslutningskraft og undgår at reproducere eksisterende magtforhold.

Respondent Arakkal udfordrer ideen om, at et globalt organ kan sikre etisk AI-brug alene. Hun mener, at både udviklere og brugere bør have etisk medejerskab, fordi beslutninger og konsekvenser sker lokalt, selv når teknologien er global. "It comes down to the individuals who are building these things too – and also individuals that are utilizing it" (Arakkal: 3, 1-3). Hun fremhæver vigtigheden af, at organisationer selv påtager sig ansvar og handler aktivt i forhold til, hvordan AI implementeres og anvendes, og ikke bare følger

retningslinjer udefra. Et andet centralt punkt for Arakkal er behovet for større gennemsigtighed omkring datagrundlag og kontekst (Arakkal: 3, 24-26). Når brugere og organisationer ikke ved, hvad modellerne er trænet på, eller hvilke antagelser de bygger på, bliver det vanskeligt at handle etisk, eller at stille kritiske spørgsmål. Her bliver transparens ikke kun en teknisk udfordring, men et etisk krav, fordi det muliggør informeret handling og ansvarlig beslutningstagning, som Holdsworth (2023) i litteraturen centralt inkluderer gennemsigtighed, hvor den menneskelige evaluering ses i udviklings- og beslutningsprocesser.

Afslutningsvis peger Arakkal på, at sådanne krav om deltagelse og repræsentation ikke realiseres af sig selv. I stedet påpeger hun vigtigheden ved etablering af rum og fora, hvor aktører fra forskellige dele af verden faktisk kan mødes og udveksle erfaringer (Arakkal: 8, 20-26). For hende er etisk AI en proces, der skal forankres i inkluderende fællesskaber, hvor repræsentationen af dem, teknologien påvirker, er en forudsætning for legitim udvikling. Med dette perspektiv placeres data etik som et fælles ansvar, hvor legitimitet ikke opstår gennem central styring, men gennem dialog, åbenhed og lokal forankring. I sidste ende handler etik ikke kun om intentioner, men om strukturer, der sikrer gennemsigtighed, inklusion og retfærdig fordeling af teknologiens udbytte, som vi ser i tråd med Floridi og Cowls' (2019) principper om 'retfærdighed' og 'gavnlighed'. Dermed skal etisk teknologiudvikling ikke blot forudsætte, at AI fremmer velfærd, men at den samtidig beskytter de mest sårbare og understøtter et globalt ansvar for, hvem der får adgang, medbestemmelse og gavn af teknologien.

7.2.4 Ansvar og aktører

Under interviewene spurgte vi alle vores respondenter om ansvar - hvem har ansvaret for at sikre, at AI udvikles forsvarligt og inkluderende? Er det et ansvar, som ligger hos tech-industrien, lovgivere eller noget tredje? Til de danske interviews spurgte vi ind til Danmarks og EU's rolle i udviklingen af ansvarlig AI. Til de internationale respondenter spurgte vi ind til samarbejdet mellem vestlige aktører og den Afrikanske Union (Bilag: Interviewguide). Til svar på disse spørgsmål trækker vi i dette afsnit på interviewene med Ebbensgaard, Thomsen, Høyrup, Rosman og Alayande, som bidrager med både internationale og nationale perspektiver.

Ebbensgaard beskriver, hvordan brugen af sprogmodeller inden for kort tid vil være almindelig i arbejdslivet, og det er op til os, om vi vælger at købe amerikanske produkter,

eller vælge en anden vej: “Vi kan ikke sige til virksomheden, at du skal lave dit produkt på en anden måde. Det er svært hvis man ikke har regler om måden de udvikler på, er ulovlige eller direkte forurenende.’’ (Ebbensgaard: 3, 5-10).

En lignende pointe fremføres af Thomsen, der afviser, at tech-industrien selv kan forventes at tage ansvar: “Ansvaret kan næppe ligge hos tech-industrien. Jeg tror ikke, at tech-industrien er en værre industri end alle mulige andre industrier” (Thomsen: 3, 1-3). For ham er det ikke et spørgsmål om moral, men om strukturelle incitamenter, og historisk erfaring viser, at markedsdrevne industrier sjældent handler etisk uden tydelig regulering udefra.

Disse udsagn indrammer det grundlæggende spørgsmål: Når teknologier har globale konsekvenser, men udvikles under private interesser og uden globale spilleregler, hvem har da det første og sidste ansvar?

Ifølge Adams (2025) er skellet mellem AI’s vindere og tabere ikke tilfældigt, men både kønnet, racialiseret og geografisk. Hun argumenterer for, at ansvaret for de negative konsekvenser bør bæres af dem, der i dag høster fordelene, for det er også disse aktører, der har størst mulighed for at forme AI som et redskab for global velfærd. Udvikles teknologien alene med henblik på profit og underholdning for en privilegeret minoritet, vil den aldrig komme de marginaliserede grupper til gavn.

Et af de perspektiver der kom frem i undersøgelsen var, at ansvaret for AI ikke nødvendigvis kun kan placeres hos tech-industrien eller stater, men også hos det enkelte individ.

Ebbensgaard udtrykker frustration over en passiv kritik af AI og opfordrer i stedet til teknologisk handlekraft: “Jamen, ved du hvad kammerat? Så lav det selv. Kom nu i gang. Lad være at tale om det. Gør det. Find ud af, hvad GitHub er” (Ebbensgaard: 11, 1-3). Det individuelle ansvar forstås her som en aktiv deltagelse, ikke alene som bruger af teknologien, men som medudvikler. Når Ebbensgaard taler for individuelt ansvar som teknologisk deltagelse, synliggør det den menneskelige handlekraft bag AI. Dog kan man med at flytte ansvaret til individet, risikere at overskygge strukturelle uligheder. Crawford (2021) advarer imod, at det kan skjule det politiske og økonomiske ansvar, der ligger hos dem, der har ressourcerne til at udvikle og styre AI-systemerne.

For Ebbensgaard handler ansvar ikke blot om at forstå og anvende teknologien, men om at mestre den: “Send børnene til Coding Pirates. Send pigerne til Hello Ada. Efteruddan medarbejdere. Vis det på skolerne” (Ebbensgaard: 11, 10-11). Teknologisk myndiggørelse bliver dermed en form for borgerdannelse, hvor kompetencer i kodning og AI-forståelse ses som en forudsætning for, at samfundet kan tage ansvar for sin teknologiske retning, nedefra.

Samtidig udstiller Ebbensgaards synspunkt en spænding i ansvarsfordelingen. Set med Sens (2006) tilgang til 'comparative justice' er det afgørende ikke blot at pege på formel adgang til teknologi, men at vurdere, hvilke reelle muligheder mennesker har for at handle. Ansvar må derfor knyttes til de kapabiliteter, borgere faktisk råder over, socialt, økonomisk og teknologisk, og ikke antages for at være ligeligt fordelt. Det fordrer en differentieret forståelse af ansvar, hvor både handlekraft og strukturelle begrænsninger medtænkes.

Flere respondenter problematiserer, at tech-industrien i dag ikke mødes med de samme krav og sikkerhedsstandarder som andre højrisiko sektorer. Thomsen undrer sig over, hvorfor AI og digital teknologi ikke reguleres med samme alvor og præcision som eksempelvis medicinal- eller kemikalieindustrien, der traditionelt underlægges streng kontrol for at beskytte borgere mod skade (Thomsen: 3, 5-7). Pointen er, at når teknologier har kapacitet til at påvirke liv og samfund i stor skala, bør de underlægges tilsvarende ansvarlighed. Thomsen mener, at lande som USA og EU har de strukturelle muskler til at forme AI-politikker, mens mindre lande som Danmark og mange lande i majoritetsverden er langt dårligere stillet: "Hvis det er svært for Danmark, så er det jo kun endnu sværere, hvis man er et afrikansk land syd for Sahara" (Thomsen: 3, 12-13). Ansvar bliver dermed et spørgsmål om geopolitisk kapacitet, hvem der kan og bør tage ansvar, og som Fairclough (2003) anerkender, hvor stater eller institutioner har agens til at ændre kursen.
Respondent Høyrup supplerer dette syn med en detaljeret forståelse af aktørens ansvar for at intervenere. Hun understreger, at både tech udviklere, distributører, brugere og myndigheder bør være omfattet af ansvar, afhængigt af deres rolle og teknologiens risikoprofil (Høyrup: 3, 3-7). Tech-giganter i USA er formelt private aktører, og til det påpeger Høyrup, at de underlægges europæisk lovgivning, når deres produkter bevæger sig ind på EU's marked: "Ligesom sociale medier er jo også underlagt dansk lovgivning, fordi de indsamler data i Danmark" (Høyrup: 2, 12-13). Hun ser derfor et delt ansvar på tværs af både tech-industrien og jurisdiktioner. Men spørgsmålet er, om det er nok. Høyrup minder om, at AI også skal være til gavn for mennesker - og at dette mål kan glide i baggrunden, hvis udviklingen styres af profit og ikke etisk ansvar: "Hvis vi ikke tænker usikkerheder og mindre kulturer ind, vil AI kun gavne den vestlige verden, som i forvejen er bedre stillet" (Høyrup: 7, 19-22).

Fra et internationalt perspektiv er det især USA, der udpeges som magtcentrum, og problem. Respondent Rosman peger på, at USA i årtier har fungeret som magnet for AI-talenter på tværs af landegrænser, hvilket har svækket kapacitetsopbygning i resten af verden: "Many other places are going to struggle to prepare their countries for what's coming" (Rosman: 7, 8-9). Hvis ansvar kræver viden og kompetencer, udpeger Rosman her en

strukturel skævhed, som kun forstærkes af tempoet i AI-udviklingen: “We don't have time, and in many cases, we don't have the money or the people.” (Rosman: 7, 17). Alayande tilføjer en klar kritik af de vestlige landes forventninger til lande i majoritetsverden. Med reference til energipolitik peger han på, hvordan man i dag kræver, at afrikanske lande udvikler bæredygtig infrastruktur, uden at have fået lov selv til først at industrialisere, og samtidig har bidraget minimalt til globalt klimaaftryk:

> *‘’China is the big contributor. The West has historically been a big contributor and now is kicking away the ladder. African countries are all of a sudden expected to build sustainable/green compute infrastructure alongside countries that are much more advanced in renewables. It's a very nice thing to build sustainable data centers and things like that. But I think there has to be room to explore whatever energy sources are suitable for African countries’ data centre expansion. We need to have as much compute power as possible first.”* (Alayande: 13, 22-26)

Ansvaret for teknologisk bæredygtighed bliver til et urimeligt krav i majoritetsverden, når det ikke ledsages af investeringer, støtte eller reel indflydelse. Samlet peger citaterne på, at spørgsmålet om ansvar ikke kan skilles fra spørgsmål om magt og struktur. Når minoritetsverden formulerer forventninger til majoritetsverden, er det et eksempel på, hvordan sprog og ansvar italesættes som styringsmekanismer - hvilket vi i tråd med Fairclough (2003) forstår som et centralt middel til magtudøvelse.
I denne kontekst placerer Thomsen ansvaret i en bredere global retfærdighedsramme, hvor teknologisk udvikling ikke kun er et spørgsmål om fremskridt, men også om moralsk forpligtelse. ”Velstående lande har meget store forpligtelser til at hjælpe mindre velstående lande - også langt større end de løfter lige nu” (Thomsen: 12, 3-4). Men han understreger også, at dette ansvar ikke er enkelt at løfte i praksis. Der må ske prioriteringer mellem behov: “Det kan godt være, de har brug for kloakering, grundskoleuddannelse eller et sygehusvæsen, det er ikke sikkert, digital infrastruktur kommer først” (Thomsen: 12, 5-11). Her tydeliggøres et dilemma mellem teknologisk transformation og basale velfærdsbehov, som skaber behovsstruktur, der kan spores til en vestlig videnskabelig tradition (Schwartz 2022).
Sen (1993) understreger, at retfærdighed ikke alene handler om at sikre adgang til bestemte ressourcer, men om at vurdere, hvilke kapabiliteter der er mest afgørende i den konkrete kontekst. Det er ikke de samme behov, der nødvendigvis skal komme først – men de behov, der muliggør andre muligheder.

Thomsen peger på en vigtig pointe i denne forstand; At ansvar ikke kun handler om, hvem der udvikler teknologien, men også om, hvem der har etiske forpligtelser til at sikre, at flere får mulighed for at drage nytte af den. I dette perspektiv bliver Vestens ansvar ikke blot teknisk, men fundamentalt politisk og moralsk.
Adams (2025) kommer med EU som et eksempel på at tage ansvar, at de har gjort det forbudt at bruge live AI-drevne biometiske systemer, der anses for at udgøre en høj risiko i konflikt med grundlæggende menneskerettigheder (Adams 2025: 53).

For flere af vores internationale respondenter handler ansvar ikke kun om at blive inkluderet i andres strukturer, men om retten til selv at definere teknologiens retning. Alayande fremhæver, at spørgsmålet om sprog og adgang i AI-udvikling i virkeligheden handler om noget dybere: "It's the question of ownership (…) a kind of cultural sovereignty in itself." (Alayande: 3, 21-22). Egen styring kræver ikke blot teknisk infrastruktur, men kontrol over, hvad AI bruges til, og hvilke værdier det afspejler. Han efterlyser også, at globale AI-processer gøres mindre elitære og mere tilgængelige for dem, der reelt bliver påvirket af teknologierne, ved at gentænke hvem der er i rummet, når vi tager beslutningerne (Alayande: 10, 19-20). Samtidig advarer han mod asymmetriske partnerskaber, hvor lokale aktører i Afrika leverer og indsamles data fra, som ender med at styrke vestlige konkurrenter (Alayande: 18, 16-17). Det viser, at egen styring ikke blot er en ambition, men en reel bevægelse, der tager ansvar og kontrol over udviklingen. Som Seger (2023) fremhæver, må demokratiseringen af AI ikke reduceres til teknisk adgang eller åbenhed alene, men forstås som et komplekst spørgsmål om, hvem der har mulighed for at udvikle, anvende, styre og drage nytte af teknologien.

7.2.5 Delkonklusion

Denne delanalyse har undersøgt, hvordan ambitionen om en retfærdig, etisk og inkluderende udvikling af generativ AI udfordres i praksis. Vi har vurderet den nuværende teknologiske udvikling efter reel og lige muligheder for at forme og drage nytte af teknologien.
Vores analyse viser, at generativ AI i dag fungerer som en geopolitisk magtfaktor: Udvikling og kontrol er koncentreret hos få dominerende aktører i minoritetsverdenen, mens majoritetsverdenen i praksis fastholdes som dataleverandør snarere end medskabere. Manglende eller utilstrækkelig regulering samt et svagt internationalt samarbejde cementerer denne ulighed og gør AI til et redskab, der tjener de teknologisk førende lande frem for et fælles globalt ansvar. Etisk teknologiudvikling kan derfor ikke begrænses til tekniske

justeringer eller symbolpolitiske tiltag. Ejerskab, teknologisk kapacitet og faktisk adgang afgør, hvem der reelt kan forme og drage nytte af AI og hvor formel adgang ikke garanterer reel handlefrihed.

Det etiske fokus må derfor rettes mod at definere acceptable grænser for hvornår teknologien påfører reel skade og etablere ansvarlige procedurer for, når systemer fejler. Samtidig må ansvarsbegrebet udvides, hvis ikke det kan placeres alene hos tech-industrien, stater eller individuelle brugere, kræver det koordinerede indsatser på tværs af sektorer, klare governance-strukturer og målrettet kapacitetsopbygning. Uden sådanne systemiske ændringer risikerer AI-udviklingen at øge global ulighed med tech-monopoler, koloniale videreførelser i form af indlejret bias og en forstærkning af vidensdominans.

8. DISKUSSION

8. Diskussion

Efter at have analyseret fund og dilemmaer frem i de foregående afsnit vil vi i dette kapitel diskutere, hvad vores undersøgelse fortæller om generativ AI's indvirkning på global ulighed. Med fokus på, hvordan adgang til teknologi, investering i AI-kompetencer og respekt for lokale normer, sprog og verdenssyn er afgørende for en retfærdig digital udvikling, har vi undersøgt hvordan bidrager udviklingen og træningen af generativ AI til global ulighed, og hvilke etiske dilemmaer opstår i bestræbelserne på at gøre AI mere retfærdig og inkluderende.
Vi vil diskutere de mest centrale og generelle problematikker, som har vist sig gennem vores analyse af problemstillingen.

AI som magtredskab
Et dilemma, der er blevet tydelig gennem vores interviews med aktører fra majoritetsverdenen, med særligt fokus på Sydafrika, er, at mange lande ser sig nødsaget til at forholde sig aktivt til AI, ikke nødvendigvis fordi de prioriterer det højt internt, men fordi det er en forudsætning for at følge med den globale digitale udvikling. Som vores respondent Alayande, forsker ved Global Center on AI Governance pointerede, er denne prioritering nødvendig for at kunne tiltrække investeringer og blive anerkendt som en ligeværdig global aktør (Alayande: 6, 26).
I flere interviews med sydafrikanske respondenter blev det som tidligere nævnt, at AI til tider kan opleves som en påtvunget agenda i lande, der stadig kæmper med basale infrastrukturelle udfordringer og presserende socioøkonomiske problemer.

På den ene side ser vi, hvordan AI-udviklingen i Afrika bliver mødt med både ambition og strategisk vilje. Dette har vi tidligere vist afspejlet i AU's AI-strategi, hvor man udtrykker et ønske om at deltage aktivt i udviklingen og positionere kontinentet som en aktør i det teknologiske landskab. Det viser, at der er både alvor og villighed til at engagere sig geopolitisk i AI. Samtidig understreger vores sydafrikanske respondent, universitetsprofessor Benjamin Rosman, hvordan interessen og udviklingen inden for AI i regionen vokser eksponentielt år for år. Det vidner også Masakhane-projektet om, som med sin vækst og fokus på lokale sprog og kapacitetsopbygning, konkretiserer ambitionen om en afrikansk tilstedeværelse i AI-udviklingen, via repræsentation af sprog i store sprogmodeller.

På den anden side rejser udbredelsen af AI en række bekymringer hos vores respondenter og i litteraturen. For mange lande i den globale majoritetsverden er det ikke blot

et spørgsmål om teknologisk interesse, men om et hurtigt voksende pres fra minoritetsverden om at prioritere AI over andre dagsordener. Denne prioritering kan dog risikere at føre til øget intern ulighed, da gevinsterne fra AI primært kommer visse befolkningsgrupper til gode - ofte dem, der allerede har adgang til ressourcer, teknologisk infrastruktur og digitale hjælpemidler.
I et land som Sydafrika, der i forvejen er et af verdens mest ulige lande, kan denne udvikling af AI risikere at forstærke eksisterende sociale og økonomiske skel endnu mere. Både Adams' (2025) begreb om digital kolonisering og Couldry og Mejias' (2019) teori om datakolonialisme belyser, hvordan tidligere magtstrukturer gentages i en ny teknologisk form, der fortsat presser befolkningsgrupper til at være afhængige af mere velstillede lande. I den forstand bliver det postkoloniale perspektiv afgørende for at forstå, hvordan intentioner om inklusion i udviklingen af store sprogmodeller kan dække over en dybere afhængighed af teknologier, som er formet af normer og interesser uden for majoritetsverdenens selvbestemmelse.

Vores respondenter udfordrer ideen om inklusion i AI-udvikling, og hvad nytter det at inkludere nogen i teknologien, hvis majoriteten reelt ikke har mulighed for at bruge, forstå, eller forme den. Respondent Morten Sodemann, professor og forskningsleder ved Center for Sundhedsfilosofi og Etik, pointerer, at teknologi ikke har nogen indbygget etik, og at teknologiske systemer er udviklet i lukkede miljøer, af fagpersoner uden blik for brugernes konkrete behov (Sodemann: 2, 28-30). Han pointerer at systemerne ikke er bygget til virkeligheden, men kan virke fremmedgørende og der mangler en ansvarstagen mellem praksis og det digitale univers. Netop her bliver kapabilitetsteorien afgørende i vores undersøgelse, med hensyn til, at det ikke handler om adgang alene, men om evnen til at omsætte teknologien til relevante AI løsninger. Den digitale kløft handler også om det udgangspunkt, teknologien udvikles fra, hvordan den designes, og hvem den fungerer for i praksis. Vi risikerer, at inklusion blot betyder, at underrepræsenterede grupper skal tilpasse sig de normer og behov, som allerede er indlejret i Vestens teknologi. På samme måde risikerer vi, at retfærdighed reduceres til et spørgsmål om, hvorvidt teknologien er tilgængelig, frem for om den er udviklet med udgangspunkt i forskellige gruppers reelle behov og udgangspunkter. Det peger også på en mere grundlæggende kritik af teknologisk determinisme, idéen om, at digitalisering og AI nødvendigvis og uundgåeligt er vejen frem.

Et af de mest påfaldende og måske mest foruroligende fund i vores undersøgelse om AI-udvikling er, hvor lidt magt og indflydelse de fleste har over teknologier, der i stigende grad former deres liv. Vores forskningslitteratur, teori og interviews peger på, at AI i dag ikke

blot er teknisk avanceret, men politisk og ideologisk betinget. Den magt, der ligger hos tech-virksomheder, er ikke blot økonomisk, men kulturel, normdannende og identitetsskabende. Med afsæt i Couldry og Mejias' (2019) teori om datakolonialisme bliver AI en måde at udøve magt på, hvor mennesker reduceres til datakilder. Magt i denne forstand udøves ikke blot over kroppe eller territorier, men over selve betingelserne for at deltage i det digitale liv.

Regulering af AI-teknologi har enten været stærkt forsinket eller meget begrænset, som det også fremgår i forskningsfeltet. Denne reguleringsmæssige tøven skyldes ikke blot teknologiens hastige udvikling, men også et bevidst strategisk valg. Ved f.eks. ikke at underskrive den fælles erklæring om en inkluderende og etisk AI-udvikling på Paris AI Summit signalerede lande som USA og England, at international teknologisk dominans vægter højere end fælles etiske forpligtelser, som ellers var intentionen med erklæringen. I denne sammenhæng bliver AI ikke blot et teknologisk fremskridt, men en platform for geopolitisk magtkamp.

Flere respondenter påpeger desuden, at vi har købt os ind på præmisser, vi ikke forstod i den tidlige fase af digitaliseringen, til en grad hvor sociale medier, algoritmer, overvågningskapitalisme og datamining er blevet accepteret uden større forbehold.
Som beskrevet i problemfeltet, sidder amerikanske virksomheder som Microsoft, NVIDIA, Apple, Google og Meta på størstedelen af verdens markedsudviklingen inden for generativ AI. Dermed har de magten til at sætte retningen for udviklingen af teknologien, som kan udgøre et demokratisk problem, hvis brugere af teknologien ikke formår at forstå og kritisk anvende teknologien. Respondent og Vice President for Ethics, Governance & Policy ved Responsible AI Future Foundation, Laura Haaber Ihle, påpeger, gælder udfordringen også for offentlige institutioner, der selv skal indkøbe og implementere avancerede teknologier, uden nødvendigvis at have de nødvendige ressourcer eller tekniske kapaciteter til at forholde sig kritisk til dem. Hun fremhæver et eksempel fra en dansk kommune, hvor en medarbejder på egen hånd skal anskaffe store sprogmodeller og træne personale i deres brug, uden nationale rammer eller retningslinjer, der understøtter ansvarlig anvendelse og håndtering af de etiske og praktiske konsekvenser (Ihle: 9, 2-11). Det rejser spørgsmålet, om den accelererende udvikling af generativ AI har været med til at give tech-virksomhederne frit råderum, og om de i dag har overskredet gængs praksis for offentlig kontrol og regulering.

Retfærdig og inkluderende AI udvikling

Afslutningsvis vil vi diskutere de etiske dilemmaer og løsninger, som vores undersøgelse har identificeret i bestræbelserne på at gøre sprogmodeller mere retfærdige og inkluderende.

Efter at have belyst de aktuelle udfordringer ved generativ AI's indvirkning på global ulighed, vil vi i dette afsnit undersøge, hvilke normative og praktiske tilgange der kan bidrage til en mere ligeværdig teknologisk udvikling.

Som defineret af Helm et al. (2024) kan bias i sprogmodeller identificeres, når teknologien fortolker eller bearbejder indhold mindre korrekt på nogle sprog sammenlignet med andre. Dette kan tvinge talere af underrepræsenterede sprog til at forsimple eller tilpasse deres kommunikation, selvfremstilling og udtryksformer, når de engagerer sig med teknologien, for at matche det mere privilegerede sprogs normer. Vores udvalgte case, Masakhane, fremhæves af både respondenter og i litteraturen som et godt eksempel på, hvordan arbejdet med generativ AI netop kan gøres mere repræsentativt og inkluderende. Organisationen inddrager modersmålstalere i udviklingen af sprogmodeller for afrikanske underrepræsenterede sprog, hvilket bidrager til mere nøjagtige og kulturelt passende outputs. Som Claus (2024) dog påpeger, kan selv en græsrodsorganisation som Masakhane, risikere at reproducere eksisterende uligheder ved primært at fokusere på større afrikanske sprog, mens mindre sprog forbliver ekskluderet. Vi identificerer dette som et cirkulært problem: Hvis der ikke findes ressourcer eller aktører til at udvikle sprogmodeller for specifikke minoritetssprog, vil disse sprog fortsat være ekskluderet fra teknologisk repræsentation, hvilket netop fastholder deres marginalisering. Vi vil dog understrege, i tråd med Crawford (2021), at fokus ikke bør være at tilpasse sprogmodeller, når bias først opstår, men snarere at rette opmærksomheden mod de grundlæggende strukturelle forhold, der skaber disse uligheder i første omgang.

Et andet centralt aspekt, vi ønsker at fremhæve, er ulighed i teknologi som et demokratisk dilemma. I AU's AI-strategi og i den eksisterende litteratur fremhæves adgang til teknologi ofte som et centralt spørgsmål. Som Seger (2023) påpeger, må en demokratisering af AI ikke reduceres til spørgsmål om teknisk adgang eller åbenhed, men også om øget teknologisk infrastruktur, om deltagelse og indflydelse i hele AI-udviklingsprocessen - fra udvikling, til ejerskab, til anvendelse.

Som både litteraturen og vores respondenter har fremhævet, rejser spørgsmålet om den mest hensigtsmæssige vej mod en mere retfærdig og inkluderende AI-udvikling en række dilemmaer.

På den ene side kan man argumentere for, at lokalt tilpassede løsninger, som Masakhane, udgør en etisk og bæredygtig tilgang, fordi de inddrager modersmålstalere og sikrer kulturel og sproglig repræsentation.

På den anden side risikerer sådanne initiativer at stå isoleret og blive overskygget af store

teknologivirksomheder, der sætter de globale standarder.

Det er vigtigt at understrege, at den ulighed, vi har beskrevet i AI-udviklingen ikke er opstået ud af ingenting, men den udspringer af en allerede eksisterende global ulighed. Mange lande i majoritetsverdenen har af flere årsager, ikke haft mulighed for at tage del i den tidlige teknologiske udvikling, og det er derfor urealistisk at forvente at der er helt lige globale vilkår i udviklingen af AI.

Samtidig bør det ikke antages, at alle samfund nødvendigvis ønsker at deltage i den igangværende udvikling af kunstig intelligens. Det er vigtigt at udvise opmærksomhed på, at idéen om teknologisk fremskridt som et universelt mål kan afspejle en vestligt forankret forståelse, som ikke nødvendigvis deles globalt. En sådan antagelse risikerer at overskygge lokale behov, værdier og prioriteter. Dette rejser væsentlige spørgsmål om kontekstuel relevans og differentierede udgangspunkter: Ikke alle samfund har de samme behov, livssyn, ressourcer eller socioøkonomiske forudsætninger for - eller lyst til - at engagere sig i teknologisk innovation på lige vilkår. Derfor bør AI-initiativer ikke anskues som universelle løsninger, men vurderes i forhold til lokale kontekster, hvor 'one size fits all'-tilgange kan være direkte uhensigtsmæssige.

Det er dog også vigtigt at anerkende, at generativ AI også rummer potentiale til at mindske global ulighed. For eksempel kan sprogmodeller, der er designet med et eksplicit fokus på at undgå bias, bidrage til mere retfærdig behandling, end mennesker er i stand til. Det har dog ikke været formålet med denne undersøgelse at afdække teknologiens lighedsskabende potentiale, men snarere at undersøge, hvordan generativ AI, i sin nuværende form, risikerer at forstærke eksisterende globale uligheder. Vores ærinde er ikke at argumentere for et urealistisk ideal om lighed i distributionen af teknologisk magt, men at pege på, hvordan AI-teknologier, som i vid udstrækning formes og distribueres af aktører med teknologisk, økonomisk og geopolitisk forspring, risikerer at reproducere og forværre de strukturelle uligheder, der allerede eksisterer.

9. KONKLUSION

9. Konklusion

Der er ingen tvivl om, at generativ AI har udviklet sig i rekordfart og skyller ind over verden som en bølge, og som reguleringsmæssige tiltag har haft svært ved at gardere sig imod.
En konsekvens af dette, er en række etiske dilemmaer, som blandt andet relaterer sig til påvirkningen af global ulighed. Derfor har vi i dette speciale undersøgt:
Hvordan bidrager udviklingen og træningen af generativ AI til global ulighed, og hvilke etiske dilemmaer opstår i bestræbelserne på at gøre AI mere retfærdig og inkluderende?

Udviklingen af generativ AI rummer et potentiale for store teknologiske fremskridt og innovative forbedringer, men rummer på samme tid også risikoen for at forstærke globale socioøkonomiske uligheder.
Ved at zoome ind på sprogmodeller, kan vi konkludere, at den teknologiske udvikling af AI risikerer at forstærke eksisterende uligheder både globalt og lokalt, hvis der ikke aktivt arbejdes for at inkludere flere underrepræsenterede perspektiver og sprog. Ellers vil teknologien primært komme de mest ressourcestærke grupper til gode og efterlade mindre privilegerede samfund uden for udviklingen.
Bias, inklusion og etik i AI er ikke blot tekniske udfordringer, men tegn på en dybereliggende ulighed i, hvem teknologien er designet til at tjene. Som Frej Klem Thomsen, seniorkonsulent ved Nationalt Center for Etik påpeger, kan det synes neutralt, at en engelsk trænet model bedst hjælper engelsktalende brugere. Men når disse fordele systematisk tilfalder én gruppe, skaber det en ulige fordeling af økonomiske og sociale gevinster, og dermed et etisk dilemma.

Ud fra et normativt rammeværk, baseret på et globalt retfærdighedsprincip, har vi vist at den nuværende udvikling af AI skaber ulige vilkår, både hvad angår ulighed i deltagelse, kontrol og ejerskab. Det gælder især for sprog i mindre velstillede lande, som i forvejen er dårligt digitalt repræsenteret og som sjældent inkluderes i teknologiens udformning.
Både danske og sydafrikanske eksperter har i vores undersøgelse belyst denne problematik.

Danske eksperter fremhæver, at EU kæmper for at finde sin plads i et globalt kapløb om AI, hvor USA primært sætter tempoet. Mens USA overgiver innovationen til markedet og lader udviklingen løbe foran både etik og lovgivning, har EU vedtaget omfattende regulering, herunder den nye AI-forordning, der primært fokuserer på ansvarlig implementering, og ikke på selve træningen af modellerne.
Sydafrikanske eksperter lægger vægt på inklusion, kapacitetsopbygning og teknologisk

selvbestemmelse – som også er prioriteter formuleret i AU's AI-strategi. Strategien markerer, at afrikanske lande selv skal have adgang til teknologier, ressourcer og politiske værktøjer, der gør dem i stand til at sætte rammerne for AI på egne præmisser. Et opgør med den rolle, som Afrika historisk har haft i globale teknologiske udviklinger, som passiv modtager frem for aktiv medskaber.
Forskellene mellem respondenternes svar og henholdsvis EU's AI-forordning og AU's AI-strategi afspejler både forskellige politiske prioriteter og syn på teknologiens rolle i samfundet. EU lægger vægt på regulering og kontrol, mens AU i højere grad fokuserer på, hvordan AI-udviklingen kan komme kontinentet til gode, frem for at regulere. Det vidner om ikke blot divergerende strategiske mål, men også om dybt rodfæstede historiske uligheder og magtstrukturer.

Lokale initiativer som Masakhane viser, at AI-udviklingen i stigende grad kan komme til at tage udgangspunkt i egne sprog og datasæt frem for vestlige standarder. Men selvom der kommer flere lokale alternativer, fører det ikke nødvendigvis til øget global lighed. Uden at minoritetsverden løfter et globalt ansvar for etisk teknologiudvikling, herunder at styrke teknologiske kapabiliteter og etablere partnerskaber, risikerer privilegerede vestlige lande med adgang til kapital, teknisk ekspertise og infrastruktur at reproducere de samme ekskluderende mønstre, hvor kløften mellem majoritet- og minoritetsverden forstørres.
I en postkolonial kontekst er gamle magtstrukturer under opløsning, men de former stadig, hvem der bestemmer AI's fremtid, og hvem teknologien reelt kommer tilgode.

Vi kan konkludere, at generativ AI risikerer at forstærke eksisterende uligheder, både globalt og lokalt, medmindre teknologien målrettet udvikles mere inkluderende og reguleres, for at sikre at AI bruges ansvarligt. En etisk tilgang til generativ AI fordrer et paradigmeskifte i måden, vi forstår magt, ansvar og teknologiudvikling på.
Det er endvidere en kompleks opgave at prioritere mellem forskellige samfundsmæssige behov. I nogle sammenhænge vil grundlæggende infrastruktur som sundhedsvæsen og uddannelsessystem have højere prioritet end investeringer i digital udvikling og teknologisk kapacitet. Dette illustrerer, at lighed i AI er en kompleks udfordring, fordi den skal navigere i dybt forankrede, globale uligheder.
Generativ AI kan kun blive en drivkraft for reel lighed hvis vi samtidig genovervejer, hvem der har magt til at definere teknologien, herunder hvilke sprog og vidensformer der tæller, og balancerer investeringer i digital infrastruktur med alle samfunds grundlæggende behov. At give mindre sprogområder på egne modersmål reel medbestemmelse, er derfor ikke blot et spørgsmål om teknisk tilpasning, men om at anerkende sprogets centrale rolle i at forme,

hvem der kan tage del i AI's positive gevinster, og hvem, der fortsat risikere at blive forbigået.

10. **PERSPEKTIVERING**

10. Perspektivering

I det følgende afsnit præsenterer vi de perspektiver, som vores analyser har synliggjort. Her anerkender vi en række øvrige dimensioner af problemfeltet, som ligger uden for specialets primære fokus, men som vil være oplagte at udforske nærmere.

Vi har koncentreret os om store sprogmodeller og belyst, hvordan sproglig og kulturel repræsentation i generativ AI-teknologi bidrager til ulighed, som blot udgør et lille hjørne af forskningsfeltet.

AI og global retfærdighed er et komplekst og hastigt voksende forskningsfelt, som rummer mange nuancer og uforudsigelige fremtidsscenarier.

Spørgsmålet om, hvordan vi håndterer globale retfærdighedsdilemmaer i AI-udviklingen, er pt. åbent og genstand for diskussioner blandt forskere, politikere og civilsamfund. Generativ AI og store sprogmodeller har først for alvor accelereret de seneste tre år, med OpenAI's lancering af ChatGPT i november 2022 som et markant vendepunkt, der har sat skub i både tekniske fremskridt og etiske overvejelser. Vores arbejde er udført i en fase, hvor internationale reguleringer og standarder stadig er fragmenterede og under udvikling.

Vi mener, at de temaer og dilemmaer, specialet identificerer, er så nye, at de kræver målrettet forskning og uddybende undersøgelser.

Der er behov for mere forskning i, hvordan ulighed manifesterer sig i forskellige dele af AI's værdikæde. Blandt andet rejser brugen af det nyere fænomen 'predictive AI' etiske bekymringer, idet teknologien ofte reproducerer eksisterende uligheder ved at trække på data, som kan fastholde marginaliserede grupper i allerede ulige positioner, f.eks. ved at forudsige højere risiko for kriminalitet i bestemte geografiske områder (Rotaru et al., 2022).

Dernæst er der voksende opmærksomhed på 'data-labelling', som ofte er et usynligt led i udviklingen af AI-systemer, præget af lavtlønnede arbejdere, der udfører mentalt belastende opgaver under vanskelige forhold.

En anden konsekvens for øget global ulighed af AI, er udlicitering af jobs. Visse jobfunktioner automatiseres og forsvinder, særligt i lande med svagere arbejdsmarkeder og større økonomisk sårbarhed. Hvor tidligere investeringer under industrialiseringen ofte efterlod fysisk infrastruktur, som lokale samfund kunne videreføre og udvikle, efterlader den 'digitale industrialisering' sjældent noget varigt. Når sprogmodeller først er færdigtrænet, og behovet for lokal arbejdskraft ophører, står de involverede samfund tilbage uden arbejdspladser, teknologisk kapacitet eller langsigtet værdi. Denne praksis er ikke blot et

udtryk for økonomisk ulighed, men en videreførelse af koloniale forhold i ny, digital form – hvor ressourcer stadig udvindes, men nu i form af data og arbejdskraft, uden reel investering i lokal udvikling.
Derudover rejser open source-strategier og den såkaldte 'demokratisering' af AI en række dilemmaer. Når tech-virksomheder stiller AI-modeller frit til rådighed på et stadig meget ureguleret marked, muliggør det anvendelser uden etiske rammer eller ansvarlighed, hvilket kan skabe nye former for magtbalancer i stedet for at nedbryde eksisterende.
Specialet lægger op til en række undersøgelser af, hvilke fremtidige løsninger der bedst adresserer de skitserede problemstillinger. Lokale og alternative initiativer som Masakhane og Lelapa AI, har selv en risiko for at reproducere ulighed, eksempelvis ved primært at fokusere på de mest udbredte afrikanske sprog og dermed undlade mindre repræsenterede sprog. Samtidig rejser manglen på data en grundlæggende udfordring for mange afrikanske sprog. Det nødvendige datagrundlag kan være snævert eller helt fraværende, hverken som nedskrevne tekster eller digitaliseret materiale, hvilket betyder, at væsentlige kulturelle, sproglige og kontekstuelle elementer risikerer at gå tabt i udviklingen af AI-systemer for underrepræsenterede sprogområder.
Der er derfor behov for øget opmærksomhed på, hvordan man kan styrke udviklingen af pluralistiske, sprog- og kulturspecifikke AI-systemer, der både leverer mere præcise resultater og fremmer etisk og respektfuldt samarbejde med lokale aktører. Det kunne eksempelvis komme fra en multilateral institution som FN, som kunne være med til at skabe en forenende kollektiv stor sprogmodel, som en måde at tage styring væk fra kommercielle interesser og sikre governance med fokus på inklusion, etik og retfærdighed. Det er dog væsentligt at understrege, at ethvert pluralistisk system først og fremmest må tage udgangspunkt i de samfund, det skal tjene. For at være meningsfuldt og legitimt må det således kunne adressere og opfylde de konkrete behov, som eksempelvis det sydafrikanske samfund står overfor.
Fremtidig forskning bør undersøge, hvordan marginaliserede aktører kan opnå reel indflydelse på AI-udviklingen, samt hvilke politiske, regulatoriske og teknologiske mekanismer der kan understøtte mere retfærdige og inkluderende AI-systemer, som kan sikre teknologiske fremskridt samtidig med man afhjælper socioøkonomisk og kulturel global ulighed.

Da specialet har undersøgt aspekter af global ulighed i relation til AI, er det oplagt at rette opmærksomheden mod, hvordan teknologisk udvikling i højere grad kan bidrage til global lighed, sådan som vi også indledningsvis rejser spørgsmålet om: Hvad skal der til, for at AI bliver en del af løsningen på globale udfordringer frem for at forstærke dem?

Problemstillingen rækker ud over AI og peger på mere grundlæggende spørgsmål om teknologisk udvikling, magtforhold og global fordeling af velstand.

11. LITTERATUR

11. Litteraturliste

Primær litteratur

Adams, R. (2025). *The New Empire of AI: The Future of Global Inequality*. Polity Press.

Alayande, A., & Adams, R. (2024). *Africa Now Has a Continental AI Strategy: What Next?* Global Center on AI Governance.

Azaroual, F. (2024). *Policy Brief: Artificial Intelligence in Africa—Challenges and Opportunities*. Policy Center for the New South. https://www.policycenter.ma/sites/default/files/2024-09/PB_23_24%20%28Azeroual%29%20%28EN%29.pdf

Bahar, D., & Michael, A. (2024). *A New Industrial Revolution: Will AI Widen or Close the Income Gap between Rich and Poor Countries?* Center for Global Development. https://www.cgdev.org/publication/new-industrial-revolution-will-ai-widen-or-close-income-gap-between-rich-and-poor-countries

Bender, E. M. (2024). "Resisting Dehumanization in the Age of 'AI'." *Psychological Science*. https://faculty.washington.edu/ebender/resistingdehumanization

Bender, E. M., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Cowls, J., Vayena, E. (2018). *AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations.* Minds and Machines.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Brinkmann, S. (2013). *Qualitative Interviewing*. Oxford University Press.

Bryman, A. (2008). *Social Research Methods* (3rd ed.). Oxford University Press.

Cazzaniga, A., et al. (2024). “Gen-AI: Artificial Intelligence and the Future of Work.” *IMF Staff Discussion Note* SDN2024/001. International Monetary Fund. https://doi.org/10.5089/9798400262548.006

Claus, H. (2024). “Now You Are Speaking My Language: Why Minoritised LLMs Matter.” The Ada Lovelace Institute. https://www.adalovelaceinstitute.org

Couldry, N., & Mejias, U. A. (2019). “Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject.” Television & New Media.

Couldry, N., & Mejias, U. A. (2019). “Making Data Colonialism Liveable: How Might Data’s Social Order Be Regulated?” Internet Policy Review.

Couldry, N., & Mejias, U. A. (2023). “The Decolonial Turn in Data and Technology Research: What Is at Stake and Where Is It Heading?” *Information, Communication & Society, 26*(4), 1–18.

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Cronin, P., Ryan, F., & Coughlan, M. (2008). “Undertaking a Literature Review: A Step-by-Step Approach.” British Journal of Nursing. https://doi.org/10.12968/bjon.2008.17.1.28059

Eke, D. O., Wakunuma, K., & Akintoye, S. (Eds.). (2023). *Responsible AI in Africa: Challenges and Opportunities*. Springer. https://doi.org/10.1007/978-3-031-08215-3

Fairclough, N. (2003). *Analysing Discourse: Textual Analysis for Social Research*. Routledge.

Fernhout, F., & Duquin, T. (2024). “The EU Artificial Intelligence Act: Our 16 Key Takeaways.” Stibbe.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). *AI4People - An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations*. Minds and Machines.

Floridi, L., & Cowls, J. (2019). “A Unified Framework of Five Principles for AI in Society.” The MIT Press. https://doi.org/10.1162/99608f92.8cd550d1

Freeman, R. (2015). “Stakeholder Management: Framework and Philosophy.” In *Strategic Management*. Cambridge University Press. https://doi.org/10.1017/CBO9781139192675.006

Hämäläinen, M. (2021). “Endangered Languages Are Not Low-Resourced.” https://doi.org/10.31885/9789515150257.1

Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2024). “Diversity and Language Technology: How Language Modeling Bias Causes Epistemic Injustice.” Ethics and Information Technology. https://doi.org/10.1007/s10676-023-09742-6

International Telecommunication Union. (2024). *AI for Good Impact Report* (ISBN 978-92-61-39741-8). https://s41721.pcdn.co/wp-content/uploads/2021/06/eBat-2402309_AI-for-Good-Impact-Report-E-v6.pdf

Kvale, S., & Brinkmann, S. (2015). *Interview: Det Kvalitative Forskningsinterview som Håndværk* (3. udg.). Hans Reitzels Forlag.

Lignos, C., Holley, N., Palen-Michel, C., & Sälevä, J. (2022). *Findings of the Association for Computational Linguistics*. Dublin, Ireland: Association for Computational Linguistics.

Nielsen, P., & Buch-Hansen, H. (2012). “Kritisk Realisme.” I S. Juul & K. B. Pedersen (Red.), *Samfundsvidenskabernes Videnskabsteori: En Indføring*. Hans Reitzels Forlag.

Nussbaum, M. C. (2003). “Capabilities as Fundamental Entitlements: Sen and Social Justice.” Feminist Economics. https://doi.org/10.1080/1354570022000077926

Prescott, A. (2016). “Avoiding the rear view mirror” *Digital Transformations: Talks and Presentations*. https://medium.com/@ajprescott/avoiding-the-rear-view-mirror-96cd7cfc794c

Rotaru, V., Huang, Y., Li, T. et al. (2022). “Event-level prediction of urban crime reveals a signature of enforcement bias in US cities”. Nat Hum Behav 6. https://doi.org/10.1038/s41562-022-01372-0

Rowley, J., & Slack, F. (2004). “Conducting a Literature Review.” Management Research News. https://doi.org/10.1108/01409170410784185

Schwartz, L. (2022). “Primum Non Nocere: Before Working with Indigenous Data, the ACL Must Confront Ongoing Colonialism.” In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics.* https://aclanthology.org/2022.acl-short.104

Schellekens, P., & Skilling, D. (2024, 17. oktober). “Three Reasons Why AI May Widen Global Inequality.” Center for Global Development. https://www.cgdev.org/blog/three-reasons-why-ai-may-widen-global-inequality

Seger, E. (2023). *Democratising AI: Multiple Meanings, Goals, and Methods*. Centre for the Governance of AI.

Sen, A. (1993). “Capability and Well-Being.” I M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (online ed., publ. 1. nov. 2003). Oxford University Press.

Sen, A. (2006). “What Do We Want from a Theory of Justice?” *The Journal of Philosophy, 103*(5), 215–238. https://doi.org/10.5840/jphil2006103517

Sen, A. (2015). “The Idea of Justice: A Response.” *Philosophy & Social Criticism*, 77–88.

Singh, A. (2008). “Natural Language Processing for Less Privileged Languages: Where Do We Come From? Where Are We Going?” I *Proceedings of the IJCNLP-08 Workshop on NLP for Less Privileged Languages*. ACL Anthology.

Spivak, G. C. (1988). “Can the Subaltern Speak?” I C. Nelson & L. Grossberg (Eds.), *Marxism and the Interpretation of Culture* (s. 24–28). Macmillan.

Suleyman, M. (2023). *The Coming Wave.* Vintage Publishing.

Tsou, J. Y., & Walsh, K. P. (2023). “Ethical Theory and Technology.” I G. J. Robson & J. Y. Tsou (Eds.), *Technology Ethics: A Philosophical Introduction and Readings*. Routledge.

Wild, S. (2021). “African Languages to Get More Bespoke Scientific Terms.” Nature. DOI: https://doi.org/10.1038/d41586-021-02218-x https://www.nature.com/articles/d41586-021-02218-x

Wendler, C., Veselovsky, V., Monea, G., & West, R. (2024). “Do Llamas Work in English? On the Latent Language of Multilingual Transformers.” Cornell University / EPFL. https://arxiv.org/abs/2402.10588

Yin, R. K. (2014). *Case Study Research: Design and Methods* (5th ed.). Sage Publications.

Sekundær litteratur

African Union. (2024). *Continental Artificial Intelligence Strategy: Harnessing AI for Africa’s Development and Prosperity.*

AscendixTech. (2025). *How many AI companies are there in the world?* Hentet 29. maj 2025 fra https://ascendixtech.com/how-many-ai-companies-are-there/

CompaniesMarketCap. (2024). *Largest AI companies by market capitalization*. Hentet 29. maj 2025 fra https://companiesmarketcap.com/artificial-intelligence/largest-ai-companies-by-marketcap/

Den Danske Ordbog. *bias.* Hentet 15. april 2025 fra https://ordnet.dk/ddo/ordbog?query=bias

Den Europæiske Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council.* EUR-Lex. https://eur-lex.europa.eu/eli/reg/2024/1689

Europa-Parlamentet og Rådet. (2024). *Forordning (EU) 2024/1689 af 13. juni 2024 om harmoniserede regler for kunstig intelligens og om ændring af forordning (EF) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 og (EU) 2019/2144 samt direktiv 2014/90/EU, (EU) 2016/797 og (EU) 2020/1828 (forordningen om kunstig intelligens).* Den Europæiske Unions Tidende, L 2024/1689, 1–X. https://eur-lex.europa.eu/legal-content/DA/TXT/?uri=CELEX:32024R1689

European Commission. (2019, 8. april). *AI HLEG: Independent expert group set up by the European Commission.* https://ec.europa.eu/digital-strategy/our-policies/artificial-intelligence_en

European Union. (2025). *Shaping Europe's digital future: Regulatory framework for AI.* Digital Strategy. https://digital-strategy.ec.europa.eu/policies/regulatory-framework-ai

Holdsworth, J. (2023, 22. december). *What is AI bias?* IBM. https://www.ibm.com/topics/ai-bias

International Monetary Fund. (u. å.). *Introduction to inequality.* https://www.imf.org/en/Topics/Inequality/introduction-to-inequality

Masakhane Research Foundation. (2024). *5-year strategy (2025–2029).* https://www.masakhane.io/

Thomsen, F. K., & Skou Olsen, L. I. (2021). *Dataetisk metode & teori.* Institut for Menneskerettigheder for Dataetisk Råd.

University of Oxford. (2025, February 14). *Expert comment: Paris AI Summit misses opportunity for global AI governance*. https://www.ox.ac.uk/news/2025-02-14-expert-comment-paris-ai-summit-misses-opportunity-global-ai-governance

United Nations AI Advisory Body (2024) https://www.gp-digital.org/news/the-final-report-of-hlab-ai-our-analysis-and-thoughts/

UN System Chief Executives Board for Coordination. (u. å.). *Inequalities.* https://unsceb.org/topics/inequalities

W3Techs. (n.d.). *Usage statistics of content languages for websites.* Hentet 24. marts 2025 fra https://w3techs.com/technologies/overview/content_language