

NOTAT

# KI: BETYDNING FOR ARBEIDSSTYRKEN

En analyse av potensialet for kunstig intelligens-drevet effektivisering i norsk næringsliv



**MENON-PUBLIKASJON NR. 127/2023**

Av Trygve Svalheim, John Oskar Skjeldrum, Sebastian Winther-Larsen, Kristoffer Midttømme og Jonas Erraia  
Forsidefoto: Menon Economics/DALLE 3



## Forord

Alle snakker om KI og hvordan det vil påvirke samfunnet. I dette notatet benytter vi etablert forskningsmetodikk på feltet for å forsøke å kvantifisere én bit av påvirkningen KI vil kunne ha på det norske samfunnet. Konkret ser vi på potensialet for å effektivisere tid brukt på å løse *dagens* arbeidsoppgaver. Teknologi generelt, og kunstig intelligens spesielt, har potensial til å omforme hvordan vi jobber og lever, og vi kan ikke ennå forutsi hvordan en slik effektiviseringsgevinst tas ut. [John Maynard Keynes spådde for snart 100 år siden](#) at teknologisk vekst ville tas ut i økt fritid, men det har vi ennå ikke sett. I notatet forsøker vi derfor ikke å si noe om hvordan, men hvor mye.

Mange har forsøkt før oss, men det unike med vår fremgangsmåte er at vi kobler etablert internasjonal forskningsmetodikk med norske data for å komme med anslag som er mer relevante for norske yrker, næringer og fylker. Selv om anslagene er beheftet med stor usikkerhet, gir de nyttig innsikt for alle som forsøker å planlegge for en usikker fremtid, eller forme den.

Menon Economics er et forskningsbasert analyse- og rådgivningsselskap i skjæringspunktet mellom foretaksøkonomi, samfunnsøkonomi og næringspolitikk. Vi tilbyr analyse- og rådgivningstjenester til bedrifter, organisasjoner og offentlig sektor. Vårt hovedfokus ligger på empiriske analyser av økonomisk politikk, og våre medarbeidere har økonomisk kompetanse på et høyt vitenskapelig nivå.

---

Oktober 2023

Jonas Erraia  
Prosjektansvarlig  
Menon Economics

# Innhold

<b>SAMMENDRAG</b>	<b>3</b>
<b>INTRODUKSJON</b>	<b>4</b>
<b>KI VIL SPARE TID FOR DE FLESTE NORSKE ARBEIDSTAKERE</b>	<b>7</b>
Gjennomsnittsnordmannen kan spare 17 prosent av arbeidstiden med KI	7
Høyere effektiviseringspotensial i høyere lønnede grupper	8
Opp mot 50 prosent effektiviseringspotensial i de mest eksponerte yrkene	9
Høyest effektiviseringspotensial i tjenesteytende næringer	11
Jevnt fordelt effektiviseringspotensial i ulike fylker, antall sysselsatte viktig	12
I offentlig sektor er det særlig stort effektiviseringspotensial innen saksbehandling	13
Effektiviseringspotensialet er verdt opp mot 500 milliarder årlig	15
<b>METODIKK</b>	<b>17</b>
En kombinasjon av amerikanske og norske data	17
Små justeringer i modellen fra GPTs are GPTs	18
Vi kobler yrkesdata med næringsdata for å analysere variasjonen	19
Vi benytter en mer konservativ kategorisering	19
<b>VEDLEGG: RUBRIKK</b>	<b>21</b>
<b>REFERANSELISTE</b>	<b>24</b>

## Sammendrag

Dette notatet presenterer funn fra Menons analyse av effektene ved bruk av KI-verktøy i den norske arbeidsstyrken. For å vurdere effektiviseringspotensialet benytter vi en database med yrker og deres tilhørende arbeidsoppgaver. Først kategoriserer vi alle arbeidsoppgaver basert på i hvilken grad KI kan øke effektiviteten i utførelsen av arbeidsoppgaven. Metodikken er en videreutvikling av en tidligere studie gjennomført av OpenAI. Analysen gir et helhetlig bilde rettet mot beslutningstakere i politikk, næringsliv og akademia, og søker å belyse de økonomiske, sosiale og politiske konsekvensene av å integrere KI i det norske arbeidsmarkedet.

Vi bruker vår modell til å beregne effektiviseringspotensialet for ulike yrker, næringer (inkludert offentlig sektor) og fylker.

Våre hovedfunn indikerer at omtrent **70 prosent av den norske arbeidsstyrken kan redusere arbeidstiden med 10 prosent** som følge av effektivisering ved bruk av KI-verktøy. For nesten halvparten av arbeidsstyrken kan denne gevinsten være så høy som 20 prosent, noe som i teorien kan redusere den gjennomsnittlige arbeidsuken med en hel dag. Vi observerer også en klar korrelasjon mellom høyere lønnsgrupper og større effektiviseringspotensial, med særlig sterke effekter innen yrkesgrupper som ingeniører, matematikere og programmerere.

På *næringsnivå* viser resultatene at tjenesteytende næringsgrupper som jurister og regnskapsmessig tjenesteyting, finansiering og administrasjon kan oppnå en effektiviseringsgrad som tilsvarer tredjedel av dagens arbeidsoppgaver. Totalt sett har 90 prosent av næringene en effektivitetsgrad over 10 prosent ved bruk av KI, noe som viser til hvor bredt teknologien kan brukes over flere næringsgrupper med forskjellige ferdigheter og behov.

Det er ikke kun privat sektor som kan høste gevinster i form av effektivitet ved å implementere KI i sine arbeidsoppgaver. Ved fullstendig utnyttelse av KI slik som teknologien står i dag, kan offentlig sektor effektivisere arbeidsoppgaver som årlig vil tilsvare 155 000 årsverk i sektorer som helse, undervisning og offentlig administrasjon.

Geografisk sett er effektiviseringspotensialet høyest i Oslo, Viken og Trøndelag, men forskjellene mellom fylkene er minimale, noe som antyder en relativt jevn yrkesfordeling på nasjonalt nivå. Når det kommer til næringer, har tjenesteytende sektorer som juridisk og finansiell tjenesteyting betydelig høyere effektiviseringspotensial, mens yrker med høy grad av menneskelig interaksjon har mindre muligheter.

Effektivisering er lønnsomt, og full utnyttelse av KI i arbeidsstyrken på måten beregnet her, kan tilsvare en årlig verdiskaping i Norge på 500 milliarder kroner. Dette tallet representerer *kun* effektivisering av dagens arbeidsoppgaver, og inkluderer ikke gevinster fra ny KI-drevet innovasjon.

# Introduksjon

Siden lanseringen av ChatGPT i november 2022, har samfunnet blitt stadig mer oppmerksomme på det enorme potensialet til store språkmodeller (LLMs) og andre former for generativ kunstig intelligens (KI). Kun måneder etter lanseringen demonstrerte OpenAI store fremskritt med en ny versjon: GPT-4. Teknologien er under stadig utvikling, men tilgjengelige versjoner demonstrerer allerede enorme evner som på sikt vil ha stor betydning for mange deler av samfunnet vårt, inkludert arbeidsmarkedet. Tidligere forskning gjort av OpenAI selv, har pekt på store gevinster ved bruk av ChatGPT og lignende verktøy i det amerikanske arbeidsmarkedet.<sup>1</sup> I dette analysenotatet utforsker vi, med en videreutvikling av OpenAIs metodikk, konsekvensene av KI-verktøy på det norske arbeidsmarkedet.

Store språkmodeller som GPT-4 er resultatet av avansert maskinlæring, hvor milliarder av parametere er finjustert gjennom trening på store mengder tekst. Disse modellene er designet for å ta inn tekst og gjette på en ny mengde med tekst basert på en matematisk likhet til teksten som ble matet inn. Til tross for at disse verktøyene er designet for å generere tilfeldig tekst, har de vist en stor grad av fleksibilitet og potensiale. De viser seg å være i stand til en rekke oppgaver, inkludert oversettelse, tekstbehandling, idé-sparring, komplekse problemløsning og kodeutvikling.

Selv om språkmodeller har sine svakheter, spesielt når det gjelder generering av nøyaktig og faktabasert informasjon, er den nyeste versjonen, GPT-4, betydelig forbedret. Denne modellen har evnen til å programmere spill fra bunnen av med et minimum av menneskelig input, bestå den amerikanske advokateksamen med gode karakterer, og snakker hebraisk bedre enn forgjengeren snakket engelsk. Disse avanserte funksjonene understreker det brede spekteret av anvendelser hvor språkmodeller kan komme til å spille en avgjørende rolle i fremtiden.

Dette analysenotatet har som mål å bidra til en bedre forståelse av hvordan KI og språkmodeller kan forme det fremtidige norske arbeidsmarkedet. Vi ønsker også å gi en ramme for politiske beslutningstakere, næringslivsledere og akademikere for å vurdere de økonomiske, sosiale og politiske implikasjonene av disse transformativ teknologiene. For å forstå effekten av kunstig intelligens (KI) på arbeidsstyrken, er det nødvendig å forstå hvordan et yrke kan bli berørt. Vi tar for oss tre hovedscenarier:

## 1. Effektivisering

I dette scenarioet fører KI til en betydelig økning i produktivitet, ved at KI effektiviserer oppgaver som tidligere krevde menneskelig innsats. I en slik situasjon vil effektiviseringen frigjøre tid til å gjøre oppgaver KI ikke egner seg til, eller jobbe mindre. Samtidig har studier rapportert at arbeidstakere opplever større grad av tilfredshet, positiv læringseffekt, og at ferdighetsnivået i bedriften blir utjevnet når de kan benytte språkmodeller til å effektivisere arbeidet sitt<sup>2</sup>. På den andre siden vil yrker der KI effektiviserer uten at etterspørselen etter arbeidet øker, oppleve en negativ effekt på lønnsnivået eller føre til oppsigelser. Politiske tiltak som omfordeler deler av produktivetsgevinsten tilbake til arbeidstakeren kan ses på som hybridalternativ som minker de negative effektene. KI-revolusjonen kan

---

<sup>1</sup> Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *Gpts are gpts: An early look at the labor market impact potential of large language models*. arXiv preprint arXiv:2303.10130.

<sup>2</sup> Erik Brynjolfsson, Danielle Li, Lindsey R. Raymond, *Generative AI at Work*. NBER Working Paper No. 31161, April 2023. DOI: [10.3386/w31161](https://doi.org/10.3386/w31161)



derfor åpne dørene for en større omstrukturering av arbeidsuken, på lik linje med den industrielle revolusjonen eller arbeiderbevegelsen.

## 2. Forsterkning

Effekten av KI-verktøyene trenger ikke bare ha en tidsbesparende effekt på allerede eksisterende arbeidsoppgaver i et yrke. Mindre kvalifiserte arbeidstakere kan også gjennomføre nye arbeidsoppgaver, som før krevde høy utdanning eller kompetanse. Dette har skjedd mange ganger gjennom historien, et tydelig eksempel er hvordan bøker ble kopiert for hånd av dyktige skrivere før trykkpressen gjorde det mulig for arbeidere med andre kvalifikasjoner å gjennomføre arbeidet. Et eksempel på hvordan denne effekten kan utarte seg i fremtiden er at mindre kvalifiserte helse- og omsorgsarbeidere kan ta over deler av diagnostiseringsarbeidet til legene. Den viktigste forskjellen mellom effektivisering og forsterkning er derfor hvordan teknologien kan styrke mindre kvalifiserte arbeidstakere.

## 3. Automatisering

I dette scenarioet fører bruken av KI til en betydelig erstatning av menneskelig arbeidskraft. Hvis de økonomiske konsekvensene blir store nok, vil dette scenarioet kreve en fremtidig balanserende politisk respons. Men fullstendig effektivisering av alle arbeidsoppgaver innenfor et yrke er ikke den eneste veien til automatisering. Et mer naturlig scenario er at automatiseringen skjer på tvers av stillinger, ved at spesifikke arbeidsoppgaver automatiseres, og én arbeidstaker får frigjort nok tid til å ta over ansvarsområdene til en annen. Det finnes allerede eksempler på at selskaper benytter KI-omstillingen til å gjøre drastiske endringer i selskapsstrukturen. Et konkret eksempel er nettavisen Gizmodo, som nylig permitterte sin spanske avdeling til fordel for KI-drevne oversettelser<sup>3</sup>.

## 4. Liten eller ingen effekt

I dette scenarioet har KI ingen merkbar effekt. Dette gjelder spesielt yrker dominert av fysisk arbeid. Utover disse er dette scenarioet mindre sannsynlig sett i lys av KIs hurtige utbredelse, men det kan heller ikke utelukkes. Faktorer som mulig stagnasjon i den teknologiske utviklingen, strenge regulatoriske tiltak eller en mer usannsynlig ideologisk retrett fra markedsledende selskaper som OpenAI, kan begrense KI-utbredelsen.

Etter seks måneder med tilgang til avanserte språkmodeller har arbeidsledighetsraten globalt sett forblitt stabil, noe en nylig OECD-studie bekrefter.<sup>4</sup> Studien finner ingen signifikant nedgang i etterspørselen etter menneskelig arbeidskraft. En sannsynlig årsak er at vi fortsatt er i en tidlig fase av teknologisk adopsjon, hvor det tar tid å identifisere de mest effektive bruksområdene. I tillegg indikerer studien at bedrifter kan være tilbakeholdne med å permittere og avskjedige ansatte, og heller lar arbeidsstyrken reduseres naturlig gjennom pensjonering og frivillige oppsigelser. Denne observerte stabiliteten i arbeidsmarkedet samsvarer godt med Everett Rogers' teori om «*Diffusion of Innovations*». Denne teorien forklarer at ny teknologi adopteres i etapper, med en liten gruppe innovatører som går foran. Resten av markedet venter ofte på å se deres suksess før de følger etter. Følgelig kan vi forvente at de mest markante endringene vil komme til syne etter hvert som teknologien modner og blir mer allment akseptert.

---

<sup>3</sup> [Gizmodo's owner replaced its Spanish language journalists with AI - The Verge](#)

<sup>4</sup> OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, <https://doi.org/10.1787/08785bba-en>.

Samtidig indikerer flere og flere studier at store språkmodeller fører til positiv gevinst i arbeidsmarkedet. For eksempel viste en studie av Erik Brynjolfsson med flere<sup>5</sup> at tilgang til en KI-basert samtaleassistent økte produktiviteten med 14 prosent i gjennomsnitt blant kundeservicearbeidere. Nybegynnere og lavt kvalifiserte arbeidere så en forbedring på 35 prosent i produktiviteten, noe som underbygger argumentet for at KI kan være en katalysator for økt effektivitet, særlig blant mindre erfarne arbeidere. En studie av Fabrizio Dell'Acqua med flere, utført med Boston Consulting Group støtter dette med å vise at konsulenter fullførte 12 prosent flere arbeidsoppgaver, 25 prosent raskere og med 40 prosent bedre kvalitet sammenlignet med en kontrollgruppe innenfor en gitt tidsramme.<sup>6</sup> Det var igjen de lavest kvalifiserte arbeiderne som så den høyeste forbedring i produktivitet og kvalitet på 43 prosent. Studien pekte også på viktigheten av å kjenne kapasitetene til KI. I oppgaver hvor KI har begrensede evner, hadde konsulenter 19 prosent lavere sannsynlighet for å utføre oppgaver riktig med hjelp fra KI sammenlignet med dem uten hjelp.

En annen studie av Shakked Noy og Whitney Zhang<sup>7</sup> fant at assisterende chatboter som ChatGPT kunne øke produktiviteten i skriveoppgaver betydelig, redusere tidsbruken med 40 prosent og forbedre output-kvaliteten med 18 prosent. Dette demonstrerer at selv i komplekse, kreative oppgaver kan KI ha en markant positiv effekt.

Motivasjonen for vår analyse ligger ikke bare i de teknologiske fremskrittene disse modellene har gjort, men også i de brede, skalerbare og praktiske anvendelsene vi ser i teknologier utviklet rundt dem. I Norge, som i resten av verden, blir bruken av KI og språkmodeller stadig mer utbredt, og det er derfor avgjørende å forstå hvordan disse teknologiene kan påvirke arbeidsstyrken og økonomien som helhet.

Vår studie følger den omtalte *GPTs are GPTs*-studiens metode, med noen tilpasninger. Studien benytter en kombinasjon av menneskelig ekspertise og GPT-4-klassifiseringer til å vurdere KIs effekt på hver enkelt arbeidsoppgave. Til forskjell fra OpenAI's studie, tar vår metode en mer konservativ tilnærming til klassifisering av arbeidsoppgaver for å gi et mer realistisk og robust bilde av teknologiens potensielle innflytelse, og vi har gjort noen justeringer i rubrikken som brukes for klassifisering.

Implementeringen av KI kan skje både gradvis og plutselig, avhengig av teknologisk modenhet og markedsdynamikk. Vår målsetting er å gi en helhetlig og nyansert forståelse av hvordan KI kan forme det norske arbeidsmarkedet, med fokus på de økonomiske, sosiale og politiske utfordringene, og mulighetene, som ligger foran oss.

---

<sup>5</sup> Erik Brynjolfsson, Danielle Li, Lindsey R. Raymond, *Generative AI at Work*. NBER Working Paper No. 31161, April 2023. [DOI: 10.3386/w31161](https://doi.org/10.3386/w31161)

<sup>6</sup> Dell'Acqua, F. et al. (2023) 'Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality', *SSRN Electronic Journal [Preprint]*. doi:10.2139/ssrn.4573321.

<sup>7</sup>Shakked Noy Whitney Zhang, *Experimental evidence on the productivity effects of generative artificial intelligence*. *Science*381, 187-192(2023). [DOI:10.1126/science.adh2586](https://doi.org/10.1126/science.adh2586)

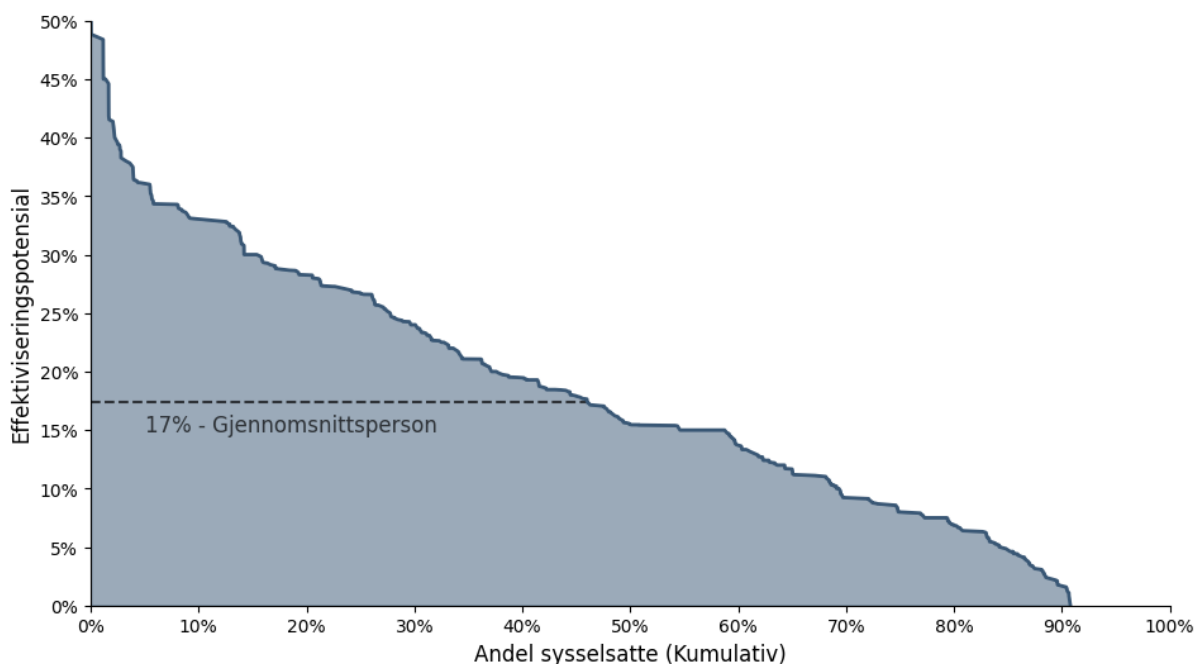
## KI vil spare tid for de fleste norske arbeidstakere

I dette kapittelet vil vi utforske og diskutere de kvantitative funnene fra vår analyse av effektiviseringspotensialet i det norske arbeidsmarkedet. Vi tar sikte på å gi en detaljert oversikt over potensialet for automatisering og effektivisering gjennom kunstig intelligens for ulike yrker og lønnsgrupper. Tallet vi presenterer for effektiviseringspotensialet i prosent er et estimat på total tidsbesparelse, basert på hvor mange arbeidsoppgaver som kan effektiviseres ved bruk av KI.

### Gjennomsnittsnordmannen kan spare 17 prosent av arbeidstiden med KI

Utslaget av effektiviseringen på det norske arbeidsmarkedet bestemmes av flere faktorer. For eksempel vil resultatet av KI-revolusjonen avgjøres av hvor mange som er berørt, og hvilke yrkesgrupper de tilhører. Dersom det er få ansatte i yrkene som ser mest effektiviseringspotensial, vil effekten på landsbasis være liten. På samme måte vil lønnsfordelingen i de berørte yrkene påvirke størrelsen på de samfunnmessige konsekvensene. For å danne et fullstendig bilde av KIs effekt på arbeidsstokken benytter vi derfor sysselsettingstall per yrkesnivå for å finne utslag per sysselsatt, og illustrerer resultatet under.

Figur 2: Effektiviseringspotensial mot kumulativ andel sysselsatte i Norge. Figuren viser hvordan effektiviseringspotensialet fordeler seg over Norges arbeidsstyrke. Effektiviseringspotensialet for en gjennomsnittlig norsk sysselsatt er markert med stiplet linje. Kilde: SSB og Menon Economics.



Figuren viser en kumulativ graf som illustrerer potensialet for effektivisering av Norges arbeidsstyrke. Den avslører at cirka 70 prosent av den sysselsatte befolkningen kan forvente en effektiviseringsgevinst på 10 prosent, mens 40 prosent kan se for seg en gevinst på 20 prosent. Videre vil omtrent 1 prosent, som tilsvarer rundt 25 tusen personer, oppnå en effektivisering på 40 prosent eller høyere.

Dette tyder på at en stor del av arbeidsstyrken står overfor mulige automatiserings- og effektiviseringseffekter. Vi har også fremhevet det gjennomsnittlige effektiviseringspotensialet for en typisk norsk arbeidstaker, som



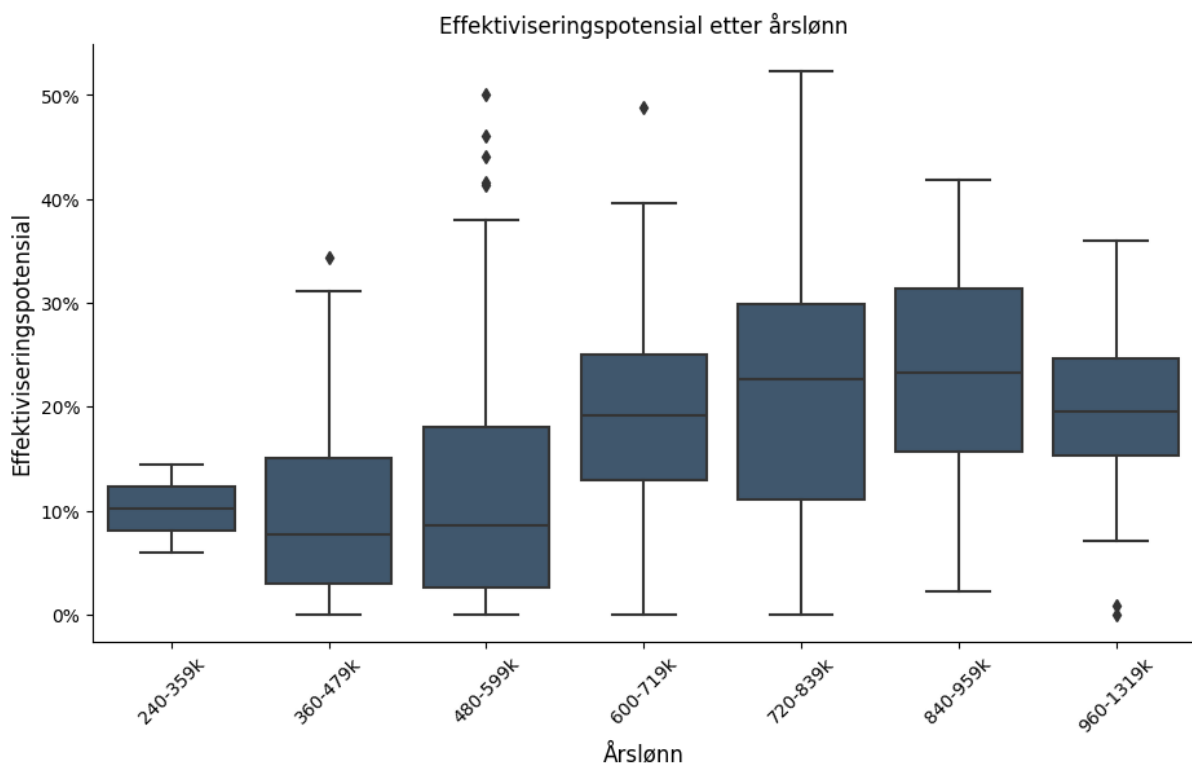
ligger på 17 prosent (beregnet som gjennomsnitt av effektiviseringspotensial per yrke vektet mot antall sysselsatte). Dette betyr at dersom den økonomiske gevinsten tillater det, kan den gjennomsnittlige nordmann redusere arbeidsuken med nesten en hel arbeidsdag.

### Høyere effektiviseringspotensial i høyere lønnede grupper

Historisk har teknologi primært favorisert høyere utdannet arbeidskraft. Dette har skjedd fordi store deler av teknologien har automatisert manuelle oppgaver, mens den har økt etterspørselen etter høykompetent arbeidskraft. Dette har bidratt til at lønnsforskjellen mellom lavt og høyt utdannet arbeidskraft har økt i de fleste vestlige økonomier. Relativt til historiske perioder med automatisering, skiller LLMs seg ved å kunne utføre og automatisere ikke-manuelle oppgaver. Alt annet likt, tilsier det at effektiviseringspotensialet vil være høyest blant analytiske yrker som ofte er bekledd av høyt utdannet arbeidskraft, mens manuelle yrker, og da særlig de som innebærer en vesentlig grad av menneskelig kontakt, ikke vil bli påvirket i like høy grad.

For å teste dette har vi beregnet det gjennomsnittlige effektiviseringspotensialet for ulike lønnsnivåer. Dette er vist i figuren under.

Figur 3: Effektiviseringspotensial inndelt i grupper etter lønn. Kilde: Menon Economics, SSB



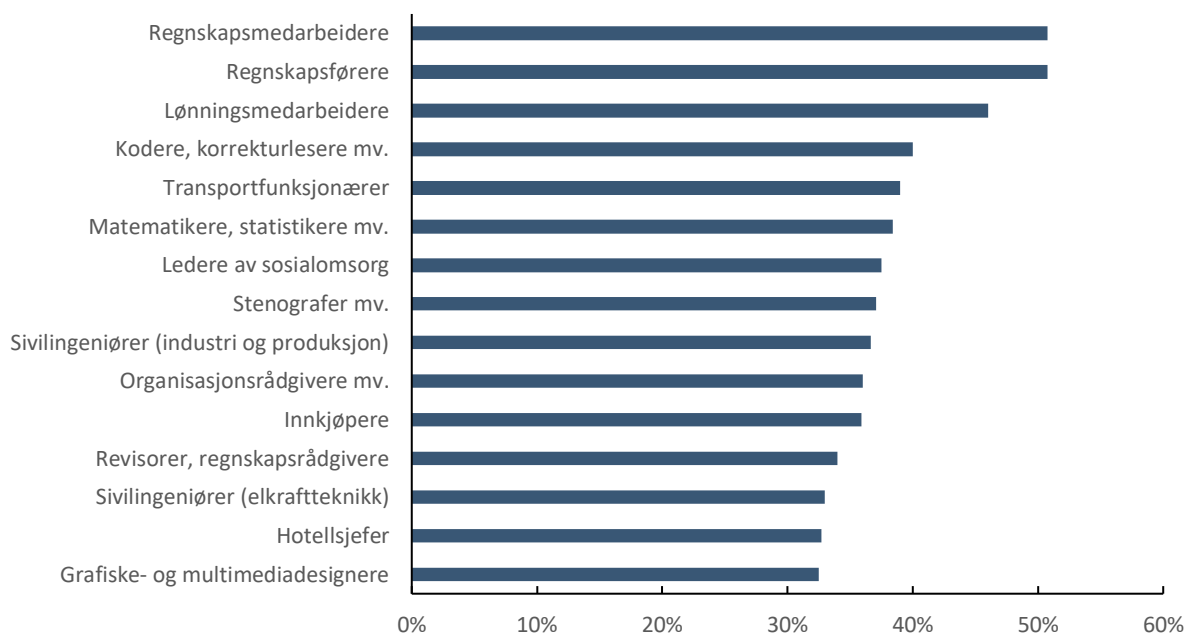
I figuren ser vi en tydelig stigende trend. Selv om det er en stor variasjon i effektiviseringspotensialet innenfor hver lønnsgruppe, ser vi en generell tendens til at høyere lønn korrelerer med større effektiviseringspotensial. Gjennomsnittlig effektiviseringspotensial ligger på omtrent 10 prosent for lønnsgrupper under 600 000 NOK, men dette tallet doubles for grupper som tjener over 600 000 NOK. Hvis vi gjentar beregningen for gjennomsnittlig tidsbesparelse i forrige delkapittel, men også tar i betraktning lønn i hvert yrke, øker gjennomsnittlig nasjonalt effektiviseringspotensial til 18.5 prosent.

Effekten av skjeve lønnsstrukturer kan manifestere seg på ulike måter. Agrawal med flere<sup>8</sup> peker på en forsterkingseffekt («augmentation»), snarere enn automatisering i mange yrkesfelt. Dette betyr at KI-verktøy kan gjøre det enklere for arbeidstakere med lavere kvalifikasjoner å utføre oppgavene sine. Et illustrerende eksempel er sjåførbransjen, hvor det var 200 000 profesjonelle taxi og limousin-sjåførere i USA i 2018. I dag kjører mer enn 10 ganger så mange for Uber alene, takket være lett tilgjengelig GPS-teknologi, noe som har ført til nedgang i lønnsnivået. Et annet eksempel er innen helse- og omsorgssektoren, der helsepersonell i økende grad kan overta diagnostiske oppgaver fra leger, med den konsekvens at de får økte lønninger og bedre karrieremuligheter.

## Opp mot 50 prosent effektiviseringspotensial i de mest eksponerte yrkene

Vi ser at det er tydelige forskjeller i effektiviseringspotensial på tvers av yrker, dette er fordi språkmodellenes evner er bedre tilpasset noen oppgaver enn andre. Lønnsstatistikken tyder på at det er en skjevhet mot oppgaver som ofte finnes i bedre betalte yrker. Vi undersøker derfor yrkene direkte, for å se hvilke grupper som blir mest berørt.

**Figur 4: Yrkesgrupper med høyest effektiviseringspotensial (potensiell tidsbesparelse i prosent). Kilde: SSB og Menon Economics**



Figur 4 viser de 15 yrkene med høyest effektiviseringspotensial, hvor maksimal tidsbesparelse er over 50 prosent. Vi merker oss også at 60 prosent effektiviseringspotensial er maksimalt oppnåelig tidsbesparelse i vår analyse. Dette gjenspeiler det faktum at arbeidsoppgaver som involverer tekstbehandling, som stilles til store språkmodeller, er særlig utsatt. Videre er KI-verktøy spesielt dyktige på kodeoppgaver, noe som forklarer hvorfor yrker som ingeniører, matematikere, statistikere og programmerere er blant de mest berørte. En annen viktig

<sup>8</sup> Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2023). *The Turing Transformation: Artificial Intelligence, Intelligence Augmentation, and Skill Premiums*. National Bureau of Economic Research, Working Paper Series, No. 31767. doi:10.3386/w31767. Retrieved from <http://www.nber.org/papers/w31767>.

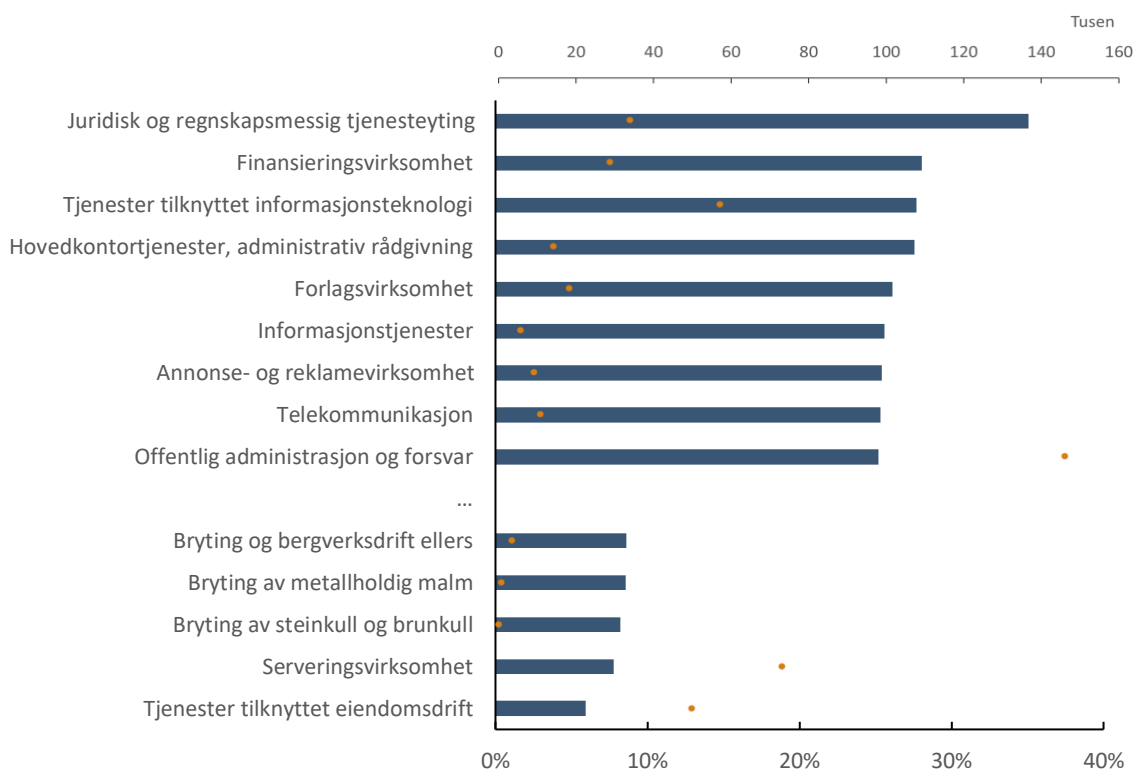
observasjon er at KI-verktøyenes evne til effektivt å planlegge og strukturere data har innvirkning på yrkesgrupper som inneholder mange ledelse- og organisasjonsrelaterte oppgaver.

Selv om telekommunikasjon og radio- og fjernsynsringkasting ved første øyekast kan virke som uventede tilføyelser, gir det mening ved nærmere ettersyn. Telekommunikasjon domineres av tekniske ingeniørstillinger, markedsførere og selgere, mens radio- og fjernsynsringkasting har en høy konsentrasjon av journalister – en yrkesgruppe med et betydelig effektiviseringspotensial. Denne sektoren inkluderer også saksbehandlere, ingeniører og oversettere, som alle kan dra nytte av KI-verktøyenes kapasiteter.

## Høyest effektiviseringspotensial i tjenesteytende næringer

Videre aggregerer vi yrkene opp til næringsnivå. Figur 5 nedenfor illustrerer effektiviseringspotensialet i prosent for de ni mest og fem minst eksponerte næringene. Den nederste akse viser effektiviseringspotensialet i prosent for næringen, hvor andelen representeres av det blå søylediagrammet. Den øvre akse viser antall sysselsatte i hver av næringene, hvor mengden er representert med de oransje prikkene.

Figur 1. Vektet eksponeringspotensial og totalt antall sysselsatte for ulike næringer. Kilde: SSB og Menon Economics



Fra figuren ser vi at mange næringer har potensial til å effektivisere en fjerdedel av arbeidstiden sin. Hovedsakelig er det tjenesteytende næringer som har et høyt effektiviseringspotensial. De mest eksponerte næringene har også like kjennetegn, som at de er tekst- og informasjonsintensive. Disse næringene inkluderer også oppgaver som innebærer store mengder data, repetitiv informasjonsbehandling eller kundeinteraksjoner, områder der kunstig intelligens har vist seg å være verdifull. Videre ser vi at de øverste næringene ofte krever høyere utdanning og generelt er bedre betalt enn næringene med lavere eksponering. Et eksempel er «Juridisk og regnskapsmessig tjenesteyting», som skiller seg ut med en vektet effektiviseringsgrad på over 35 prosent. Kunstig intelligens har vist seg å være svært effektiv til å lese, bearbeide og sammenfatte store mengder litteratur og tall, noe som kan være nyttig for jurister for å raskt trekke ut viktig informasjon fra rettsavgjørelser og lover, noe som ville tatt lengre tid for mennesker.

«Finansvirksomhet» har også mange oppgaver som kan effektiviseres, ettersom arbeidsoppgavene i denne sektoren er preget av mye dataanalyse og beslutningstøtte – områder der kunstig intelligens har vist seg å ha gode egenskaper. Effektiviseringspotensialet i denne næringen er like under 30 prosent. Vi ser også tjenester knyttet til informasjonsteknologi og informasjonstjenester på listen. Dette er en sektor som generelt har vært tidlig ute med å omfavne digitalisering, og kunstig intelligens vil sannsynligvis bare forsterke denne effekten.

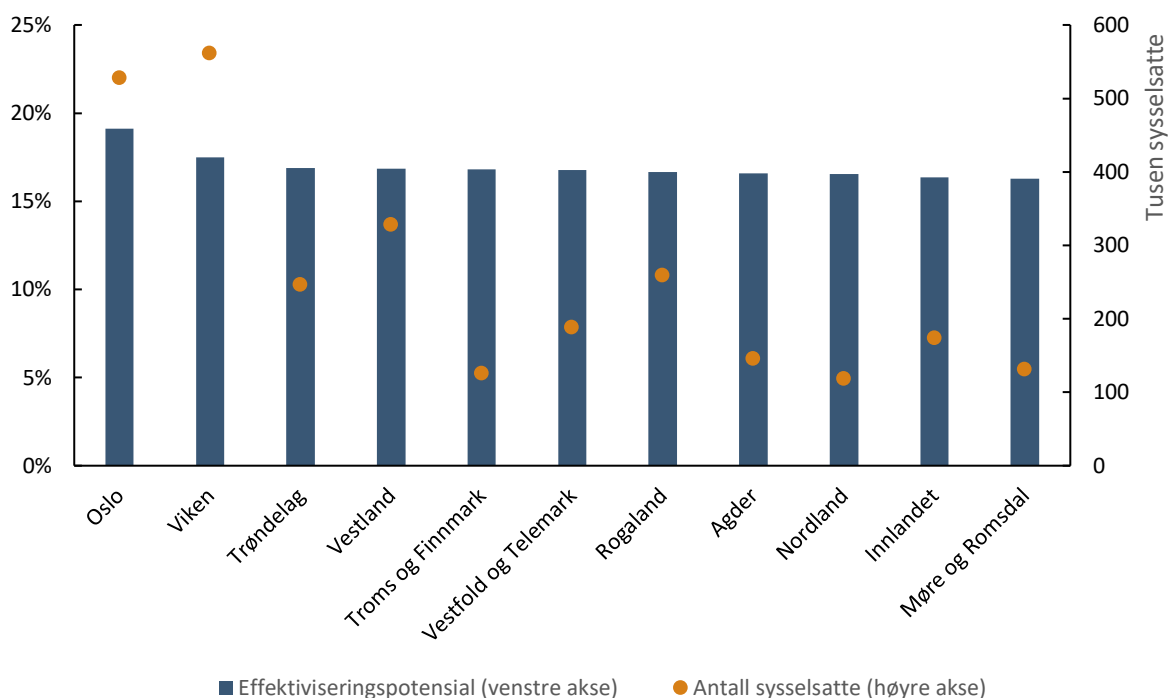
De gjenværende næringene synker gradvis fra 25 prosent ned til det laveste på 6 prosent, som er «Tjenester knyttet til eiendomsforvaltning». I motsetning til de mest eksponerte næringene, har de minst eksponerte næringene arbeidsoppgaver som krever direkte menneskelig interaksjon, fysisk arbeid og menneskelige ferdigheter, noe kunstig intelligens foreløpig ikke kan utføre. Disse næringene har også et større antall unike og mindre repetitive arbeidsoppgaver, og møter ofte flere uforutsigbare hendelser som må håndteres av mennesker.

Umiddelbart er ikke det prosentvise effektiviseringspotensialet den viktigste faktoren for å oppnå gevinst ved implementering av kunstig intelligens, det avhenger av det samlede gevinstpotensialet som også avhenger av antall sysselsatte hvis arbeidsoppgaver kan effektiviseres. Den øvre aksene representerer antall sysselsatte i de forskjellige næringene, og vi ser for eksempel at «Offentlig administrasjon og forsvar» har over 140 000 sysselsatte med over 25 prosent effektiviseringspotensial. Dette tilsvarer 35 000 sysselsatte i næringen, et tall som tilsvarer alle sysselsatte i Ålesund. Implementering av kunstig intelligens i denne næringen vil dermed kunne ha en stor total effekt på deres produksjon.

### Jevnt fordelt effektiviseringspotensial i ulike fylker, antall sysselsatte viktig

Vi har også har undersøkt den geografiske fordelingen av effektiviseringspotensial på både fylkesnivå. I figur 7 nedenfor viser vi det vektete effektiviseringspotensialet for alle norske fylker.

Figur 6. Effektiviseringspotensial i Norges elleve fylker. Kilde: SSB og Menon Economics



Fra figuren ser vi at Oslo er mest eksponert for effektivisering drevet av kunstig intelligens. Vår modell indikerer et effektiviseringspotensial på 19 prosent av arbeidsstyrken, tilsvarende 100 tusen sysselsatte. Rett etter følger Viken og Trøndelag, med et effektiviseringspotensial på henholdsvis 17,5 og 17 prosent, som tilsvarer 98 tusen sysselsatte i Viken og 41 tusen i Trøndelag. Fylkene med lavest eksponeringsgrad er Nordland, Innlandet, og Møre og Romsdal, med en eksponeringsgrad litt over 16 prosent. Dette tilsvarer 20-30 tusen sysselsatte. I figuren ser

vi også at det generelt er slik at fylker med høyere antall sysselsatte også har høyere effektiviseringspotensial.

Det er ikke stor forskjell i effektiviseringspotensialet fra Oslo, med det høyeste på 19 prosent, til Møre og Romsdal, med det laveste på 16 prosent. Dette kan indikere at yrkesfordelingen i fylkene er ganske lik, i hvert fall med tanke på fordelingen av effektiviseringspotensialet i prosent. Naturligvis vil fylker som har mange sysselsatte innenfor de mest eksponerte næringene også ha større effektiviseringspotensial på fylkesnivå. For å undersøke dette har vi brukt statistikk fra SSB på næringsfordeling blant sysselsatte i forskjellige fylker. Når vi undersøker næringsgruppen *faglig, vitenskapelig og teknisk tjenesteyting* – som inkluderer yrker som juridisk og regnskapsmessig tjenesteyting og administrativ rådgivning, ser vi selv etter korrigering for befolkningsstørrelse at Oslo og Viken har 55 prosent av alle sysselsatte i denne næringen. De tre minst eksponerte fylkene representerer kun 6 prosent av sysselsatte i samme næringskategori. Samtidig representerer Oslo og Viken en mindre, men betydelig del av næringene med lavest effektiviseringspotensial, med 33 prosent av alle sysselsatte i de 10 næringene med lavest effektiviseringspotensial. Dette bidrar til en mindre forskjell i effektiviseringsgrad mellom fylkene.

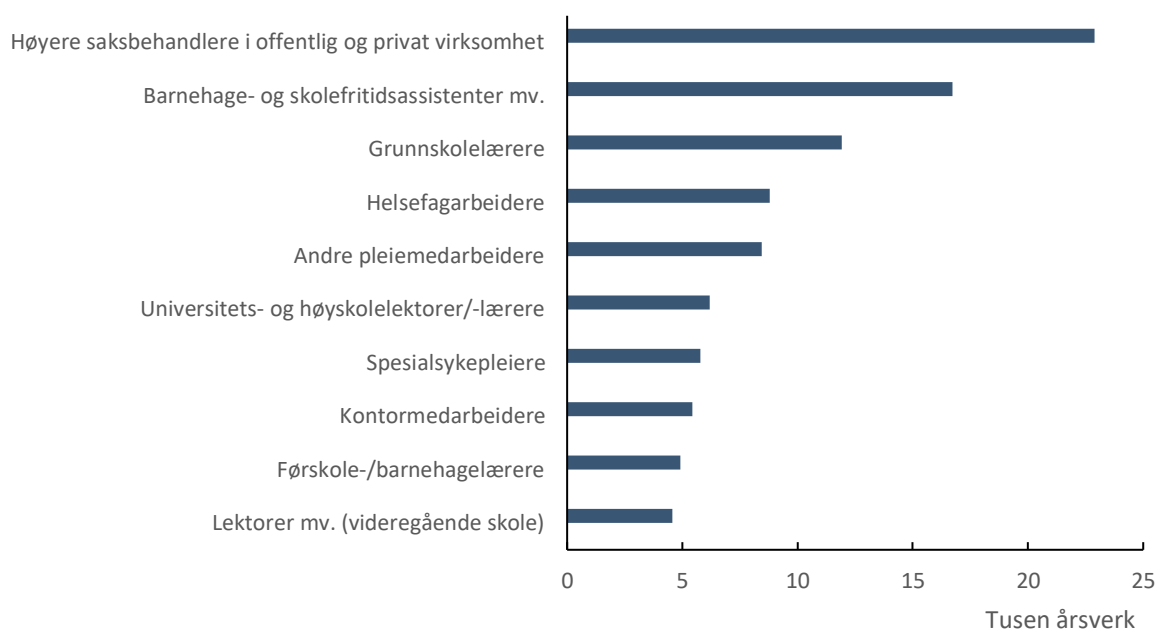
Overordnet kan alle fylker høste et effektiviseringspotensial på over 16 prosent ved at næringene utnytter kunstig intelligens i sine arbeidsoppgaver. Det er imidlertid mest sannsynlig at teknologien først vil bli implementert i større fylker hvor næringene har flere sysselsatte, siden det ofte er oppstartskostnader og stordriftsfordeler ved å implementere kunstig intelligens i arbeidsflyten.

### **I offentlig sektor er det særlig stort effektiviseringspotensial innen saksbehandling**

Figuren under illustrerer potensialet for økt effektivitet i offentlig sektor ved å benytte kunstig intelligens (KI). Aksen i figuren representerer antall årsverk KI muligens kan dekke hvis potensialet blir fullt utnyttet. Verdiene er funnet ved å multiplisere effektiviseringspotensialet med antall ansatte i yrket. Dette betyr umiddelbart at yrkene med de høyeste verdiene ikke nødvendigvis må ha størst effektiviseringspotensial i prosent, som vi har sett på tidligere.



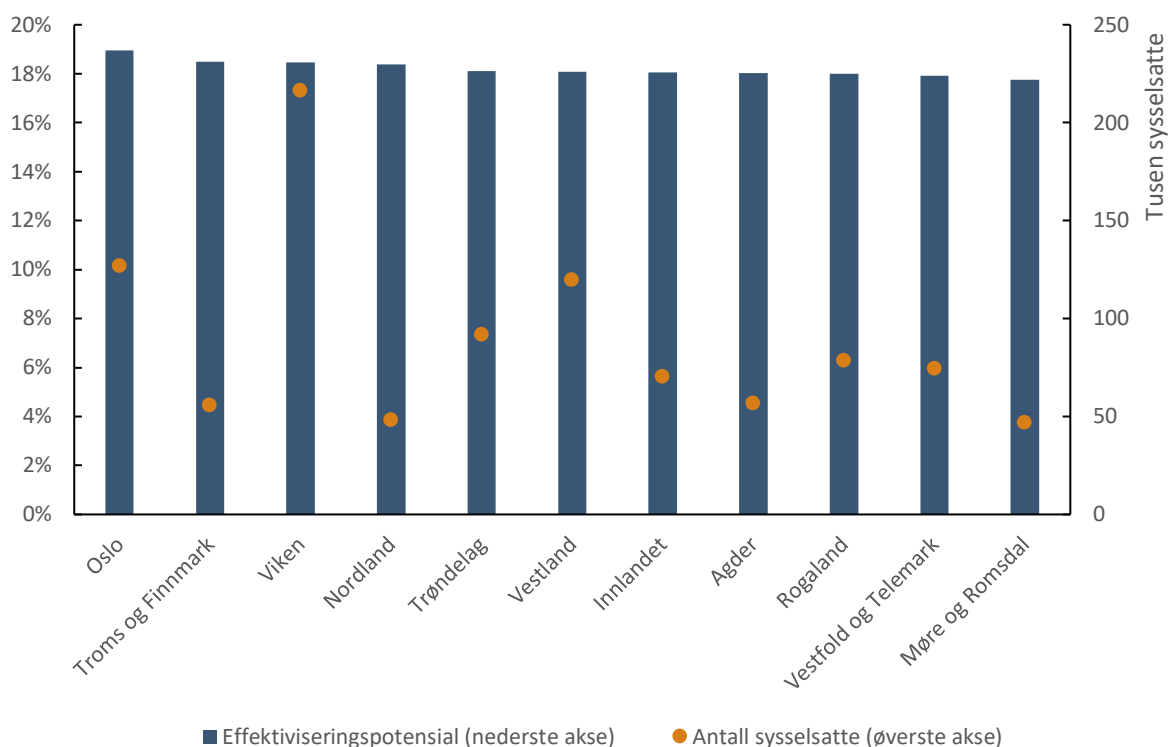
Figur 7. Top ti yrker basert på effektiviseringspotensial multiplisert med antall sysselsatte. Kilde SSB og Menon Economics



På toppen av listen finner vi yrket «Høyere saksbehandlere i offentlig og privat virksomhet», hvor KI kan dekke 22 000 årsverk, noe som tilsvarer en tredjedel av arbeidsstyrken i 2022. Dette yrket innebærer å utrede og besvare søknader, dokumenter og andre henvendelser, arbeidsoppgaver hvor KI har vist seg å være særdeles effektiv. Vi finner også læreryrker som «Barnehage- og skolefritidsassistenter m.fl.», «Grunnskolelærere» og «Universitets- og høyskolelektorer» høyt på listen, med henholdsvis 16 000, 11 000 og 6 100 årsverk med effektiviseringspotensial mellom 15 og 20 prosent. Dette er yrker som består av mange oppgaver hvor KI enda ikke er effektiv grunnet nødvendigheten av menneskelig interaksjon og ferdigheter i omgang med barn og eldre. Derimot er det også mange arbeidsoppgaver som innebærer koordinering og kommunikasjon, både internt i organisasjonen og ut mot brukerne, hvor KI kan gi store gevinster ved å øke effektiviteten og spare tid. Deretter følger omsorgsyrker som «Helsefagarbeider», «Pleiemedarbeider» og «Spesialsykepleier», hvor KI kan dekke mellom 6000 og 9000 årsverk. Omsorgsyrker krever også mye personlig kontakt med mennesker og pasienter, men sammenfaller med mye koordinering og kommunikasjonsarbeid der KI kan gi støtte. Eksempler på dette kan være journalskriving, oppfølging av pasienter og behandling av informasjon. Lenger ned på listen finner vi yrker som «Kontormedarbeidere», med 5 400 sparte årsverk, hvor Kis evne til å utføre administrative oppgaver, organisere dokumenter og assistere med kommunikasjon kan gi gevinster i form av effektivisering.

I figuren under ser vi også på de fylkene som har størst effektiviseringspotensial, kun med hensyn til de næringene vi definerer som offentlig sektor.

Figur 8. Effektiviseringspotensial i Norges elleve fylker for offentlig sektor. Kilde: SSB og Menon Economics



Fra figuren kan vi trekke mange av de samme konklusjonene som når vi analyserte fylkene generelt. Først ser vi at alle fylkene har en noe høyere effektiviseringsgrad når vi ser isolert på næringsgruppene i offentlig sektor, sammenlignet med alle norske næringer. Vi ser også at forskjellene mellom kommunene nå er mindre. Dette er ikke overraskende siden vi kun ser på offentlig sektor, hvor fordelingen av sysselsatte vil være mer lik mellom de forskjellige fylkene. Igjen er Oslo mest eksponert for KI-drevet effektivisering. I motsetning til den tidligere figuren ser vi nå at fylkene Troms og Finnmark, samt Nordland, er høyere på listen, mens Trøndelag, Rogaland, Vestfold og Telemark har falt ned. Dette kan indikere at yrkene som tidligere ble nevnt er mer konsentrert blant sysselsatte i offentlig sektor.

Modellen vår tilsier at det er store gevinster å hente ved å innføre KI i offentlig sektor i alle norske fylker. Hvis alle yrkesgrupper i offentlig sektor implementerte KI i den grad potensialet indikerer, kunne dette bidratt til en tidsreduksjon på 155 000 årsverk årlig. Dette tallet representerer nesten 18 prosent av alle sysselsatte i de valgte næringene og kan være en løsning på mange norske virksomheters mangel på arbeidstakere og helsepersonell.<sup>9</sup>

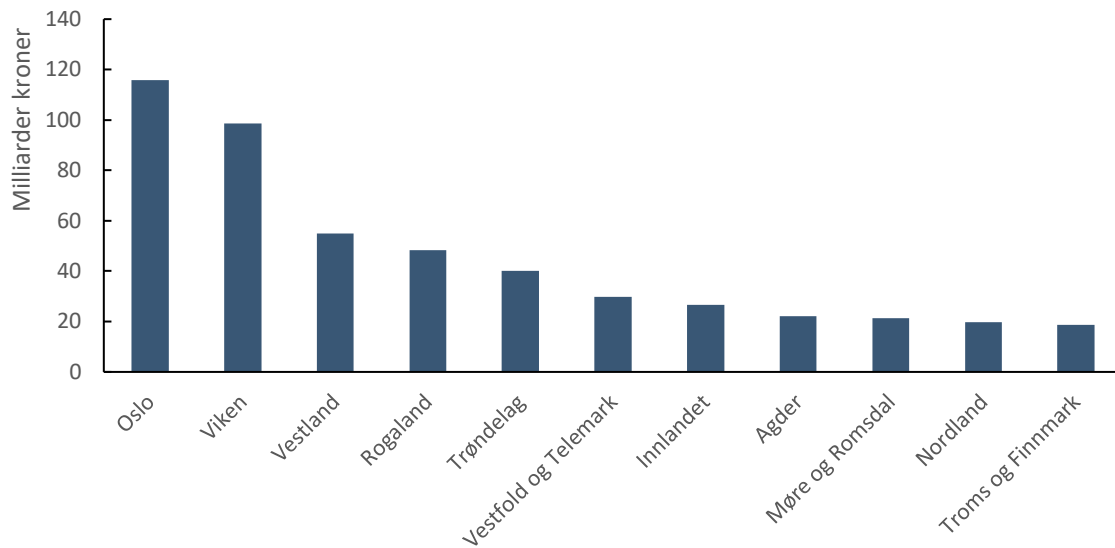
### Effektiviseringspotensialet er verdt opp mot 500 milliarder årlig

Som diskutert tidligere kan denne effektiviseringen enten resultere i økt produksjon for den samme innsatsmengden, eller i økt fritid. I begge tilfeller medfører effektiviseringen samfunnsøkonomisk nytte. Vi kan estimere størrelsen på denne nytten basert på tall fra Menons regnskapsdatabase, nasjonalregnskapet og det beregnede effektiviseringspotensialet. Samlet estimerer vi dette til 500 milliarder kroner årlig.

<sup>9</sup> [Arbeidsmarkedet for helsepersonell fram mot 2040 \(ssb.no\)](https://ssb.no)

Fordelt på fylker finner vi verdien av den KI-drevne effektiviseringen, som vist på figuren under:

**Figur 10. Mulig økning verdiskaping med bruk av KI. Kilde: SSB, Menon Economics**



For å kvantifisere potensialet som kan realiseres ved å bruke store språkmodeller, har vi estimert en potensiell verdiskaping ved å implementere KI fullstendig i alle næringer på fylkesnivå. Det er viktig å understreke at disse estimatene kommer med et forbehold om at KI kun effektiviserer de arbeidsoppgavene som ble utført i 2021, og ikke tar hensyn til de ekstra gevinstene KI kan gi ved å muliggjøre arbeid vi i dag ikke får utført i det hele tatt, noe som kan gi enda større gevinster. I figuren over ser vi at KI i 2021 kunne ha bidratt til en verdiskaping i Oslo på 115 milliarder kroner årlig. Videre ser vi at Viken og Vestland har mulighet til å generere henholdsvis 98 og 54 milliarder kroner årlig. De nordligste fylkene har lavest potensial for verdiskaping med KI, med en potensiell verdiskaping på omtrent 20 milliarder kroner. Dette er bare estimater, men de kan gi innsikt i de potensielle gevinster KI kan gi i kroner og øre.

## Metodikk

Formålet er å vurdere effektiviseringspotensialet til ulike yrker som følge av generativ KI, hvor vi definerer effektiviseringspotensial til å være graden av tidsbesparelse i et yrke. For å komme frem til dette tallet har vi brutt yrker ned i sine enkelte arbeidsoppgaver og brukt KI til å klassifisere hver arbeidsoppgave etter dennes eksponering for KI-verktøy. Vi deler arbeidsoppgavene i fire kategorier: minimal effekt (ME), god effekt (GE), effekt med ekstraverktøy (EE) og effekt med synsevner (ES). De to siste kategoriene betyr at arbeidsoppgaven kan effektiviseres godt dersom språkmodellen har et ekstra økosystem bygget rundt seg, eller evnen til å se bilde eller film. Vi bruker så kategoriene til å kvantifisere effektiviseringspotensial basert på antall arbeidsoppgaver som kan effektiviseres i hvert yrke, dette vil være den beskrivende faktoren videre i analysen. I dette kapitlet gir vi en grundig og teknisk forklaring av vår fremgang til data og modell.

### En kombinasjon av amerikanske og norske data

I vår analyse bruker vi en videreutvikling av metodikken beskrevet i “*GPTs are GPTs*” av OpenAI <sup>10</sup>. I likhet med OpenAI, benytter vi oss av O\*NET-databasen <sup>11</sup>, som inneholder en liste med over 923 yrker og tilhørende arbeidsoppgaver. Hver yrkesgruppe i databasen har i gjennomsnitt 20 forskjellige arbeidsoppgaver beskrevet.

#### **Input:**

**Yrke:** Dispatchers, Except Police, Fire, and Ambulance

**Task type:** Core

**Task:** Schedule or dispatch workers, work crews, equipment, or service vehicles to appropriate locations, according to customer requests, specifications, or needs, using radios or telephones.

#### **GPT-4 kategorisering:**

**Label:** ES (effekt med synsevner)

**Forklaring:** While the LLM can assist in organizing and scheduling tasks, it would require additional software to integrate with existing systems to dispatch workers, crews, equipment, or vehicles to appropriate locations.

O\*NET-databasen er en amerikansk database og tilsvarer ikke yrkesdataene som utgis av SSB. Vi bruker derfor et kryssreferansedatasett fra *US Bureau of Labor Statistics*<sup>12</sup> for å konvertere de ulike yrkeskodene til ISCO, som harmonerer med de norske yrkesdataene.

Videre blir supplerende statistikk for hver yrkesgruppe hentet fra Statistisk sentralbyrå (SSB) for videre analyse.

---

<sup>10</sup> Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *Gpts are gpts: An early look at the labor market impact potential of large language models*. *arXiv preprint arXiv:2303.10130*.

<sup>11</sup> [https://www.onetcenter.org/dictionary/20.1/excel/task\\_statements.html](https://www.onetcenter.org/dictionary/20.1/excel/task_statements.html)

<sup>12</sup> [https://www.bls.gov/soc/ISCO\\_SOC\\_Crosswalk.xls](https://www.bls.gov/soc/ISCO_SOC_Crosswalk.xls)

- **08536: Kjønn- og næringsfordeling (88 grupper) blant sysselsatte (15-74 år). 4. kvartal (K) 2008 – 2022:** Denne tabellen beskriver sysselsettingen for bosatte i Norge på et detaljert nivå. Dataene er brukt til å undersøke effektiviseringspotensialet på næring- og fylkesnivå.
- **Sysselsatte 15-74 år, etter arbeidsstedsfylke. Person og prosent. 4. kvartal:** Denne tabellen beskriver sysselsettingen for bosatte i Norge på et detaljert og regionalt nivå.
- **09391: Hovedtall fylkesfordelt nasjonalregnskap, etter region, statistikkvariabel og år:** Denne tabellen viser det fylkesfordelte nasjonalregnskapet.
- **Spesialbestilling fra SSB:** Denne tabellen inneholder enda mer finmasket data på sysselsatte i Norge enn vanlig tabeller SSB leverer. Her finner vi antall sysselsatte innenfor ett spesifikt yrke for alle næringene. Tabellen inneholder også en ISCO-08-Kode, som er brukt til å knytte tabellen til vår data på effektiviseringspotensialet i prosent.

### Små justeringer i modellen fra GPTs are GPTs

Modellen vi benytter for å vurdere effektiviseringspotensialet i ulike yrker er basert på kategorisering av arbeidsoppgaver ved hjelp av en språkmodell, mer spesifikt GPT-4-motoren. Hver arbeidsoppgave blir kategorisert i én av fire mulige kategorier (ME, GE, EE og ES) ved hjelp av et detaljert instruksjonsark (Se vedlegg 1), som klargjør betingelsene for hver kategori.

For å øke nøyaktigheten i vår analyse, blir resultatene av GPT-4s kategorisering sammenlignet med et mindre utvalg av manuelt kategoriserte oppgaver, som hjelper å identifisere viktige motsetninger mellom menneskers og GPT-4s valg. I tillegg blir resultater evaluert løpende gjennom analyseprosessen for å avdekke eventuelle inkonsistenser. For eksempel blir noen spesielt utfordrende yrkesgrupper, blant andre «Dispatchers, Except Police, Fire, and Ambulance» og «Telemarketers» undersøkt for hver kategorisering for å evaluere nøyaktigheten. Disse to yrkene inneholder oppgaver som inkluderer verbal kommunikasjon, der GPT-4 har for vane å foreslå effektivisering ved å generere manus fortløpende. Det var nøyaktig denne egenskapen som ble testet av Brynjolfsson et. Al.<sup>13</sup>, som konkluderte med god effekt for uerfarne ansatte, men lite til ingen for erfarne.

To metriske indikatorer,  $\alpha$  og  $\beta$ , brukes deretter for å kvantifisere en samlet score per yrke.  $\alpha$  beregnes som en vektet andel av oppgaver kategorisert som «God effekt», der kjerneoppgaver vektet dobbelt i henhold til O\*NET-databasen.  $\beta$  er en utvidelse av  $\alpha$  og inkluderer også «effekt med ekstraverktøy» og «effekt med synsevner», men kun med 50% vekt. Formlene uten vekting blir da henholdsvis,

$$\alpha = \frac{n_{\text{god effekt}}}{n_{\text{total}}}, \quad \beta = \frac{n_{\text{god effekt}} + 0.5 \times (n_{\text{effekt med ekstraverktøy}} + n_{\text{effekt med synsevner}})}{n_{\text{total}}}$$

Det endelige resultatet er et datasett med 923 yrkesrader, hver med en score beregnet ved hjelp av  $\alpha$ - eller  $\beta$ -indikatorerne. Til slutt beregner vi effektiviseringspotensial ved å multiplisere  $\beta$ -scoren med 0.6. Dette gjøres for å få et tall som kan tolkes som direkte tidsbesparelse, fordi kategoriene er definert som tidsreduksjon mellom 50 og 100 prosent. 60 prosent er valgt som et konservativt anslag.

---

<sup>13</sup> Erik Brynjolfsson, Danielle Li, Lindsey R. Raymond, *Generative AI at Work*. NBER Working Paper No. 31161, April 2023. [DOI: 10.3386/w31161](https://doi.org/10.3386/w31161)

## Vi kobler yrkesdata med næringsdata for å analysere variasjonen

Videre kobles datasettet mot sysselsettingstall for det norske arbeidsmarkedet, og gir dermed viktige indikasjoner på hvordan store språkmodeller kan komme til å påvirke arbeidsstyrken i Norge. Vi har brukt sysselsettingstabeller fra SSB og koblet sammen yrkenes effektiviseringspotensial med yrkeskomposisjonene i forskjellige norske næringskoder<sup>14</sup>. For å bestemme effektiviseringspotensialet i de ulike næringene, multipliserte vi yrkenes effektiviseringsgrad med antall ansatte per yrke og summerte resultatet for hver næring. Inndelingen av næringsstandard og næringskoder er basert på SN2007 fra SSB.<sup>15</sup>

$$\text{Effektiviseringspotensial i næring} = \sum_{i=1}^n \frac{\beta_i * \text{antall sysselsatte i yrkesgruppen}_i}{\text{antall ansatte i næringen}}$$

Vi har videre brukt disse dataene til å gjøre analyser på fylkesnivå ved å koble på dataene med relevante tabeller på næringsfordeling fra SSB. Det har også blitt gjort en analyse med vekt på offentlig sektor, ved å definere offentlig sektor som næringene «Offentlig administrasjon og forsvar, samt trygdeordninger underlagt offentlig forvaltning», «Undervisning» og «helse og sosialtjenester». Disse næringene består hovedsakelig av sysselsatte i offentlig sektor, selv om det også kan forekomme sysselsatte i private sektorer, for eksempel private helsetjenester. Den mulige inkludering av private tjenester gjør også at totalt antall sysselsatte er noe høyere enn for offentlig sektor alene.

I offentlig sektor finner vi ofte de samme yrkene som vi gjør i næringslivet generelt, men fordelingen av antall sysselsatte i yrkene er ofte annerledes. For å hindre at vi kun gjentar resultatene vi finner på yrkesnivå, har vi valgt å multiplisere effektiviseringsgraden med antall ansatte i hvert yrke for å estimere antall «årsverk» KI kan dekke. Dette betyr at det ikke kun vil være effektiviseringspotensialet i prosent som vil ha innvirkning, men også antall sysselsatte i yrket.

Et viktig forbehold er at disse kalkulasjonene kun er estimater, da koblingen mellom effektiviseringsgrad i prosent for næringer koblet over til geografiske effekter kan bli upresis. Et eksempel er at ett fylke kan ha en annen intern yrkesfordeling i en næringsgruppe enn fordelingen er på nasjonalt nivå, og da vil vår gruppering feilaktig vekte arbeidsoppgavene i enten for stor eller for lite effektivisering enn det som er den faktiske i det fylkes næringsgruppe.

## Vi benytter en mer konservativ kategorisering

En nøkkelforskjell mellom vår metode og den benyttet av OpenAI's artikkel, er knyttet til kategoriseringen av hver enkelt arbeidsoppgave. I motsetning til OpenAI, gjennomfører vi analysen fem ganger for hver arbeidsoppgave, noe som gir litt forskjellig svar hver gang. Variasjonene i svarene kan tilskrives flere faktorer. En av dem er språkmodellens innebygde stokastiske egenskaper, som betyr at identiske input ikke alltid vil generere nøyaktig samme output. Videre har det også vært små justeringer i instruksjonene underveis i studien. Da kategoriseringen har foregått over en seksmånedersperiode, kommer noen av variasjonene som en naturlig konsekvens av kontinuerlig utvikling og oppdatering av GPT-4.

Vår konservative strategi innebærer at vi systematisk velger det mest pessimistiske av de fem svarene som modellen genererer. Hvis et av fem svar er «Minimal effekt», blir dette det foretrukne svaret. På denne måten

---

<sup>14</sup> [Standard for næringsgruppering \(SN\) \(ssb.no\)](#)

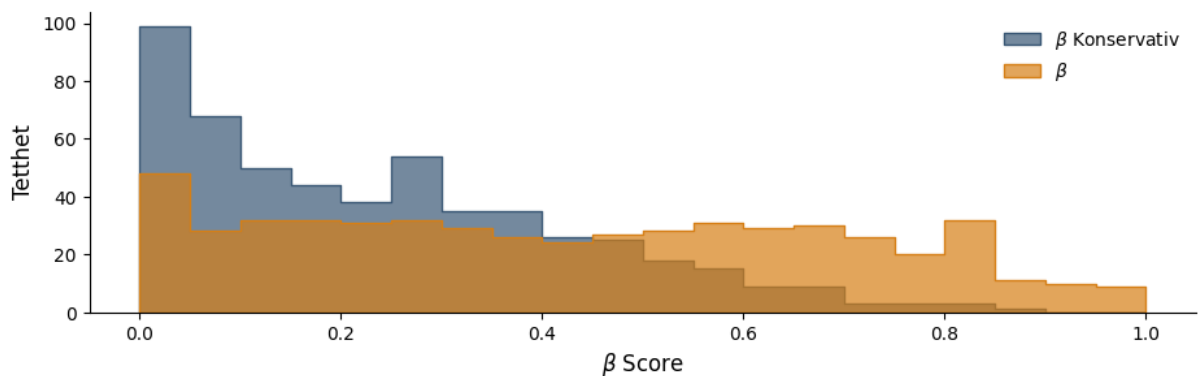
<sup>15</sup> [Næringsstandard og næringskoder - SSB](#)



mener vi å gi et mer realistisk og robust bilde av hvordan avanserte språkmodeller som GPT-4 kan påvirke effektiviteten i forskjellige yrker.

En sammenligning mellom vår konservative metode og den benyttet av OpenAI er visualisert i Figur 1. Den viser fordelingen av  $\beta$ -score (fordelingen vil være lik for effektiviseringspotensialet, men med verdier mellom 0 og 0.6). Fra denne figuren kan vi se tyde et visst optimistisk bias, hvor flere yrker får alle sine arbeidsoppgaver berørt. Vi tolker at GPT-4 har en tendens til å overvurdere sin egen evne til å automatisere eller effektivisere oppgaver og mener derfor den konservative fordelingen vi benytter, er mer realistisk.

**Figur 2: Effektiviseringspotensial i det norske arbeidsmarkedet ved bruk av to metodologiske tilnærminger. Kilde: Menon Economics.**



Det er viktig å merke seg at  $\alpha$  og  $\beta$ , som går fra 0 til 1, er mål på andelen arbeidsoppgaver som er berørt av KI. De representerer altså *ikke* den prosentvise reduksjonen i arbeidstid. For å gjøre et anslag på tidsbesparelse har vi gjort følgende antagelser:

1. Hver kjerneoppgave opptar like mye tid av arbeidsdagen i et yrke.
2. Hver tilleggsoppgave opptar like mye tid av arbeidsdagen i et yrke.

*Som en konsekvens av 1 og 2 antar vi derfor at  $\alpha$  og  $\beta$  er gitt et estimat på andel arbeidsoppgaver berørt.*

3. Den gjennomsnittlige tidsbesparelsen for kategorier med målbar effekt (GE, EE og ES) er 60 prosent.

*En oppgave faller i disse kategoriene dersom tidsbesparelse vurderes til å være mellom 50 og 100 prosent. 60 prosent er derfor å anse som et konservativt anslag.*

- 4.

Samlet betyr dette at effektiviseringsscore  $\alpha$  og  $\beta$  kan oversettes til

$$\text{Effektiviseringspotensial} = 0.6\beta,$$

Hvor tallet kan tolkes som direkte tidsbesparelse.

## Vedlegg: Rubrikk

Under følger rubrikken som er benyttet som instruksjon for språkmodellen. Vi har brukt OpenAI sin metode som utgangspunkt, og kategoriene har derfor litt andre navn. I rubrikken tilsvarer E0, E1, E2 og E3 våre kategorier ME, GE, EE og ES.

### # E Exposure Taxonomy

Consider the most powerful OpenAI large language model (LLM) This model can complete many tasks

that can be formulated as having text input and text output where the context for the input can be captured in 2000 words. The model also cannot draw up-to-date facts (those from <1 year ago) unless they are captured in the input.

Assume you are a worker with an average level of expertise in your role trying to complete the given task. You have access to the LLM as well as any other existing software or computer hardware tools mentioned in the task. You also have access to any commonly available technical tools accessible via a laptop (e.g. a microphone, speakers, etc.). You do not have access to any other physical tools or materials.

Please label the given task according to the taxonomy below. ## E0 – No exposure

Label tasks E0 if direct access to the LLM through an interface like ChatGPT or the OpenAI playground cannot reduce the time it takes to complete this task with equivalent quality by half or more.

If a task requires a high degree of human interaction (for example, in person demonstrations) then it should be classified as E0.

Label as E0 or E2 if the task requires real-time verbal correspondence or audio communication via radio or telephone, even if an LLM could assist by writing scripts.

Very specialized and repetitive tasks are likely performed frequently by the worker, so the utility of the LLM may be limited to the initial learning phase of the job and should be labeled E0. Label as E0 if an LLM only contributes to time reduction the first time the task is done.

### ## E1 – Direct exposure

Label tasks E1 if direct access to the LLM through an interface like ChatGPT or the OpenAI playground alone can reduce the time it takes to complete the task with equivalent quality by at least half. This includes tasks that can be reduced to: - Writing and transforming text and code according to complex instructions, - Providing edits to existing text or code following specifications, - Writing code that can help perform a task that used to be done by hand, - Translating text between languages, - Summarizing medium-length documents,

- Providing feedback on documents, - Answering questions about a document, or - Generating questions a user might want to ask about a document.

### ## E2 – Exposure by LLM-powered applications

Label tasks E2 if having access to the LLM alone may not reduce the time it takes to complete the task by at least half, but it is easy to imagine additional software that could be developed on top of the LLM that would reduce

the time it takes to complete the task by half. This software may include capabilities such as: - Summarizing documents longer than 2000 words and answering questions about those documents - Retrieving up-to-date facts from the Internet and using those facts in combination with the LLM capabilities - Searching over an organization's existing knowledge, data, or documents and retrieving information

Examples of software built on top of the LLM that may help complete worker activities include: - Software built for a home goods company that quickly processes and summarizes their up-to-date internal data in customized ways to inform product or marketing decisions - Software that is able to suggest live responses for customer service agents speaking to customers in their company's customer service interface - Software built for legal purposes that can quickly aggregate and summarize all previous cases in a particular legal area and write legal research memos tailored to the law firm's needs - Software specifically designed for teachers that allows them to input a grading rubric and upload the text files of all student essays and have the software output a letter grade for each essay - Software that retrieves up-to-date facts from the internet and uses the capabilities of the LLM to output news summaries in different languages

### ## E3 – Exposure given image capabilities

Suppose you had access to both the LLM and a system that could view, caption, and create images. This system cannot take video media as inputs. This system cannot accurately retrieve very detailed information from image inputs, such as measurements of dimensions within an image. Label tasks as E3 if there is a significant reduction in the time it takes to complete the task given access to a LLM and these image capabilities: - Reading text from PDFs, - Scanning images, or - Creating or editing digital images according to instructions.

### ## Annotation examples:

Occupation: Inspectors, Testers, Sorters, Samplers, and Weighers Task: Adjust, clean, or repair products or processing equipment to correct defects found during inspections. Label (E0/E1/E2/E3): E0 Explanation: The model does not have access to any kind of physicality, and more than half of the task (adjusting, cleaning and repairing equipment) described requires hands or other embodiment.

Occupation: Computer and Information Research Scientists Task: Apply theoretical expertise and innovation to create or apply new technology, such as adapting principles for applying computers to new uses. Label (E0/E1/E2/E3): E1 Explanation: The model can learn theoretical expertise during training as part of its general knowledge base, and the principles to adapt can be captured in the text input to the model.

Activity: Schedule dining reservations. Label (E0/E1/E2/E3): E2 Explanation: Automation technology already exists for this (e.g. Resy) and it's unclear what an LLM offers on top of using that technology (no-diff). That said, you could build something that allows you to ask the LLM to make a reservation on Resy for you. (E3)

Activity: Negotiate purchases or contracts. Label (E0/E1/E2/E3): E2 Explanation: You could have each party transcribe their point of view and then feed this to an LLM to resolve any disputes (E3). That said, many people would need to buy into using new technological tools to accomplish this (system).

Occupation: Allergists and Immunologists Task: Prescribe medication such as antihistamines, antibiotics, and nasal, oral, topical, or inhaled glucocorticosteroids. Label (E0/E1/E2/E3): E2 Explanation: The model can provide guesses for different diagnoses and write prescriptions and case notes. However, it still requires a human in the loop using their judgment and knowledge to make the final decision.

Output list with items "index: {TaskID }label: {label} explanation: {5 word explanation (Without "LLM can" or similar)}" separated by ";"

## Referanseliste

Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2023). The Turing Transformation: Artificial Intelligence, Intelligence Augmentation, and Skill Premiums. *National Bureau of Economic Research, Working Paper Series, No. 31767*. <https://doi.org/10.3386/w31767>.

Arbeidsmarkedet for helsepersonell fram mot 2040. (n.d.). Hentet fra <https://www.ssb.no/arbeid-og-lonn/sysselsetting/artikler/arbeidsmarkedet-for-helsepersonell-fram-mot-2040>.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at Work. *NBER Working Paper No. 31161*. <https://doi.org/10.3386/w31161>.

Dell'Acqua, F. et al. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *SSRN Electronic Journal* [Preprint]. <https://doi.org/10.2139/ssrn.4573321>.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.

Gizmodo's owner replaced its Spanish language journalists with AI – The Verge. (n.d.). Hentet fra <https://www.theverge.com/2023/9/1/23856029/gizmodo-shuts-down-spanish-language-site-ai-translations>

ISCO to SOC crosswalk – [https://www.bls.gov/soc/ISCO\\_SOC\\_Crosswalk.xls](https://www.bls.gov/soc/ISCO_SOC_Crosswalk.xls).

Næringsstandard og næringskoder – SSB. (n.d.). Hentet fra <https://www.ssb.no/virksomheter-foretak-og-regnskap/naeringsstandard-og-naeringskoder>.

OECD. (2023). OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market. *OECD Publishing, Paris*. <https://doi.org/10.1787/08785bba-en>.

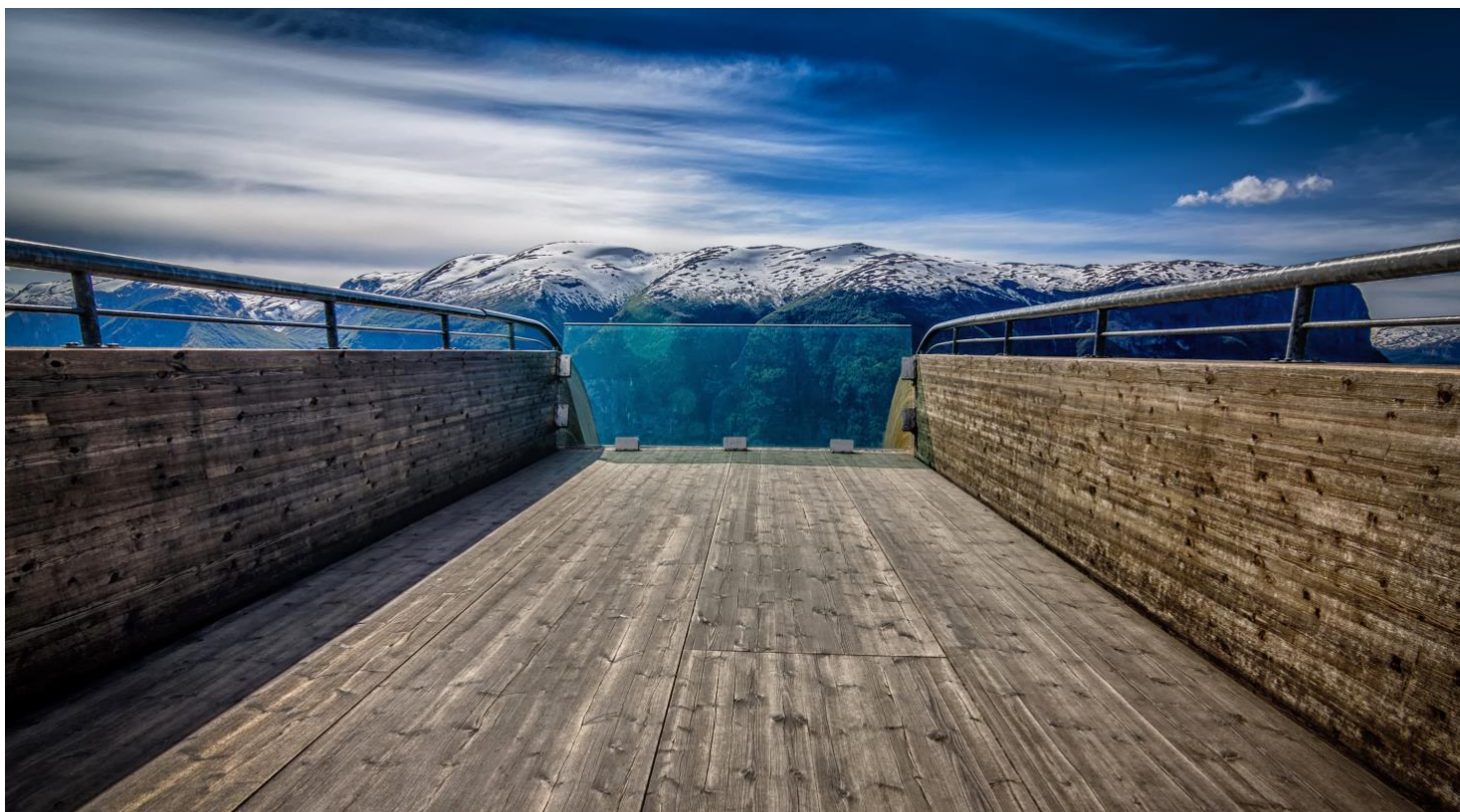
O\*NET Occupations and tasks - [https://www.onetcenter.org/dictionary/20.1/excel/task\\_statements.html](https://www.onetcenter.org/dictionary/20.1/excel/task_statements.html).

Shakked Noy, & Whitney Zhang. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381, 187-192. <https://doi.org/10.1126/science.adh2586>.

Standard for næringsgruppering (SN). (n.d.). Hentet fra <https://www.ssb.no/klass/klassifikasjoner/6>.







Menon Economics analyserer økonomiske problemstillinger og gir råd til bedrifter, organisasjoner og myndigheter. Vi er et medarbeidereiet konsultentselskap som opererer i grenseflatene mellom økonomi, politikk og marked. Menon kombinerer samfunns- og bedriftsøkonomisk kompetanse innenfor fagfelt som samfunnsøkonomisk lønnsomhet, verdsetting, nærings- og konkurranseøkonomi, strategi, finans og organisasjonsdesign. Vi benytter forskningsbaserte metoder i våre analyser og jobber tett med ledende akademiske miljøer innenfor de fleste fagfelt. Alle offentlige rapporter fra Menon er tilgjengelige på vår hjemmeside [www.menon.no](http://www.menon.no).

+47 909 90 102 | [post@menon.no](mailto:post@menon.no) | Sørkedalsveien 10 B, 0369 Oslo | [menon.no](http://menon.no)