



ChatGPT: Stoler vi på det nye vidunder-verktøyet?

En kvantitativ studie av tillit til ChatGPT sammenlignet med Google og Wikipedia

Ine Ervik & Ingrid Oliva Sandblåst Hoem

Veileder: Helge Thorbjørnsen

Selvstendig masterutredning innen Økonomisk Styring og
Strategi og Ledelse

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer innestår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

KI erklæring

I arbeidet med denne masteroppgaven har vi benyttet kunstig intelligens (KI), mer spesifikt ChatGPT, som et støtteverktøy for systematisk språkvask ved å forbedre tekstkvalitet. ChatGPT er brukt for å hjelpe oss med å uttrykke oss klarere og mer presist. Vi har også brukt ChatGPT til å hjelpe oss med å programmere i R. All faglig vurdering og analyse er utført selvstendig. Endelig tekstbearbeiding og ansvar for innholdet er vårt eget.

Sammendrag

November 2022 ble ChatGPT lansert for offentligheten. På svært kort tid fikk chatboten en enorm brukermasse, og utfordret flere ulike markeder, blant annet markedet for informasjonsinnhenting. Google og Wikipedia, som tidligere var de største aktørene, fikk nå konkurranse fra kunstig intelligens. Kunstig intelligens, som ChatGPT, er imidlertid ikke feilfri, noe som har skapt bekymring rundt ukritisk bruk av teknologien, da den kan produsere feilinformasjon. Hvor stor ChatGPT kommer til å bli fremover er uklart, men allerede nå ser vi at omfanget av plattformen stadig øker. Med dette som utgangspunkt har formålet med oppgaven vært å undersøke tilliten til ChatGPT sammenlignet med Google og Wikipedia. I tillegg har vi undersøkt om en advarsel av at ChatGPT kan ta feil påvirker brukernes tillit.

Vi har benyttet kvantitativ metode, og utarbeidet seks sett med hypoteser, som er oppsummert i en forskningsmodell. Hypotesene er basert på seks ulike tillitsdimensjoner: tre etablerte dimensjoner (*evne*, *velvilje* og *integritet*) samt tre nyere dimensjoner tilpasset teknologier (*funksjonalitet*, *hjelpsomhet* og *pålitelighet*). For å teste hypotesene har vi gjennomført et eksperiment i form av en spørreundersøkelse, og rekruttert respondenter via Profilic. Ved å gjennomføre ANOVA, Tukey HSD, lineære regresjonsanalyser og Welch t-test, har vi avdekket hvordan tillitsnivået varierer mellom de ulike informasjonsverktøyene.

Hovedfunnene i denne oppgaven viser at det er signifikante tillitsforskjeller mellom ChatGPT og Google når det gjelder de nyere tillitsdimensjonene. Google oppfattes som mer funksjonelt og pålitelig enn ChatGPT, men begge verktøyene er sett på som like hjelpsomme. ChatGPT ansees som mer hjelpsom enn Wikipedia, men det er ingen signifikante forskjeller i funksjonalitet og pålitelighet. Vi får ingen utslag når vi tester tillit mellom verktøyene med de tradisjonelle tillitsdimensjonene som utgangspunkt.

Flere faktorer påvirker tillitsnivået på tvers av gruppene. Vi finner at både hvor lenge og hvor hyppig et individ bruker et verktøy har en positiv effekt på tillit til verktøyene. Videre er det en positiv sammenheng mellom tillit og generell tilbøyelighet til å stole på nye teknologier. Vi finner ingen funn på at en advarsel om at ChatGPT kan ta feil har effekt på tillit.

Forord

Denne oppgaven er skrevet i forbindelse med mastergradsstudiet ved Norges Handelshøyskole (NHH). Masteroppgaven utgjør 30 studiepoeng, og er en utredning innen hovedprofilene Økonomisk Styring (BUS) og Strategi og Ledelse (STR).

Som studenter på en av Norges ledende handelshøyskoler, var det ikke overraskende at majoriteten av studentene hev seg på ChatGPT-trenden da chatboten ble lansert i november 2022. Vi gikk fra å skrive alt tekstmateriale selvstendig, og klø oss i hodet over at idemyldringen gikk sakte eller at Google Translate ikke oversatte korrekt fra norsk til engelsk, til å plutselig ha en maskin som kunne løse alle våre utfordringer mer effektivt, nøyaktig og riktig. Trodde vi. ChatGPT gikk fra å være et vidunder-verktøy og nåtidens «Jesus-øyeblikk», til å stadig vise at den ikke var ufeilbarlig. Men, observerte vi endringer i våre medstudenters atferd med chatboten? Nei. Til tross for flere nyheter om at man burde være forsiktig med bruken av ChatGPT, hørte vi stadig om studenter som hadde levert fullverdige oppgaver produsert ene og alene av ChatGPT. Dette vekket en nysgjerrighet i oss. Var det slik at studenter på NHH er de eneste som er naive og fokuserer på de positive sidene ved ChatGPT, og neglisjerer de negative? Eller er dette en allmenn holdning? Hvordan er *egentlig* tilliten til ChatGPT?

I midten av vår masterskriving fikk vi muligheten til å stille som hårmodeller. Her møtte vi både unge og eldre frisører som aldri engang hadde hørt om ChatGPT. Vi lurte på hva de brukte for å innhente informasjon, og de svarte Google eller Wikipedia. Dette ga oss inspirasjon til å teste tillitsnivået for både etablerte informasjonsverktøy, men også markedets nykommer.

Våre kropper består nok av 70% koffein etter denne avhandlingen, og vi føler det er på sin plass å rette en takk til Circle K som har holdt dørene åpne døgnet rundt, og alltid hatt fire ulike kaffemaskiner å velge mellom. Uten ørten antall kaffepåfyll i løpet av en dag, er det ikke sikkert vi hadde kommet oss gjennom skrivingen. Eller vedlikeholdt det gode vennskapet oss mellom. Vi ønsker også å rette en stor takk til vår veileder, Helge Thorbjørnsen, som har gitt oss gode tilbakemeldinger underveis og vist stor interesse for vårt tema.

Bergen, desember 2024

Ine Ervik og Ingrid Oliva Sandblåst Hoem

Innholdsfortegnelse

KI erklæring	2
Sammendrag.....	3
Forord.....	4
Innholdsfortegnelse	5
1. Innledning	8
2. Bakgrunn.....	11
2.1 Kunstig intelligens	12
2.2 Generativ KI og ChatGPT	12
2.3 Store språkmodeller	12
2.4 ChatGPT som chatbot-applikasjon.....	13
2.5 Informasjonsverktøy	13
3. Litteratur og forskning	14
3.1 Tillit	15
3.1.2 De tradisjonelle dimensjonene av tillitsfull overbevisning	16
3.1.3 Tidligere forskning	18
3.2 Tillit til teknologi	19
3.2.1 HCI.....	19
3.2.2 Tilbøyelighet til å stole på teknologier.....	20
3.2.3 Brukserfaring.....	20
3.2.4 De nye dimensjonene av tillitsfull overbevisning	21
3.3 Tilskrivning av menneskelige tillitsdimensjoner til teknologi	23
3.4 Tillit til KI.....	24
3.4.1 Tidligere forskning	25
3.5 Tillit i beslutningssituasjoner.....	27
3.5.1 Tillit til KI i beslutningssituasjoner.....	28
3.6 Faktorer som svekker tillit	28
3.6.1 Algoritmeaversjon	29
3.7 Tillit til informasjonsverktøy	31
3.7.1 Viktigheten av tillit i informasjonsinnhenting.....	31
3.7.2 Redaktørstyrt innhold	33
3.7.3 Brukergenerert innhold.....	33
3.7.4 Tillit til ChatGPT, Google og Wikipedia	34
4. Hypoteser og modell	35
4.1 Hypotesesett 1: Systemtillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia	35
4.2 Hypotesesett 2: Mennesketillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia	38
4.3 Hypotesesett 3: ChatGPT sammenlignet med ChatGPT med en advarsel	40
4.4 Hypotesesett 4: Brukserfaring sin påvirkning på tillit	42
4.5 Hypotesesett 5: Generell tilbøyelighet til å stole på nye informasjonsverktøy sin påvirkning på tillit	44
4.6 Hypotesesett 6: Oppfattet grad av menneskelignende egenskaper sin påvirkning på tillit	45
4.7 Forskningsmodell	47

5. Metode.....	48
5.1 Forskningsdesign	48
5.2 Forskningsstrategi.....	48
5.3 Utforming av eksperimentet.....	50
5.4 Datainnsamling	53
5.4.1 Utvalg.....	53
5.4.2 Pilotstudie.....	54
5.4.3 Praktisk gjennomføring	54
5.5 Variabler	54
5.5.1 Uavhengig variabel: Grad av menneskelignende egenskaper.....	54
5.5.2 Uavhengig variabel: Generell tilbøyelighet til å stole på nye teknologier.....	55
5.5.3 Uavhengig variabel: Brukserfaring	55
5.5.4 Uavhengig- og interaksjonsvariabel: Gruppetilhørighet	55
5.5.5 Avhengig variabel: Tillit	55
5.5.6 Kontrollvariabler	56
5.6 Databehandling og analysering av data	58
5.6.1 Normalfordeling	59
5.6.2 Statistiske tester.....	59
5.7 Forskningsetikk og personvern.....	61
5.8 Reliabilitet og validitet	62
5.8.1 Reliabilitet	62
5.8.2 Validitet.....	63
6. Resultater.....	66
6.1 Deskriptiv statistikk.....	66
6.1.2 Alder og kjønn.....	66
6.1.3 Politisk tilhørighet	67
6.1.4 Utdanning og IT-bakgrunn.....	67
6.1.5 Klimafornektelse	67
6.1.6 Tilbøyelighet til å stole på nye informasjonsverktøy	67
6.1.7 Brukserfaring.....	67
6.1.8 ChatGPT 3.5 og ChatGPT 4.....	68
6.2 Diskusjon av hypotesesettene	69
6.2.1 Hypotesesett 1: Systemtillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia	69
6.2.2 Hypotesesett 2: Mennesketillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia	74
6.2.3 Hypotesesett 3: ChatGPT sammenlignet med ChatGPT med advarsel	77
6.2.4 Hypotesesett 4: Brukserfaring sin påvirkning på tillit.....	78
6.2.5 Hypotesesett 5: Generell tilbøyelighet til å stole på nye informasjonsverktøy sin påvirkning på tillit.....	82
6.2.6 Hypotesesett 6: Oppfattet grad av menneskelignende egenskaper sin påvirkning på tillit	84
6.2.7 Demovariabler sin påvirkning på tillit til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med advarsel.....	86
7. Oppsummerende diskusjon	89

8. Konklusjon	91
8.1 Begrensninger	93
8.2 Teoretiske implikasjoner	94
8.3 Praktiske implikasjoner	96
8.4 Anbefalinger til videre forskning.....	97
9. Litteraturliste	100
Vedlegg	124
Vedlegg: Spørreundersøkelsen	124
Vedlegg: Advarsel	135
Vedlegg fra R.....	136
Vedlegg 1: En oversikt over kjønnsfordeling i de ulike gruppene	136
Vedlegg 2: Histogram over aldersfordeling i utvalget	136
Vedlegg 3: Gjennomsnittlig generell tilbøyelighet til å stole på nye informasjonsverktøy	136
Vedlegg 4: Oversikt over brukstiden respondentene har anvendt de respektive verktøyene.....	137
Vedlegg 5: Brukshyppighet av de ulike verktøyene	137
Vedlegg 6: ANOVA-test over systemtillit mellom alle gruppene	137
Vedlegg 7: Tukey HSD-test over systemtillit mellom alle gruppene	137
Vedlegg 8: Ekvivalens-test av oppfattet <i>hjelpsomhet</i> mellom ChatGPT og Google	138
Vedlegg 9: Forskjeller i systemtillit for ChatGPT og Wikipedia.....	138
Vedlegg 10: Forskjeller i systemtillit med brukstid som uavhengig variabel	139
Vedlegg 11: Sammenhengen mellom systemtillit og brukstid for ChatGPT	139
Vedlegg 12: Forskjeller i systemtillit med brukshyppighet som uavhengig variabel	140
Vedlegg 13: Sammenhengen mellom systemtillit og brukshyppighet for ChatGPT	140
Vedlegg 14: Forskjeller i mennesketillit med brukstid som uavhengig variabel	141
Vedlegg 15: Forskjeller i systemtillit med tilbøyelighet til å stole på nye teknologier som uavhengig variabel	141
Vedlegg 16: Forskjeller i mennesketillit med tilbøyelighet til å stole på nye teknologier som uavhengig variabel	142
Vedlegg 17: Sammenhengen mellom systemtillit og oppfattet grad av menneskelignende egenskaper for ChatGPT	142
Vedlegg 18: Sammenhengen mellom mennesketillit og oppfattet grad av menneskelignende egenskaper for ChatGPT	143
Vedlegg 19: Sammenheng mellom mennesketillit og kontrollvariablene	143
Vedlegg 20: Forskjeller i systemtillit for kontrollvariablenes påvirkning for Google	144
Vedlegg 21: Sammenhengen mellom systemtillit og klimasyn for ChatGPT	144
Vedlegg 26: Korrelasjonsmatrise	145

1. Innledning

1.1 Motivasjon for oppgaven

Kunstig intelligens (KI) har vært en stor pådriver for utviklingen av den fjerde industrielle revolusjonen, hvor økt bruk av KI har ført til mer innovasjon og vekst på tvers av ulike sektorer (Bristol, et al., 2024). KI-revolusjonen forandrer ikke bare det politiske landskapet og verdensøkonomien, men også utdanning og arbeidsplasser både nå og i tiden fremover (Walsh, 2017).

Før lanseringen av ChatGPT var kjennskapen til KI relativt begrenset, og enda færre hadde tatt i bruk KI-tjenester aktivt. Med ChatGPTs inntreden i markedet har denne situasjonen endret seg dramatisk, og ført til en akselerasjon av KI-revolusjonen (Malik, 2023). Verktøyet har gjort KI til et kjent begrep for de fleste, og nevnes stadig i den offentlige debatten. Eksempler på dette er bruk av ChatGPT i utdanning (Jambak, 2024), helsesektoren (NTB, 2024), laksenæringen (Lorentzen, 2024), musikkindustrien (Kinapel, 2024), moteverden (Kopperud, 2024) og i journalistikken (Moe & Bjørgan, u.d.).

Med sin inntreden i november 2022 nådde ChatGPT én million brukere på bare fem dager (Marr, 2023). Til sammenligning tok det henholdsvis ti måneder for Facebook og mer enn tre år for Netflix å oppnå samme brukerantall (Meer, 2023). Dette illustrerer hvor revolusjonerende ChatGPT ble oppfattet av allmennheten. I dag har ChatGPT 180,5 millioner månedlige brukere (Meer, 2023), noe som understreker den raske og omfattende veksten på knappe to år.

Den store brukerbasen medfører dog et betydelig ansvar. Flere er redde for at vi bruker ChatGPT ukritisk og ikke validerer informasjonen, og det er bekymringer rundt hvorvidt brukeren er klar over at informasjonen kan være feil eller partisk (Buchanan & Hickman, 2024) (Buchanan & Hickman, 2024). Enkelte hevder til og med at ChatGPT er en lystløgner som dikter opp fakta i stor skala (Melteig, 2023). Ettersom ChatGPT er en probabilistisk modell, kan den generere feilaktige svar fordi den bygger på sannsynlighet. Dette kan føre til feilinformasjon, som potensielt kan ha store konsekvenser. For eksempel presenterte to advokater i New York oppdiktede rettsreferanser generert av ChatGPT, og ble som følge av dette tildømt sanksjoner (Merken, 2023). Dette illustrerer at misbruk av chatboten kan få alvorlige konsekvenser.

For å sikre trygg og tillitsvekkende bruk av KI lanserte EU i august 2024 den omfattende EU AI Act (Den europeiske kommisjonen, 2024). Denne lovgivningen reflekterer et økt behov for regulering og innsikt i hvordan tillit til KI bygges. Per dags dato kan det tenkes at flere millioner bruker KI uten at de nødvendigvis har høy tillit til KI-verktøy. Motsetningsvis, kan det også være at de benytter KI nettopp fordi de har stor tillit til teknologien.

Mange bruker først og fremst ChatGPT til informasjonssøk (Choudhury & Shamszare, 2023), noe som har utfordret eksisterende informasjonsplassformer som Google og Wikipedia. Dette har ført til økt konkurranse i markedet for informasjonssinnhenting, hvor Google som respons har intensivert implementering av KI i sine tjenester (Burr, 2023). Kun tre måneder etter lanseringen av ChatGPT, lanserte Google KI-chatboten Bard 21. mars 2023 (Pichai, 2023), som nå er kjent som Gemini (Google DeepMind, u.d.). Flere selskaper har slengt seg på KI-bølgen, og det eksisterer nå en rekke ulike KI-chatboter, slik som Microsofts Copilot, Claude, og Perplexity for å nevne noen.

Fremtiden til informasjonssøk er preget av usikkerhet, da det stadig lanseres nye KI-chatboter som utfordrer de etablerte aktørene. Markedet for KI-chatboter er estimert til å nå en verdi på 252 milliarder norske kroner innen 2031, med en gjennomsnittlig årlig vekstrate (CAGR) på 27,8% fra 2024 til 2031 (Meticulous Research, 2024). OpenAI ble tidligere i år verdsatt til 1660 milliarder kroner (Knudsen, 2024). Disse tallene illustrerer at markedet for KI-chatboter er i rask vekst og forventes å ekspandere betydelig i de kommende årene.

Dette raskt voksende markedet utfordrer markedet for informasjonssinnhenting, hvor etablerte aktører må tilpasse seg nye konkurranseforhold når KI-chatboter inntar markedet. Det blir derfor stadig viktigere å forstå hvordan tilliten til ulike informasjonsverktøy utvikler seg. I dagens litteratur eksisterer det lite forskning som sammenligner KI-chatboter, som ChatGPT, med etablerte aktører som Google og Wikipedia. Dette kan skyldes at ChatGPT er relativt nytt, og at det derfor vil ta tid før mer omfattende forskning på tillit til KI-chatboten blir publisert.

Så vidt vi kjenner til, er det kun én studie som har sett på tilliten mellom de ulike plattformene (Jung, et al., 2024). Jung et al. sin studie er kun basert på et lite utvalg (n=14), noe som svekker generaliserbarheten og gjør det vanskelig å trekke valide konklusjoner som kan anvendes i andre kontekster og på større populasjoner. Utover dette er forskningen på tillitsforskjeller mellom ChatGPT, Google og Wikipedia fortsatt begrenset. Andre forskere har riktignok sammenlignet de tre plattformene, men med fokus på andre aspekter enn tillit, som hvordan de formulerer seg (Luo, et al., 2024) og bibliotekarers syn på de tre verktøyene (Hurley, 2024).

Mangelen på kvantitativ forskning som utforsker tillitsnivået mellom ChatGPT, Google og Wikipedia representerer et gap i forskningen som denne oppgaven søker å dekke. Våre funn kan bidra til en bedre forståelse av hvilket verktøy individer har høyest tillit til, samt hvilke faktorer som påvirker denne tilliten. Dette er viktig å få innsyn i, da tillit er avgjørende for hvilke verktøy brukerne velger å benytte seg av (Nicolaou & McKnight, 2006). I en tid preget av utfordringer som feilinformasjon og spredning av falske nyheter (Kumar & Shah, 2018), er det essensielt å få et øyeblikksbilde av tillit til informasjonsverktøy. Videre er det relevant å kartlegge tillitsnivået til ChatGPT spesielt, ettersom verktøyet kan generere feilinformasjon, og et høyt tillitsnivå kan indikere at brukerne stoler ukritisk på plattformen.

ChatGPT er som nevnt probabilistisk og kan derfor begå feil. Plattformen fremhever dette selv gjennom advarselen «*ChatGPT kan gjøre feil. Sjekk viktig informasjon*», som vises nederst i grensesnittet. Vi finner derimot ingen tidligere forskning som undersøker om denne advarselen påvirker brukernes tillit til tjenesten. Dette representerer et ytterligere gap i forskningen som det er viktig å dekke for å forstå hvorvidt åpenhet om feil påvirker tillit, eller om advarselen i praksis har liten betydning for brukerens atferd på plattformen.

Med dette som bakteppe har denne studien som formål å undersøke tillit til ChatGPT sammenlignet med Google og Wikipedia, samt å vurdere om tilliten til ChatGPT påvirkes av en advarsel om at verktøyet selv fremhever at den kan ta feil. Dette leder oss til oppgavens forskningsspørsmål:

Hvordan er tillit til ChatGPT sammenlignet med Google og Wikipedia, og vil tilliten påvirkes av en advarsel om at ChatGPT selv informerer om at den kan ta feil?

Svaret på forskningsspørsmålet har praktisk betydning for alle situasjoner hvor mennesker søker å innhente informasjon. Det vil bidra til å belyse om mennesker stoler mer på en KI-chatbot enn på tradisjonelle informasjonsverktøy som Google og Wikipedia. Dette kan være særlig verdifullt for ledere, som kan bruke innsikten til å vurdere hvordan tillit til KI-verktøy påvirker deres organisasjon. Basert på våre resultater kan ledere eksempelvis velge å øke eller redusere implementeringen av KI-verktøy i informasjonsarbeid, avhengig av hvorvidt tilliten til slike verktøy vurderes som høy eller lav.

1.2 Avgrensninger

I vår oppgave har vi valgt å fokusere på KI-chatboten, ChatGPT. Dette er fordi ChatGPT var en av de første på markedet, og har den største markedsandelen på nærmere 60% av KI-chatbot-markedet (Bailyn, 2024). Det er en rekke andre KI-chatboter vi kunne inkludert i vår studie, slik som Microsoft Copilot eller Google Gemini, men på grunn av omfanget på oppgaven har vi valgt å avgrense til å kun fokusere på ChatGPT, som er den mest etablerte og brukte KI-chatboten.

Grunnen til at vi ikke sammenligner ChatGPT med Gemini er at vi undersøker markedet for informasjonssøk og ikke markedet for KI-chatboter. Selv om disse markedene kan overlappe, har vi valgt å fokusere på Google-søk fremfor Gemini, ettersom Google per dags dato er en av de mest veletablerte aktørene innen informasjonssøk. Google har også en rekke andre tjenester, som Google Maps og Google Drive. Vi avgrenser oppgaven til å se på Google som søkemotor. Når vi omtaler Google i oppgaven, refererer vi derfor utelukkende til Google-søk.

En annen avgrensning er at vi behandler ChatGPT som én samlet enhet. Når vi omtaler ChatGPT gjelder derfor dette både gratisversjonen ChatGPT 3.5 og den betalte versjonen ChatGPT 4. Dette valget er begrunnet i studiens formål om å undersøke tillit til ChatGPT som informasjonsverktøy, uavhengig av spesifikke funksjonsforskjeller.

1.3 Disposisjon

I neste del av oppgaven skal vi kort forklare hva ChatGPT, Google og Wikipedia er. Deretter skal vi presentere tillitsbegrepet, og relevant faglitteratur som omhandler tillit. Videre skal vi gjennomgå vår forskningsmetode, som er eksperiment, og begrunne dette metodevalget. Vi presenterer selve utformingen og gjennomføringen av eksperimentet, før vi viser funnene våre og kommer med en analyse av resultatene. Avslutningsvis vil vi komme med en konklusjon, samt trekke frem begrensninger ved egen oppgave, praktiske og teoretiske implikasjoner, og komme med oppfordring til videre forskning.

2. Bakgrunn

I den følgende seksjonen skal vi forklare hva ChatGPT er. For å få en dypere innsikt og forståelse av denne applikasjonen er det nødvendig med en kort introduksjon til kunstig intelligens, generativ KI og store språkmodeller. Dette fordi ChatGPT er bygget på en stor språkmodell innenfor det generative KI-feltet. Deretter skal vi presentere hvordan Google og

Wikipedia som informasjonsverktøy konkurrerer mot ChatGPT, for slik å rettferdiggjøre, samt begrunne relevansen, av valget om å sammenligne ChatGPT med disse to verktøyene.

2.1 Kunstig intelligens

Kunstig intelligens er ikke et nytt fagfelt, men har tvert imot eksistert helt siden 1956, da John McCarthy kom opp med terminologien «artificial intelligence (AI)». Denne terminologien er kjent som kunstig intelligens på norsk (Stryker & Kavlakoglu, 2024). Gjennomgående i masteroppgaven skal vi forholde oss til den norske terminologien.

Kunstig intelligens har til hensikt å skape maskiner som etterligner menneskelig adferd og intelligens (Strümke, 2023, s.25) med bruk av teknologier som anvender algoritmer, maskinlæring, probabilistiske modeller og store datasett (UiB, 2023).

2.2 Generativ KI og ChatGPT

Generativ KI, best kjent som «GenAI» på engelsk, er en kategori under KI som omfatter modeller som produserer nytt innhold. Produksjon av nytt innhold innebærer at vi ikke har sett innholdet tidligere (Rigby, 2024). Det kan være nytt innhold i form av bilder, koder, lyd, tekst og videoer (McKinsey & Company, 2024). ChatGPT er et eksempel på et generativ KI-verktøy fordi den produserer tekst som ikke er blitt skrevet noe sted tidligere og kan generere bilder som ikke finnes fra før.

ChatGPT har vært avgjørende for å løfte generativ KI frem i offentlighetens søkelys, sannsynligvis på grunn av dens evne til å kommunisere på en menneskelig måte (Jackson, 2023). Fra å være et nytt eksperiment innen KI-språkprosessering, har KI-chatboten utviklet seg til å bli et verktøy som er utbredt i flere sektorer og anvendt av millioner av mennesker. ChatGPT er et komplekst verktøy og det er vanskelig for mennesker å spore hvor informasjonen kommer fra. Dette er fordi applikasjonen er bygget på store språkmodeller, også kjent som «Large Language Models (LLMs)» på engelsk.

2.3 Store språkmodeller

Store språkmodellerer en viktig grunnsten bak de nye teknologiene vi har fått innenfor genererende KI. En LLMs er en type nøytralt nettverk, som har blitt trent på store kvantum av umerket tekst (Rigby, 2024). I praksis innebærer det at LLMs har blitt trent av store deler av data fra internettet (OpenAI, 2024). OpenAI sine LLMs-modeller er utviklet ved å bruke tre hovedkilder til informasjon: (1) offentlig informasjon tilgjengelig på internett, (2) informasjon lisensiert fra tredjeparter, og (3) informasjon som brukerne og trenere av ChatGPT har gitt

(OpenAI, 2024). Modellen lagrer ikke kopier av informasjon som den lærer av, og er å regne som stor fordi den gjør en massiv mengde kalkulasjoner for hver prediksjon den gjør. Videre er det en språkmodell fordi den predikerer sekvenser av ord (Rigby, 2024).

Den store språkmodellen bak gratisversjonen av ChatGPT er per i dag (desember 2024) GPT 3.5, mens den betalte versjonen av ChatGPT bygger på GPT 4. GPT 4 er den nyeste generasjonen av i OpenAI sin familie av LLMs kalt GPT. Denne inkluderer endringer som gjør at modellen har lavere sannsynlighet for å generere støtende eller farlige resultater, har evnen til å søke på internett i sanntid, bedre styring og kontrollerbarhet (Marr, 2023).

En sentral forskjell mellom GPT 3.5 og GPT 4 er mengden parametere. GPT 4 har 1,76 billioner, noe som er nesten 100 ganger så mye som GPT 3 med 175 milliarder. Parametere fungerer som neuronene i hjernen ved at de behandler informasjon (Microsoft Azure, 2024). På samme måte som det er komplekst å forstå hvorfor hjernen beslutter som den gjør, gjelder dette også for tjenester som bygger på store språkmodeller. Det er dette som refereres til som en «black-box-modell». Det innebærer at vi legger inn informasjon i den ene enden, og at det deretter kommer et svar ut i den andre enden, uten at vi mennesker forstår hvordan utdataen eller resultatet ble oppnådd (Tzur, 2024; Seon, u.d.).

2.4 ChatGPT som chatbot-applikasjon

ChatGPT er en virtuell KI-chatbot (OpenAI, 2024). Virtuell-KI er en type kunstig intelligens som kan forstå og svare på både skriftlige og muntlige spørsmål fra brukerne (McCarthy, u.d.). Brukeren kan blant annet anvende ChatGPT til å organisere, oppsummere eller produsere ny tekst (OpenAI, 2024). Tjenesten er utviklet slik at maskinen skal forstå og svare på spørsmål og instruksjoner fra brukeren. ChatGPT gjør dette ved å behandle en stor mengde eksisterende tekst og lære hvordan ord hører sammen med andre ord. Deretter bruker ChatGPT denne lærdommen til å forutsi det neste mest sannsynlige ordet eller sekvensen, som vises som svar på spørsmålet eller instruksjonen fra brukeren (OpenAI, 2024). På denne måten er ChatGPT probabilistisk, noe som betyr at den bygger på sannsynlighet og at det er usikkerhet involvert i svarene den gir (Haraldsen, 2021). Dette er trolig en av grunnene til at OpenAI har valgt å inkludere følgende advarsel på plattformen: «*ChatGPT kan gjøre feil. Sjekk viktig informasjon*».

2.5 Informasjonsverktøy

Brukere kan reagere på informasjon ulikt, basert på hvilken kilde som er opprinnelsen til innholdet (Jung, et al., 2024). Vår studie fokuserer på tre ulike kilder til

informasjonsinnhenting: Wikipedia, Google og ChatGPT. Den førstnevnte er bruker-drevet, i motsetning til de to sistnevnte som drives av algoritmer. Googles søkemotor er en anerkjent informasjonskilde som er kjent for å ha algoritmer som prioriterer relevante resultater for sluttbrukeren (Google, u.d.). Wikipedia er på sin side et informasjonsverktøy som baseres på at mennesker redigerer og publiserer innhold (Ayers, et al., 2008). Det er altså bidrag fra mange ulike mennesker som er grunnlaget for informasjonen som publiseres på Wikipedia. ChatGPT er derimot en algoritmebasert kilde som autonomt genererer tekst og innhold, og mennesker er ikke involvert i hverken produksjon eller redigering av innhold (Jung, et al., 2024).

ChatGPT, Google og Wikipedia kan alle anees som verktøy for informasjonsinnhenting. Selv om ChatGPT er utviklet for å generere kreativt innhold, er det også oppfattet som et informasjonsverktøy av brukerne (Jung, et al., 2024). I en studie med 607 respondenter, ble det avdekket at ChatGPT hovedsakelig benyttes for å finne informasjon (Choudhury & Shamszare, 2023).

En annen studie fant antydninger til at mennesker i større grad velger ChatGPT som deres førstevalg når de ønsker svar på spørsmål i situasjoner der de tidligere ville benyttet et Google-søk (Tutella, 2024). Google har lenge vært den dominerende søkemotoren for informasjonsinnhenting, men mye tyder på at plattformen i større grad enn tidligere mottar reell konkurranse, og at dette er fra ChatGPT. Likeså har Wikipedia lenge vært et av de mest brukte informasjonsverktøyene, spesielt for studenter, verden over (Selwyn & Gorard, 2016). Nå utfordres også denne plattformen av ChatGPT, som i økende grad anvendes i utdanningssektoren (Lo, 2023).

3. Litteratur og forskning

I denne oppgaven ser vi det som hensiktsmessig å gi en utdypende forklaring på tillitsbegrepet, og hvilke faktorer som kan påvirke individers tillit til teknologier, KI og informasjonsverktøy. Dette teoretiske bakteppe danner grunnlaget for vårt eksperiment og videre analyser og diskusjon.

Først skal vi gi en dypere innsikt i hva tillit egentlig er, samt hvilke tillitsdimensjoner begrepet kan deles inn i. Her skiller vi mellom de tradisjonelle og etablerte mennesketillitsdimensjonene *evne, velvilje og integritet*, og de nyere systemtillitsdimensjonene *funksjonalitet, hjelpsomhet og reliabilitet*.

3.1 Tillit

I denne seksjonen skal vi gå nærmere inn på hva tillit er, samt gi en grundigere forklaring på de tradisjonelle og etablerte tillitsdimensjonene *evne*, *velvilje* og *integritet*. Vi skal også vise til hvordan disse dimensjonene har blitt studert i andre settinger, og hvordan de knyttes opp mot det overordnede tillitsbegrepet.

Det finnes en rekke ulike definisjoner på hva tillit er (Cahour & Forzy , 2009; Hardin, 2006; Gambetta, 2000; Seligman, 1997), og begrepet kan være vanskelig å forklare (Simpson , 2012), noe som er naturlig ettersom tillit er et kompleks og bredt begrep. Likevel virker det å være en enighet om at tillit kan defineres som «en psykologisk tilstand som innebærer en intensjon om å akseptere sårbarhet basert på positive forventninger til tillitsobjektets intensjoner eller atferd» (Rousseau, et al., 1998). Et individ har altså tillit til et annet objekt dersom individet selv oppfatter og stoler på at objektet handler ut fra individets beste, uten at en har mulighet til å kontrollere det. For eksempel: Hvis du forteller noen en hemmelighet i fortrolighet, har du tillit til at vedkommende (objektet) ikke deler hemmeligheten videre. Et annet eksempel er når du deler informasjon på nett i god tro om at dette ikke deles videre.

Tillit er med på å predikere individers handlinger, og er derfor essensielt å kartlegge da tillitsnivå kan påvirke handlingsmønstre. Studier har vist at tillit blant annet er med på å avgjøre om en bruker gjennomfører kjøp på nett og deler personlig informasjon (Bhattacharjee, 2002) og om individer tar i bruk telefonbanktjenester og nettbank (Dimitriadis & Kyrezi, 2010). Fra et bedriftsperspektiv har også tillit vist seg å være viktig for kjøpsintensjoner, lojalitet, betalingsvillighet og bedriftens prestasjon (Khamitov, et al., 2024). Tillit bidrar også at et samfunn kan investere mer kapital og gjennomføre transaksjoner mer effektivt (Evans & Krueger, 2009). Overordnet er tillit sentralt å kartlegge da det er en viktig del av de fleste økonomiske og sosiale interaksjoner hvor det er usikkerhet (Pavlou, 2016).

Det er mulig å skille mellom tre hovedtyper av tillitsbegreper: (1) tillitsfulle overbevisninger, (2) tillitsfulle intensjoner og (3) tillitsfull adferd (McKnight, 2005). Disse begrepene omhandler ulike stadier av tillit og henger sammen. (1) Tillitsfull overbevisning omhandler at individet er overbevist om at tillitsobjektet har egenskapene *evne*, *velvilje* og *integritet*. (2) Tillitsfulle intensjoner handler om at individet har en intensjon om å gjøre seg avhengig av og bli sårbar overfor tillitsobjektet. (3) Tillitsfull adferd er å demonstrere gjennom handlinger en vilje til å gjøre seg avhengig av og sårbar overfor tillitsobjektet (McKnight, 2005). Når et individ oppfatter at et tillitsobjekt har tilstrekkelig *evne*, *velvilje* og *integritet*, kan dette skape

en intensjon om å stole på objektet. Denne tillitsfulle intensjonen kan føre til at individet, gjennom sine handlinger, gjør seg både sårbar og avhengig av tillitsobjektet.

I vår studie vil førstnevnte, (1) tillitsfull overbevisning være mest aktuell, da vi kommer til å måle individers oppfatning av ulike informasjonsverktøy, og ikke måle faktisk intensjon eller adferd.

3.1.2 De tradisjonelle dimensjonene av tillitsfull overbevisning

For å forstå i hvilken grad et individ stoler på ChatGPT, Google og Wikipedia, er det sentralt å avdekke hvilke faktorer som påvirker tillit. Mayer, Davis og Schoormans modell (1995) er et solid utgangspunkt, ettersom den er anerkjent i akademien og tar for seg både faktorene som påvirker tillit og konsekvensene av tillit. Modellen er også unik da den beskriver ulike tillitsdimensjoner og hvordan de konkret påvirker partene i tillitsforholdet.

Modellen deler tillit inn i tre dimensjoner: *evne*, *velvilje* og *integritet* (Mayer, et al., 1995). Det argumenteres for at alle dimensjonene er viktige for tillit, og at det er essensielt å skille mellom hver av de enkelte dimensjonene, da de varierer uavhengig av hverandre. I den videre delen av teoriseksjonen skal vi derfor ta for oss alle tre komponentene hver for seg.

Evne

Evne er ferdighetene, kompetansen og egenskapene som gjør at en part har innflytelse innenfor et spesifikt felt eller område (Mayer, et al., 1995). Dersom en part er kompetent på et område, kan den ifølge modellen få tillit til oppgaver tilknyttet det spesifikke området. Parten kan derimot være lite egnet til å utføre oppgaver på et annet felt, og dermed få mindre tillit i andre situasjoner. For eksempel er det naturlig at man har høy tillit til at en frisør kommer til å utføre en god jobb når man skal klippe seg, men at man har lav tillit til at frisøren kan fikse et elektrisk anlegg. I vår studie kan dette knyttes opp mot at brukeren har høy tillit til informasjonsverktøyet, dersom brukeren oppfatter at det har tilstrekkelig evne til å finne og gi informasjonen brukeren søker om et spesifikt tema.

I Mayer et al. (1995) sin modell er *evne* oppgave- og situasjonsbestemt, noe som betyr at en parts evner avhenger av oppgaven som skal gjennomføres, og i hvilken situasjon. For eksempel kan det tenkes at en bruker har høy tillit til at ChatGPT kan hjelpe til med å generere tekst og oversette språk, men lav tillit til at chatboten kan diagnostisere om du har en sykdom. Dette gjenspeiler at kontekst og oppgave er viktig for hvorvidt et individ har tillit til et objekt. Videre spiller individets oppfatning av hvor risikofyllt en situasjon er en viktig rolle for tillit (Kim, et

al., 2008). Eksempelvis når det gjelder diagnostisering av sykdommer, som har høy grad av risiko dersom diagnostiseringen er feil.

Velvilje

Velvilje handler om at en part ønsker å gjøre noe godt for motparten uten egosentriske profittmotiv (Mayer, et al., 1995). Denne faktoren av tillit bygger på at partene har en relasjon, og velvilje kommer til syne dersom den ene vil hjelpe den andre selv om det ikke innebærer en ytre belønning. Velvilje kan knyttes opp mot altruisme, empati og omsorg for andre, og er vist å være grunnlaget for tillit i flere studier (Larzelere & Huston, 1980; Leonard, 1960). Et eksempel på at noen viser velvilje kan være lærere som utdanner seg med intensjonen om å hjelpe elever, fordi de bryr seg om elevenes suksess og læringskurve. Dette reflekterer lærerne som har gode intensjoner og ønsker det beste for elevene. Disse lærerne kan oppfattes som mer tillitsvekkende av elevene.

I vår studie kan man forstå denne tillitsfaktoren som hvorvidt brukeren har tillit til at de ulike informasjonsverktøyene vil brukeren vel. For eksempel kan brukerens tillit påvirkes av om brukeren mener Google primært er opptatt av å selge brukerdata eller manipulere brukernes internett-trafikk og adferdsmønster for egen vinning. Dette viser lav velvilje. Motsetningsvis, kan brukerens oppfatning av Google som velvillig være høy dersom brukeren mener Google prioriterer å vise de mest relevante nettsidene ved et Google-søk.

Integritet

Integritet handler om at en part oppfatter at den andre følger et sett med regler som er akseptable (Mayer, et al., 1995). Integritet kan defineres som «en person eller institusjons evne og vilje til å handle selvstendig, ærlig og redelig uten å ta hensyn til uvedkommende, utenforliggende interesser» (NAOB, u.d.). Faktorer som påvirker grad av integritet er tidligere atferd og handlinger, hvorvidt det gis troverdig informasjon om parten fra andre kilder, grad av oppfattet rettferdighetssans, og om partens handlinger sammenfaller med deres uttalelser. Integritet kan kobles opp mot ærlighet, konsistens og moralske prinsipper. For eksempel kan en politiker love å gjennomføre en viktig reform for å vinne stemmer, men aldri ha intensjoner om å følge opp løftet. Dette er et eksempel på lav integritet fordi politikeren bevisst bryter tilliten til velgerne, fordi handlingene ikke sammenfaller med uttalelsen.

I vår studie avhenger denne tillitsfaktoren av om brukeren oppfatter at aktøren tilbyr troverdig og nøytral informasjon. Dette påvirkes av om brukeren synes aktøren kommer med informasjon som fremstår som ærlig og faktabasert, og at det ikke er skjevheter i svarene. Med skjevheter i

svarene viser vi til at aktøren gir partisk informasjon, for eksempel at svarene heller mot én politisk retning. For eksempel kan brukeren mene at ChatGPT har lav integritet dersom plattformen svarer på et medisinsk spørsmål med stor selvsikkerhet, som om den er en ekspert, men samtidig gir unøyaktige eller potensielt skadelige råd. Dersom ChatGPT ikke tydeliggjør at den ikke er en lege eller autorisert medisinsk kilde, kan dette føre til misforstått tillit hos brukeren.

3.1.3 Tidligere forskning

Det har vært forsket mye på de tre tradisjonelle dimensjonene av tillit: *evne*, *velvilje* og *integritet*. Tidligere forskning har i stor grad benyttet de tre tillitsdimensjonene for å måle relasjoner mellom mennesker, og ikke mellom mennesker og teknologi.

Tillitsdimensjonene har primært blitt anvendt i forskning som undersøker tillitsforholdet mellom ledere og ansatte. Forskere har benyttet dimensjonene for å studere tillit til ledere, ansatte og kolleger (Knoll & Gill, 2011; Tan & Lim, 2009). Tilliten arbeidstakere har til arbeidsgiver har vist seg å være sterkt korrelert med alle de tre tillitsdimensjonene (Tan & Tan, 2000). Det er vist at det er sterkere relasjon mellom oppfattelsen en ansatt har til sin overordnedes *evne*, *velvilje* og *integritet*, enn tilliten de ansatte har til selve organisasjonen. Forskning viser at de tre dimensjonene direkte påvirker ansattes tillit til sin overordnet (Poon, 2013), og at jo mer sårbar en ansatt er, jo viktigere er det for lederen å sørge for at egen atferd reflekterer integritet og evne (Lapidot, et al., 2007).

Videre er det funnet at de ulike tillitsdimensjonene påvirker nettverksinteraksjon, det vil si kommunikasjon i individets sosiale nettverk (Munson & Macri, 2009). Velvilje er spesielt relevant for å fremme åpen og ærlig kommunikasjon og kunnskapsdeling i nettverket (Svare, et al., 2019). Svare et al. (2019) finner også at tillit basert på velvilje legger til rette for bedre samarbeid og mer innovasjon.

Som vist ovenfor, er det omfattende forskning på tillit mellom mennesker. Det er en rekke andre studier som også har forsket på dette tidligere (Freitag & Bauer, 2016; Katz & Rotter, 1969; Chandra, et al., 2018; Rolfe, et al., 2014). Selv om enkelte studier undersøker hvordan de tre tillitsdimensjonene påvirker tillit til teknologi (Madhavan & Wiegmann, 2007; Basu & Singhal, 2016; Baker, et al., 2018; Jones & Marsh, 1997), er det relativt lite forskning på individers tillit til KI sammenlignet med tradisjonelle teknologier. Dette reflekterer et gap i forskningen, og understøtter at det er et behov for å forske på om tillitsbegrepene kan anvendes i andre situasjoner enn mellommenneskelige relasjoner. Eksempler på dette er forskning som

kartlegger om tillitsdimensjonene gir utslag i menneske-maskin-relasjoner og menneske-KI-relasjoner.

3.2 Tillit til teknologi

I denne seksjonen skal vi gå nærmere inn på tillit til teknologi. Innledningsvis skal vi forklare hva tillit til teknologi predikerer, før vi kort presenterer «Human–computer interaction» (HCI). Deretter skal vi gå nærmere inn på to variabler som påvirker tillit til teknologi, før vi gir en grundig forklaring på de nyere tillitsdimensjonene *funksjonalitet*, *hjelpsomhet* og *reliabilitet*. Vi skal også vise hvordan disse dimensjonene har blitt studert tidligere, og hvordan de knyttes opp mot tillitsbegrepet.

For å undersøke brukeres tillit til ChatGPT sammenlignet med Google og Wikipedia er det sentralt å utforske teori om tillit til teknologi, da de nevnte informasjonskildene er teknologiske verktøy. Dette er relevant fordi tillit til teknologi er avgjørende for om en bruker faktisk kommer til å anvende teknologien (Korzyński, et al., 2024). Grad av tillit kan predikere hvor avhengig en bruker er av en teknologi, og samsvaret mellom brukers tillit og teknologiens evner kan påvirke resultatene av hvordan teknologien brukes (Glikson & Woolley, 2020).

Tillit er videre viktig å måle, da dette er en av hovedfaktorene som er med på å påvirke om et individ velger å benytte seg av nye teknologier, som for eksempel selvstyrte biler (Choi & Ji, 2015) og anbefalingsagenter, det vil si systemer eller tjenester som hjelper internettbrukere med å finne, evaluere og velge produkter basert på deres behov og preferanser (Komiak & Benbasat, 2006)). Andre eksempler er adopsjon av digitale forvaltningstjenester fra myndighetene (Warkentin, et al., 2010), eller om individer benytter seg av netthandel (Van Slyke, et al., 2004). Til tross for at noen forskere har tvilt på om individer i det hele tatt kan stole på en teknologi (Howe, 2000), viser disse studiene at tillit er med på å forklare hvorfor et individ velger å ta i bruk ulike teknologiske tjenester.

3.2.1 HCI

Tillit kan påvirke interaksjonen mellom menneske og maskin, og det er dermed sentralt å ha en forståelse av HCI-feltet (Hancock, et al., 2011). «Human–computer interaction» (HCI) er et bredt forskningsfelt som har utviklet seg raskt over de siste tre tiårene (Carroll, 2005). HCI omfatter en grundig utforskning av vitenskapelige implikasjoner og praktiske aspekt tilknyttet samspillet mellom mennesker og datamaskiner, eller mennesker og intelligente systemer (Ren & Yanwei, 2019). Feltet gir altså en innføring i interaksjonen mellom mennesker og intelligente

systemer, slik som KI. I disse interaksjonene anvender mennesker flere evner, blant annet lesing, kommunikasjon, fysisk berøring og skriving (MacKenzie, 2024).

De viktigste begrepene innenfor HCI-feltet er *funksjonalitet* og brukervennlighet, hvor *funksjonalitet* referer til tjenestene et system tilbyr og brukervennlighet handler om hvorvidt brukeren klarer å anvende systemets funksjoner uten problemer (Mathew, et al., 2011). Både *funksjonalitet* og brukervennlighet kan variere mellom ulike systemer. Det er vist at teknologiske forbedringer, slik som for eksempel flere funksjoner eller økt brukervennlighet, er med på å øke brukerens HCI-opplevelse (Welsh, et al., 2009).

3.2.2 Tilbøyelighet til å stole på teknologier

Tilbøyelighet til å stole på andre, enten mennesker eller objekter, er et trekk ved menneskers personlighet som kan forklare hvorfor noen mennesker stoler mindre på teknologiske objekter enn andre (Gefen, et al., 2011). Tilbøyelighet til å stole på andre refererer til en individs villighet til å ha tillit til teknologier, selv før noen form for interaksjon har funnet sted. Denne mekanismen er særlig viktig i begynnelsen av relasjoner, hvor tillit ofte er mer basert på en generell tilbøyelighet til å stole på den andre, enn på faktiske erfaringer. Etter hvert som relasjonen utvikler seg, vil erfaringene og interaksjonene mellom partene få større betydning for tillitsnivået (Gefen, et al., 2011).

Forskning har vist at tilbøyelighet til tillit kan være genetisk (Evans & Krueger, 2009; Jang, et al., 1998). I en studie av eneggede og toeggede tvillinger fant Cesarin et al. at gener påvirker beslutninger i det såkalte tillitsspillet (2008). Felles gener viste seg å ha større betydning for tilbøyelighet til tillit, enn å vokse opp i samme miljø. Koblet opp mot teknologi, kan dette indikere at mennesker med samme gener, og som har lav tilbøyelighet til tillit, har lavere tillit til teknologi generelt.

3.2.3 Brukserfaring

Hvor lenge et individ har benyttet en teknologi kan påvirke tilliten individet har til teknologien, og denne brukserfaringen er med på å påvirke fremtidig bruk av verktøyet (Hernández-Ortega, 2011). Første gangen man benytter en teknologi, er det naturlig at man danner seg en mening om teknologien. Denne meningsoppfatningen kan påvirke videre bruk og tillit til teknologien, og potensielt føre til en fortsettelse av bruk dersom individet stoler på teknologien (Karahanna, et al., 2003). På den andre siden kan meningsoppfatningen være i strid med tidligere syn eller tro tilknyttet teknologiens egenskap eller prestasjon, noe som reduserer tillitsnivå og bruk av

teknologien. Overordnet gir dette en indikasjon på at brukserfaring kan ha en effekt på tillit, og at dette forholdet avhenger av individets oppfatning av verktøyet.

Individets oppfatning av teknologien og kvaliteten på den informasjonen som gis, er avgjørende for tillit, ettersom dette påvirker brukerens totale erfaring med verktøyet (Nicolaou & McKnight, 2006). I tillegg kan brukshyppighet, altså hvor hyppig et individ benytter et verktøy eller et system, ha påvirkning på tillit. En studie som utforsket tillit til nettbankbruk, fant at jo sjeldnere brukerne benyttet banken, jo høyere tillit hadde de (Aldas-Manzano, et al., 2011). Dette illustrerer at sjeldnere bruk av tjenester kan føre til høyere tillit, noe som kan være fordi brukeren har færre muligheter til å oppleve problemer som kan redusere tillit. Brukere som ikke like ofte bruker en tjeneste, kan ha en mer idealisert oppfatning av hvor pålitelig den er. Med dette mener vi at brukerne, på grunn av færre interaksjoner med tjenesten, er mer tilbøyelige til å oppfatte den som pålitelig sammenlignet med brukeren som har hyppigere interaksjoner.

3.2.4 De nye dimensjonene av tillitsfull overbevisning

Som tidligere nevnt har vi de tre hovedtypene tillitsbegreper: (1) tillitsfulle overbevisninger, (2) tillitsfulle intensjoner og (3) tillitsfull adferd (McKnight, 2005). Disse begrepene kan også brukes om tillit til informasjonsteknologi (IT). Den eneste forskjellen er tillitsobjektet – en teknologi i stedet for et menneske. Man kan for eksempel stole på at GPS-en til viser riktig vei på samme måte som man kan stole på at et menneske kan vise veien (McKnight, 2005).

(1) Tillitsfulle overbevisninger omhandler at individet er overbevist om at teknologien har fordelaktige egenskaper som er sterke nok til å skape tillitsfulle intensjoner. Dette kan være at teknologien er *funksjonell, hjelpsom og pålitelig*. (2) Individet har tillitsfulle intensjoner hvis det har en intensjon om å stole på og gjøre seg sårbar overfor teknologien. (2) Tillitsfulle intensjoner kan lede til (3) tillitsfull adferd hvis individets vilje til å gjøre seg sårbar overfor teknologien er sterk nok. (3) Tillitsfull adferd omhandler at individet gjennom handling demonstrerer at det stoler på eller er gjør seg avhengig av teknologien, i stedet for å forsøke og kontrollere den. Et eksempel på dette er en brukers vilje til å være sårbar overfor ChatGPT, for eksempel ved å dele sensitiv personlig informasjon på plattformen eller stole på dens svar for å ta beslutninger uten å verifisere informasjonen. Som nevnt innledningsvis kommer vårt fokus hovedsakelig til å omhandle (1) tillitsfulle overbevisninger.

Tillit til teknologi skiller seg fra tillit til mennesker fordi forholdet handler om menneske-til-teknologi tillit fremfor menneske-til-menneske tillit (Lankton, et al., 2015). Historisk sett har

det vært normalt å kategorisere teknologi som et objekt og ikke noe menneskelig (Lankton, et al., 2015). Likevel har mange forskere tidligere målt tillit til teknologi på samme måte som de har målt tillit til mennesker. Disse forskerne har benyttet de nevnte menneskelige tillitsdimensjonene *evne*, *velvilje* og *integritet* (Vance, et al., 2008; Benbasat & Wang, 2005). En av grunnene til at forskerne kan ha valgt å benytte menneskelige tillitsdimensjoner når de måler tillit til teknologi er fordi mennesker har en tendens til å tildele teknologier menneskelige egenskaper og antropomorforisere de (Reeves & Nass, 2003). Antropomorfisme handler om å overføre menneskelige egenskaper til ikke-mennesker, som i dette tilfellet til en teknologi (Løø, 2019).

Andre forskere har valgt å måle tillit til teknologi gjennom å bruke tillitsdimensjonene som er tilpasset systemer: *funksjonalitet*, *hjelpsomhet* og *pålitelighet* (McKnight, et al., 2011). De systembaserte tillitsdimensjonene er nyere og korresponderer med de menneskelige tillitsdimensjonene (Lankton, et al., 2015). *Funksjonalitet* er avledet fra *evne*, *hjelpsomhet* fra *velvilje* og *pålitelighet* fra *integritet*.

Funksjonalitet er troen på at teknologien har de nødvendige egenskapene som behøves for å utføre en oppgave, på lik linje som et menneske har de nødvendige evnene til å fullføre en oppgave. *Hjelpsomhet* er troen på at teknologien gir tilstrekkelig og rask hjelp, samme som et menneske viser velvilje gjennom sine handlinger overfor andre. *Pålitelighet* er troen på at teknologien fungerer som forventet, slik som at mennesker har integritet om de oppfører seg som forventet overfor andre (McKnight, et al., 2011). Lav pålitelighet kan redusere tillit (Hoff & Bashir, 2015).

Dagens litteratur er ikke tydelig på når forskere bør benytte de menneskelige eller systembaserte tillitsdimensjonene. Dette er kritisk, ettersom valget mellom tillitsdimensjonene kan utgjøre avgjørende empiriske forskjeller i studier (Lankton, et al., 2015). Individuer oppfatter grad av menneskelighet tilknyttet ulike teknologier forskjellig, og påvirkningen av tillit til teknologien avhenger av hvor menneskelig teknologien er (Lankton, et al., 2015).

Lankton, McKnight og Tripp (2015) mener fremtidige studier bør være oppmerksomme på at teknologier kan variere i menneskelighet og at forskere som studerer tillit derfor bør være bevisst på at dette kan påvirke tillitsoverbevisning. Lankton et al. (2015) har funnet at én tillitstype, enten menneskelig eller systembasert, ikke passer for alle respondenter da de har ulike oppfatninger av en teknologi sin menneskelighet. Forskerne argumenterer for at de menneskelige tillitsdimensjonene vil ha sterkere effekt for menneskeliggende teknologier, og

de systembaserte tillitsdimensjonene vil ha en sterkere innflytelse på systemlignende teknologier.

3.3 Tilskrivning av menneskelige tillitsdimensjoner til teknologi

Mennesker og teknologier har både likheter og ulikheter. Når det gjelder likheter har begge egenskapen *evne* i en eller annen form med varierende kvalitet. Å ha tillit til en persons *evne* innebærer å stole på at personen har kompetansen til å utføre en oppgave eller oppfylle et ansvar. Å ha tillit til en teknologi sin *evne* innebærer å stole på at en teknologi har den *funksjonaliteten* som kreves for å utføre oppgaven (McKnight, 2005).

Når det gjelder ulikheter, kan mennesker ha egenskaper som *velvilje* og *integritet*. Imidlertid er disse egenskapene ved tillit vanskelig å overføre til teknologier. Å tilegne en teknologi slike menneskelige egenskaper er en form for antropomorfisme. Ifølge McKnight kan man ikke tilskrive egenskaper som *velvilje* og *integritet* til IT-systemer – kun *funksjonalitet* og *pålitelighet* (2005). Av den grunn kan man ikke si at teknologi har evnen til å bry seg (vise *velvilje*) eller fortelle sannheten (vise *integritet*). Etersom McKnight (2005) argumenterer for at teknologier mangler moralske evner, reflekterer tillit til teknologi en tro på teknologiens funksjonaliteter og evner, mer enn en tro på *velvilje* og moralske motiver.

Forskere har ulike tilnærminger, og det en spredning i forskningsfeltet rundt hvorvidt man kan tilskrive teknologier menneskelige egenskaper. Til tross for at McKnight (2005) hevder at *velvilje* og *integritet* er egenskaper som ikke kan tilskrives en teknologi, har Bedué og Fritzsche på sin side funnet flere underkategorier til egenskapene *evne*, *velvilje* og *integritet* som kan tilknyttes en teknologi (2021). Underkategoriene for *evne* var: tillit til kunnskap, transparens (åpenhet), forklarbarhet, datakvalitet og *pålitelighet*. For *velvilje* peker studien på underkategorier som sosial ansvarlighet og etisk adferd. Underkategoriene for *integritet* er: standarder og retningslinjer, sertifiseringer og reguleringer fra myndighetene (Bedué & Fritzsche, 2021). Disse underkategoriene illustrerer at det er mulig å tildele teknologier tillitseenskapene man vanligvis assosierer med mennesker, ved at man tilpasser tillitsbegrepene spesifikt for teknologier. Dette er i tråd de nevnte tilpasningene Lankton et al. (2015) har gjort ved å avlede *funksjonalitet*, *hjelpsomhet* og *pålitelighet* fra henholdsvis *evne*, *velvilje* og *integritet*.

3.4 Tillit til KI

På bakgrunn av nevnte teori om tillit til teknologi, kan tillit til KI forstås som hvorvidt et individ stoler på, og er overbevisst om at KI-en er *funksjonell, hjelpsom og pålitelig* i sin utføring av en oppgave. Tillit til KI avhenger av dataintegritet, samsvar med sosiale normer og verdier, og overholdelse med regelverk og retningslinjer (Banavar, 2016). Sistnevnte er spesielt relevant i tråd med at EU AI Act trådte i kraft 1.august 2024 (Den europeiske kommisjonen, 2024).

Det er viktig å få en forståelse av tillit til KI fordi det er en avgjørende faktor for faktisk bruk av KI-tjenester (Korzyński, et al., 2024). Tillit er vist å ha den største signifikante effekten på faktisk bruk av generativ KI, sammenlignet med (1) tilretteleggende betingelser og (2) forventet innsats (Korzyński, et al., 2024). (1) Tilretteleggende betingelser omhandler ytre ressurser og støttende strukturer som gjør det enklere for noen å bruke teknologi. (2) Forventet innsats omhandler hvor lett eller vanskelig en bruker forventer at det vil være å bruke teknologien.

Bedué og Fritzsche (2021) peker på at tilgang til kunnskap, forklarbarhet, åpenhet, sertifisering og selvpålagte standarder er viktige faktorer for å øke tillit til KI. I det følgende skal vi gå nærmere inn på forklarbarhet og åpenhet, da disse fremheves som viktige underkategorier for tillitdimensjonen *evne* (Bedué & Fritzsche, 2021).

Mangel på forklarbarhet i KI-systemer er identifisert som en utfordring, spesielt når det gjelder ansvarliggjøring av svar fra KI. I tilfeller hvor KI kommer med feil svar eller løsninger man tviler på, er det vanskelig å vite hvem som skal stå til ansvar for disse. Åpenhet, eller transparens, er generelt viktig for de fleste teknologier, men det spiller en enda viktigere rolle for tillit til KI (Bedué & Fritzsche, 2021). Dette er fordi åpenhet er avgjørende for adopsjon av KI.

Et KI-verktøy kan være transparent på to måter: (1) forklare hvordan algoritmen fungerer og (2) reflektere over egen pålitelighet på en måte som er forståelige for alle brukerne (Glikson & Woolley, 2020). Åpenhet er essensielt i etableringen av kognitiv tillit. Kognitiv tillit handler om at brukeren anser verktøyet som *pålitelig* og *funksjonelt*. Følgelig vil åpenhet kunne påvirke brukerens oppfatning av KI-ens pålitelighet og funksjonalitet. I forskning på kognitiv tillit til KI, måler man hvorvidt brukerne er villige til å handle ut fra informasjonen og rådene fra KI-en, og om de oppfatter KI-en som hjelpsom, funksjonell eller nyttig (Glikson & Woolley, 2020).

Åpenhet er mer kompleks for KI enn for andre teknologier, spesielt for genererende KI-verktøy som bygger på en «black-box modell». Som tidligere nevnt gjør dette at det er vanskelig for brukeren å vite hvor informasjonen stammer fra (Glikson & Woolley, 2020). KI-ens evne til å lære av seg selv er en annen kompleksitet som reduserer KI-ens transparens (Bedué & Fritzsche, 2021). Mangel på åpenhet kan derfor være en faktor som reduserer tillit til KI-verktøy, sammenlignet med andre verktøy som i større grad kan vise hvor informasjonen kommer fra.

Til tross for at det kan være vanskelig å forstå hvordan KI løser oppgaver og kommer med løsninger, kan åpenhet om dens *funksjonalitet* og *pålitelighet* påvirke brukerens forventning til KI sin utførelse av en oppgave (Glikson & Woolley, 2020). Slik åpenhet kan redusere urealistisk høye nivåer av initial tillit, forhindre et drastisk fall i tillit, og beholde en stabil tillit på lang sikt. Glikson og Woolley (2020) mener derfor at videre forskning bør utforske åpenhet eller *pålitelighet* til teknologier som kan ha inkonsistent utførelse av identiske oppgaver, slik som ChatGPT.

3.4.1 Tidligere forskning

En omfattende analyse av tidligere forskning på tillit har vist at tillit til virtuell-KI, slik som ChatGPT, ofte er høy i starten, men faller med erfaring (Glikson & Woolley, 2020). Det er vist at oppgavetype påvirker individets tillitsnivå, noe som peker på at kontekst og situasjon har en effekt for et individs tillit overfor KI. Mennesker har generelt en oppfatning av at algoritmer er bedre på kalkulasjonsoppgaver enn sosiale oppgaver. De har derfor større tillit til KI dersom oppgaven som skal gjennomføres ikke krever sosial eller emosjonell intelligens (Glikson & Woolley, 2020). Glikson og Woolley hevder tillit til KI sannsynligvis avhenger av hvor intelligent KI-en oppfattes å være i brukerens øyne (2020).

Forskning viser at høy ekspertise innenfor situasjonen eller tematikken hvor man benytter KI, kan redusere tillit til KI-systemer, da eksperter ofte stoler mer på egen kunnskap enn på KI-anbefalinger (Siegrist & Cvetkovich, 2002). På den andre siden kan mennesker stole mer på KI-er sitt råd dersom de ikke selv har god nok kunnskap på feltet (Küper & Krämer, 2024). Dette illustrerer at det kan være relevant å undersøke om det er forskjeller i tillit mellom de som har sterk kompetanse om konteksten teknologien brukes i sammenlignet med andre.

Mennesker med en høyere tilbøyelighet til å stole på teknologier er mer tilbøyelige til å stole på råd fra KI (Alexander, et al., 2018). Dette er fordi en høyere tilbøyelighet til å stole på en teknologi fører til økt tillit til KI-systemer, noe som igjen øker sannsynligheten for at brukere

følger råd fra KI (Küper & Krämer, 2024). Generell tillitstilbøyelighet avhenger av kultur, alder, kjønn og andre personlighetsstrekk, og dermed spiller alle disse faktorene inn på tilliten et individ har til KI.

I Kapania et.al (2022) sin studie var deltakernes tillit til KI i stor grad basert på at de oppfattet KI-en som kompetent og *velvillig*. De anså verktøyet som *pålitelig*, god på å ta riktige beslutninger, ufeilbarlig og følelsesløs. Respondentene mente at KI sjeldnere tar feil enn mennesker, og at KI ikke har insentiver til å lyve, i motsetning til mennesker. De mente også at dersom en KI skulle vise seg å ta feil, er det menneskene bak verktøyet som er ansvarlige, ikke selve verktøyet. Deltakerne viste generelt høy tillit til KI, og flere påpekte at KI sjeldent tar feil ettersom den er bygd på et sett med regler. Dersom bruk av KI førte til et ugunstig utfall for deltakerne, skyldte de på at det var deres egen feil, nettopp fordi de mente KI sjeldent tar feil.

Deltakerne i studien var i stor grad takknemlig overfor KI-en fordi den reduserte deres avhengighet av andre individer. Et eksempel som ble nevnt var at verktøyet kunne hjelpe til med å finne informasjon om et tema, fremfor at individet var nødt til å oppsøke andre mennesker for å tilegne seg kunnskap, eller selv finne informasjon. Majoriteten av deltakerne var villige til å dele data med KI-en, i bytte mot pålitelige anbefalinger eller beslutninger. Videre hadde kvinner høyere aksept overfor KI-systemer sammenlignet med menn (Kapania, et al., 2022).

Overordnet er det viktigste funnet fra Kapania et.al (2022) at tillit til KI i stor grad påvirkes av brukernes oppfatning av KI som ufeilbarlig og kompetent. I tillegg påvirkes tilliten av at KI oppfattes som en upartisk aktør, som ikke har egne interesser eller motivasjoner, og at eventuelle feil tilskrives menneskene bak teknologien. Dette indikerer at individer ikke skylder på maskiner dersom det oppstår feil, men heller aktørene bak verktøyet, noe som er viktig å være klar over når man undersøker tillit til ChatGPT. Dette fordi individer kan skille mellom tillit til selve chatbot-applikasjonen og menneskene bak ChatGPT.

Dette funnet samsvarer med Inga Strümkes argumentasjon i boken «Maskiner som tenker»:

Det er ingeniørene og utviklerne som bestemmer avveiningene bak KI-verktøyet. Når et KI-verktøy er diskriminerende eller bias, ligger ikke forklaringen i at maskinen er diskriminerende. Forklaringen ligger i avveiningene som ble gjort da dataene ble innsamlet og maskinen trent opp. Dette er en viktig forståelse å ha i bakhodet når man skal forske på bruk av KI-verktøy. KI kan brukes til å løse problemer, men det er menneskene bak verktøyene som har ansvar for at den løser problemene på en etisk riktig måte (Strümke, 2023, s.11).

Etttersom ChatGPT er en tjeneste som er utviklet av mennesker, kan man derfor argumentere for at *evne*, *velvilje* og *integritet* kan tilskrives chatbot-applikasjonen. Dette skjer ved at egenskapene tilordnes utviklerne og eierne av ChatGPT. Dette er sentralt da det er vist at tillit til menneskene bak en teknologi også påvirker tilliten til selve teknologien.

3.5 Tillit i beslutningssituasjoner

Vi har til nå gjort rede for tillit, tillit til teknologi og tillit til KI, samt tilhørende underdimensjoner og underkategorier. I denne seksjonen skal vi først gå nærmere inn på hvordan tillit er sentralt i ulike beslutningssituasjoner. Deretter skal vi se hvordan tillit også er relevant i beslutningssituasjoner hvor KI er involvert.

Tillit er sentralt i en rekke ulike beslutningssituasjoner. For eksempel spiller tillit en essensiell rolle i kjøpsbeslutninger og har positive effekter spesielt i sammenheng med netthandel (Gefen, et al., 2011). En av de viktigste effektene er at tillit øker forbrukerens intensjon om å foreta kjøp. Tillit er også viktig i transaksjonsforhold mellom kjøper og selger, spesielt når det inkluderer en viss risiko, som ved interaksjon med en online selger (Gefen, et al., 2011).

Tillit er viktig i både sosiale og økonomiske transaksjoner fordi mennesker har behov for å forstå sine sosiale omgivelser – å vite hva, når, hvorfor og hvordan andre oppfører seg (Gefen, et al., 2011). Når regler og skikker ikke kan kontrollere en sosial situasjon, bruker folk tillit som en strategi for å håndtere kompleksiteten. Ved å stole på andre, reduserer folk opplevd sosial kompleksitet og aksepterer en tro som kan være irrasjonell, men som reduserer risikoen for uønsket oppførsel fra den de stoler på. Det samme gjelder for internett, hvor forbrukere må stole på online selgeren i mangel av effektiv regulering. Uten effektiv regulering må forbrukere stole på online selgeren de handler fra, og anta at selgeren vil opptre etisk og på en sosial

akseptabel måte. Hvis ikke, vil den sosiale kompleksiteten gjøre at de unngår å foreta kjøpet (Gefen, et al., 2011).

Tidligere forskning viser at tillit øker kjøpsintensjoner, både direkte og indirekte ved å redusere oppfattet risiko (Gefen, et al., 2011). En undersøkelse gjort av Reichheld og Scheffer viser at tillit i større grad styrer online atferd sammenlignet med pris (2000). Når forbrukere stoler på en online selger, er de mer tilbøyelige til å dele personlig informasjon. Konsekvensene av manglende tillit er derimot at forbrukerne vil handle et annet sted (Reichheld & Scheffer, 2000).

I lys av at tillit er en viktig faktor bak beslutninger som å handle på nett eller bruke algoritmebaserte verktøy, er en konsekvens av manglende tillit at brukeren avstår fra å foreta en kjøpsbeslutning eller adoptere algoritme-baserte teknologier.

3.5.1 Tillit til KI i beslutningssituasjoner

I beslutningssituasjoner hvor brukere skal anvende verktøy med «black-box»-algoritmer er de ikke nødvendigvis kjent med verktøyets indre funksjonaliteter (Logg, et al., 2019). Mennesker som bruker verktøy med «black-box»-algoritmer er villige til å stole på algoritmene til tross for manglende informasjon om algoritmenes funksjoner. I slike situasjoner vil brukeren lage seg egne teorier om hvilken type informasjon algoritmen bruker som inndata, hvordan informasjonen er prosessert, og verdien av utdata (Logg, et al., 2019).

I en studie av indiske KI-brukere ble det vist at nesten 80% aksepterer beslutninger utført av KI, og at KI hadde makt til å påvirke respondentenes handlinger (Kapania, et al., 2022). 84,1% av respondentene var villige til å stole på beslutninger fra KI i situasjoner hvor risiko har lite betydning, slik som sanganbefalinger og anbefalt kjørerute. I høyrisiko situasjoner, slik som medisinsk diagnostisering eller beslutninger tilknyttet ansettelse, stolte 73,8% av respondentene på beslutninger fra KI. I beslutninger relatert til lånevurderinger, svarte intervjuobjektene at de foretrakk at KI vurderte søknaden deres fremfor mennesker. Denne studien indikerer at grad av risiko i beslutningssituasjoner påvirker tillit til KI.

3.6 Faktorer som svekker tillit

I foregående seksjoner har vi pekt på faktorer som kan påvirke tillit. I denne seksjonen skal vi gå nærmere inn på momenter som kan redusere tillit.

Generelt reduseres tillit når den sosiale distansen mellom individet og tillitsobjektet øker (Evans & Krueger, 2009). Når individet er tett tilknyttet sosialt med tillitsobjektet øker tilliten (Glaeser, et al., 2000). Vi stoler på andre mennesker dersom vi føler at vi er en del av samme

gruppe, noe som kan forstås som sosial tillit (Brewer, 2008). Når vi ikke opplever at vi tilhører samme gruppe som et annet individ, har vi mindre tillit til individet. Sosial identitetsteori er dermed med på å forklare tilliten mennesker har til hverandre, hvor vi i større grad stoler på inn-grupped medlemmene, altså individene som er en del av samme gruppe som oss selv (Tanis & Postmes, 2005). Det er vist at dersom individer er klar over andre menneskers gruppetilhørighet, velger de enstemmig å stole på individer fra inn-gruppen (Foddy, et al., 2009). Dette kan overføres til teknologier: Dersom inn-gruppen til et individ benytter en teknologi, kan det være nærliggende at individet har tillit til denne teknologien fordi gruppen har etablert en felles forståelse av teknologien som tillitsfull.

I truende situasjoner, eller situasjoner hvor det er høy risiko for negative utfall, har mennesker en tendens til å velge å benytte seg av verktøy som er troverdige og pålitelige, og ofte velge mennesker fremfor teknologier (Cahour & Forzy, 2009). For å øke tillit til teknologier er det derfor nødvendig å redusere opplevd risiko tilknyttet situasjonen hvor teknologien benyttes (Muir, 1994).

En annen grunn til at mennesker ikke stoler på teknologier, kan være fordi de har teknologiangst. Teknologiangst handler om brukerens mentale tilstand i forhold til deres evne og vilje til å bruke teknologi-baserte verktøy (Meuter, et al., 2003). Teknologiangst påvirker brukerens tilfredshet med teknologien. Meuter et al. (2003) fant at teknologiangst i større grad påvirker et individs villighet til å benytte teknologier sammenlignet med demografiske faktorer slik som alder, kjønn, utdanning og inntekt. Frykt for teknologi har også en negativ påvirkning på pasienters aksept for helseteknologi (Calvin & Karsh, 2009). Dette indikerer at dersom en bruker har et generelt høyt nivå av teknologiangst, kan det tenkes at de er mindre tilbøyelige til å stole på teknologier, som eksempelvis ChatGPT.

3.6.1 Algoritmeaversjon

Tidligere har vi presentert teori som viser at tillit er en avgjørende faktor for adopsjon av teknologi og KI (Choi & Ji, 2015; Korzyński et al., 2024). Imidlertid kan manglende tillit også være en barriere for at mennesker adopterer algoritmebaserte system (Logg, et al., 2019). Et eksempel på dette er algoritmeaversjon, som refererer til menneskers tendens til å være skeptiske til algoritmer og foretrekke vurderinger gjort av mennesker fremfor maskiner (Dietvorst, et al., 2018).

En konsekvens av manglende tillit til algoritmer er at mennesker heller vil velge å rådføre seg med mennesker fremfor algoritmer, selv i situasjoner hvor algoritmen ville løst oppgaven bedre

(Dietvorst, et al., 2018). Det finnes flere studier som støtter disse funnene, og viser til at mennesker generelt foretrekker å få råd fra andre mennesker fremfor algoritmer (Önkal, et al., 2009; Longoni, et al., 2019).

Mennesker kan miste tillit til en algoritme når de observerer hvordan den fungerer, selv om algoritmen objektivt sett leverer bedre resultater enn et menneske i en bestemt oppgave (Dietvorst, et al., 2018). Dette betyr at selv om algoritmen presterer bedre enn mennesker i mange situasjoner, er det ikke alltid nok til å overvinne individers skepsis og preferanse for beslutninger tatt av mennesker. Årsaken er at mennesker har en tendens til å raskere miste tillit til en algoritme enn til et menneske etter å ha sett dem gjøre samme feil.

I tråd med funnene til Dietvors et. al (2018), viser Cadario et. al (2021) at mennesker er villige til å stole på råd fra algoritmen før de ser den feile. Dette tyder på at det er først når individet opplever feil ved algoritmen, at tilliten til den svekkes. Følgelig er det relevant å undersøke om det å fremheve at algoritmebaserte verktøy kan gjøre feil også påvirker tilliten negativt.

Mennesker ser ut til å raskere miste tilliten til en algoritme når de ser den gjøre små feil, samtidig som de er mer tilgivende overfor mennesker som gjør større feil (Dietvorst, et al., 2015). For å forklare nærmere kan vi vise til et enkelt eksempel. Se for deg at du opplever trafikk på din vanlige rute til jobben, og derfor tar en annen rute for å spare tid. Det viser seg at den nye ruten fører til at du kommer enda senere enn om du hadde forholdt deg til din vanlige rute. I dette scenarioet ville de fleste fortsatt stole på egen beslutningsevne til tross for denne feilkalkuleringen. Se for deg samme scenario, men at det er GPS-en som endrer rute slik at du kommer for sent. Dette gjør at du trolig får redusert tillit til GPS-en, og at du er mindre villig til å bruke denne i fremtidige situasjoner.

En direkte konsekvens av manglende tillit til algoritmebasert teknologi er at individer velger bort algoritmer til fordel for mennesker (Dietvorst, et al., 2018), også i tilfeller hvor de menneskelige vurderingene har høyere risiko for feil. Dietvors et. al (2018) foreslår at denne atferden spiller en sentral rolle bak algoritmeaversjon, noe som betyr at algoritmeaversjon helt eller delvis avhenger av menneskers erfaring med algoritmen.

Mennesker viser også lavere tillit til algoritmebaserte anbefalinger når de har mer kunnskap om konteksten eller opplever at de selv har tilstrekkelig kompetanse til å ta beslutningen (Logg, et al., 2019). Dette gjør at mennesker som er eksperter i sine fagfelt er mer skeptiske til å søke råd fra algoritmer på fagfeltet. En studie viste at en kortfattet oppsummering som visualiserer

og forklarer hvordan et algoritmebasert verktøy fungerer kan øke aksepten for verktøyet (Cadario, et al., 2021).

Samtidig som flere studier undersøker algoritmeaversjon (Reich, et al., 2022), har andre studier funnet at mennesker også verdsetter algoritmer, såkalt algoritmeverdsettelse. I en nylig studie fant Logg, Minson og Moore at respondentene stolte mer på det samme rådet dersom det ble gitt av algoritmer enn av mennesker (Logg, et al., 2018).

3.7 Tillit til informasjonsverktøy

I vår studie skal vi som nevnt sammenligne tillit til ChatGPT med informasjonsverktøyene Google og Wikipedia. Av den grunn er det derfor relevant å se på hvorfor tillit er viktig når man skal innhente informasjon.

3.7.1 Viktigheten av tillit i informasjonsinnhenting

Det er viktig å kartlegge tillit til informasjonskilder da vi benytter oss av informasjon til å ta beslutninger i en rekke ulike settinger. Tillit har stor betydning for vår forståelse av hvordan mennesker håndterer informasjon (Kelton, et al., 2008). Som Kelton et al. skriver: «Hvis ikke tillit, hva da?». Tillit er selve grunnlaget for hvilke valg vi tar når vi innhenter informasjon og avgjør hvilken informasjonskilde vi skal benytte. Kelton et al. (2008) argumenterer derfor for at det er et stort behov for ytterligere forskning på tillit til informasjon, og mener tillit bør få en større plass i forskningen.

Vi henter konstant informasjon i hverdagen, i arbeidslivet og i utdanning. I hverdagen finner vi informasjon som kan hjelpe oss med å ta beslutninger angående praktiske problemer, som hva man skal ha til middag, hvordan man lager en middagsrett og hvor man kan få tak i ingrediensene. I arbeidslivet henter ansatte og ledere informasjon om kunder og markedet, og bruker informasjon som grunnlag for beslutningstaking. I utdanningen bruker studenter informasjon for å lære seg mer om et tema, samt skrive oppgaver og forberede seg til eksamen. Overordnet er informasjonsinnhenting en grunnleggende menneskelig aktivitet som foregår i en rekke ulike situasjoner og settinger, og informasjon legger grunnlaget for beslutninger.

I dagens samfunn henter vi stadig mer informasjon online (Fallows, 2005) fremfor å oppsøke informasjon i fysiske bøker eller annet fysisk materiale. Nesten halvparten av amerikanske internettbrukere i 2006 fortalte at de hadde brukt internett for å ta store beslutninger eller navigere seg gjennom større hendelser i livet (Horrigan & Rainie, 2006). Dette utgjorde på daværende tidspunkt 60 millioner amerikanere, og viser til at markedet for online

informasjonssøking allerede var enormt for 18 år siden. Studien illustrerer hvor sterk innflytelse informasjon fra internett har på individer i beslutningssituasjoner.

I mange tilfeller er internett det første stedet man søker etter informasjon (Eysenbach, 2008). Når vi leter etter informasjon på internett, er det vi som sluttbrukere som må avgjøre hvorvidt vi stoler på informasjonen (Metzger, et al., 2003). I dag eksisterer det lite filtrering av feilinformasjon, noe som resulterer i at informasjonen på internett kan være utdatert, ufullstendig eller direkte feil (Rieh & Danielson, 2008). Det er også få standarder for kvalitetskontroll av materiale som publiseres på nett (Metzger & Flanagin, 2013).

Den enorme mengden av informasjon på internett har ført til bekymringer om troverdighet, ettersom kvaliteten og påliteligheten til kildene nå er mer usikre enn tidligere (Metzger & Flanagin, 2013). «Fake news», eller falske nyheter, er et begrep som har eksistert over lengre tid, men fikk et kraftig oppsving under det amerikanske presidentvalget i 2016 og innsettelsen av Donald Trump som president (Orgeret & Henrik, 2020).

Falske nyheter er nyheter som kan virke troverdig, men som inneholder misvisende informasjon og ofte har til hensikt å være villedende. Å bli utsatt for falske nyheter kan føre til redusert tillit til media, men kan øke tilliten til myndigheter (Ognyanova, et al., 2020). Falske nyheter spres gjerne på sosiale medier (Orgeret & Henrik, 2020), og det kan være vanskelig for brukerne å skille mellom troverdig versus ikke-troverdig informasjon, eller sanne versus falske nyheter.

Med den økte internettilgangen, har det også åpnet muligheter for å publisere innhold som har falsk informasjon. Dette innholdet kan være brukergenerert (som for eksempel på Wikipedia), redaktørstyrt (som for eksempel på Store norske leksikon), eller algoritmegenerert (som for eksempel på ChatGPT). En av grunnene til at det er viktig å avdekke om individer har tillit til en informasjonskilde er fordi man ønsker å undersøke om individer stoler på kilder som kan gi feilinformasjon.

Det er flere grunner til at dette er viktig. Når det gjelder ChatGPT, er det spesielt interessant å undersøke om individer stoler blindt på verktøyet fordi det (1) kan være partisk, som oppstår på grunn av skjevheter i treningsdataene, (2) kan gi brukeren misledende eller feilinformasjon, (3) føre til mindre grad av kritisk tenking og selvstendig problemløsning, og (4) generere skadefullt materiale slik som falske nyheter (Ray, 2023; Gosline et al., 2024).

Overordnet ser vi at tillit til informasjonsverktøy er avgjørende fordi informasjon danner grunnlaget for beslutningstaking i hverdagen, arbeidslivet og utdanningen. Internett har lenge vært den primære plattformen for informasjonsinnhenting, men manglende kvalitetskontroll og risiko for feilinformasjon utfordrer brukernes tillit.

3.7.2 Redaktørstyrt innhold

For å forhindre feilinformasjon og sikre kvalitet på informasjon, har mange plattformer redaktørstyrt innhold. Redaktørstyrt innhold er informasjon som er kvalitetssikret av en redaktør før publisering, hvor redaktøren fungerer som portvokter (Elevkanalen, u.d.). Tradisjonelt sett er en portvokter et individ som kontrollerer informasjonsflyten inn i en organisasjon fra omverden (Macdonald & Williams, 1994). Redaktørens oppgave er å sørge for at informasjonen er pålitelig, noe som påvirker tilliten til innholdet. Wikipedias portvoktere er brukerne som redigerer innholdet, mens for Google er algoritmene ansvarlige for prioritering og filtrering av informasjon.

Falske nyheter har, som nevnt, blitt mer vanlig og svekker tilliten til redaktørstyrte plattformer, da det kan være utfordrende å skille mellom ekte og falsk informasjon. Å ha en kompetent redaktør som sikrer at informasjonen er sann, er derfor kritisk.

Wikipedias innhold redigeres av brukerne selv, men det kan være utfordrende å vurdere deres pålitelighet (Jean, 2024). Tilliten til Wikipedia avhenger av brukernes tillit til redaktørene, noe som stadig forskes på i takt med økende publisering av artikler (Latif & Jaffry, 2013).

3.7.3 Brukergenerert innhold

I tillegg til å være redaktørstyrt, har Wikipedia også brukergenerert innhold. Brukergenerert innhold viser til informasjon som er generert av plattformens egne brukere (Hochstein, et al., 2023).

Muligheten til å redigere artikler og se endringene umiddelbart bidro til Wikipedias raske vekst etter lanseringen i 2001 (Luyt, et al., 2008; Maggio, et al., 2017). Til tross for tidligere spørsmål om pålitelighet (Giles, 2005; Denning, et al., 2005; Luyt, et al., 2008; Lim, 2009), bruker stadig flere Wikipedia som kunnskapsressurs. For eksempel viste en studie at 75% av studenter i Liverpool brukte Wikipedia (Knight & Pryke, 2012). Både studenter og lærere anerkjenner verktøyet i større grad etter de lærer hvordan det fungerer og viktigheten av å være kritisk til innholdet på plattformen (Bravo & Young, 2011).

Selv om Wikipedia er godt integrert i studenters hverdag, er det ikke deres primære informasjonskilde (Knight & Pryke, 2012). Forventninger til Wikipedias kvalitet påvirkes av tidligere erfaringer med verktøyet, følelsesmessig tilstand, tilbøyelighet til å tro på informasjonen, og innholdets nytteverdi.

Det er argumentert for at Wikipedia har høyere pålitelighet sammenlignet med andre leksikon, og at Wikipedia foretrekkes fremfor alternative informasjonsverktøy brukere ville benyttet dersom Wikipedia ikke hadde eksistert (Fallis, 2008). Den økte bruken av verktøyet skaper et behov for å kartlegge tillit til innholdet, spesielt fordi brukere kan redigere innhold (Lucassen & Schraagen, 2010).

Huang et al. (2015) argumenterer for at tillit til Wikipedia handler om en bruker har tillit til informasjonen som er generert av brukerne på plattformen, og ikke om brukeren stoler på selve plattformen. Dette er i tråd med tidligere argumentasjon om at tillit kan tilegnes menneskene bak en teknologi.

3.7.4 Tillit til ChatGPT, Google og Wikipedia

Litteraturen har tidligere sett på individers tillit til ChatGPT, Google og Wikipedia (Jung, et al., 2024). Jung et al. (2024) fant at Google-brukere stoler på plattformen fordi den ansees som en god portvokter («gatekeeper»). Googles søkemotor er en digital portvokter da den bestemmer hvilke nettsider som først kommer opp i søkeresultatene. Gjennom algoritmer filtrerer og prioriterer Google hvilke kilder som skal vises til brukerne, gjerne basert på relevans, troverdighet og popularitet. Individer stoler på Google fordi informasjonen som presenteres kommer fra flere ulike kilder, og ofte ansees som pålitelig (Jung, et al., 2024). Tillit til Google påvirkes positivt av at plattformen tydelig markerer annonsert innhold fordi det reflekterte åpenhet.

I motsetning til dette, fant Jung et al. at individer har varierende grad av tillit til Wikipedia (Jung, et al., 2024). Enkelte stoler mer på nettsiden nettopp fordi den baseres på bidrag fra et stort antall mennesker, som muliggjør et omfattende og mangfoldig innhold. I studien stolte flere på Wikipedia fordi plattformen viser hvor informasjonen opprinnelig kommer fra. På den andre siden stolte færre på Wikipedia fordi muligheten til å redigere innhold åpner for at brukere uten kompetanse kan bidra, noe som reduserer innholdets pålitelighet (Huang & Sundar, 2020).

Studien fra Jung et al. (2024) fant at deltakerne generelt hadde mest tillit til Google, og deretter Wikipedia. Hovedgrunnen til at ChatGPT ble stolt minst på var fordi denne plattformen ikke viste til hvor den hentet informasjonen fra. Forskerne hevdet en synliggjørelse av hvor ChatGPT får informasjonen fra, kan øke tilliten individer har til dette verktøyet sammenlignet med andre informasjonsverktøy. Forskerne konkluderte med at Google er verktøyet brukerne tyr til først, men at dette kan endres i fremtiden da KI-verktøy brukes i økende grad (Jung, et al., 2024) (Jung, et al., 2024).

4. Hypoteser og modell

I denne seksjonen skal vi presentere seks hypotesesett som er utformet basert på tidligere nevnte bakgrunn og teori. Hypotesesettene legger grunnlaget for videre analyse. Vi skal deretter presentere forskningsmodellen vår, som er en visualisering av hypotesesettene, og viser sammenhengen mellom de uavhengige variablene og den avhengige variabelen.

Fremover kommer vi til å benytte oss av begrepene *systemtillit* og *mennesketillit*. Systemtillit viser til en samling av systemtillitsdimensjonene: *funksjonalitet*, *hjelpsomhet* og *pålitelighet*. Når vi omtaler mennesketillit referer vi til tilliten som bygger på mennesketillitsdimensjonene: *evne*, *velvilje* og *integritet*, ref. teoridel.

4.1 Hypotesesett 1: Systemtillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia

I dette første hypotesesettet skal vi ta for oss tillit til ChatGPT sammenlignet med Google og Wikipedia for systemtillitsdimensjonene.

ChatGPT sammenlignet med Google for systemtillit

Når det gjelder tillit til ChatGPT sammenlignet med Google, fant Jung et al. (2024) i sin studie at deltakerne generelt stolte mest på Google, sammenlignet med Wikipedia og ChatGPT (Jung, et al., 2024). Forskerne fant at deltakerne i større grad stolte på Google enn ChatGPT, fordi ChatGPT ikke viser hvor den henter informasjonen fra. I tillegg ble Google oppfattet som mer pålitelig da verktøyet presenterer en rekke ulike kilder til brukeren, slik at brukeren selv kan velge hvilken kilde de ønsker å benytte seg av.

Jung et al. sin studie er relativt ny, noe som gir grunnlag for å hevde at oppfatningen av Google som mer tillitsfull enn ChatGPT ikke har endret seg nevneverdig. Etersom Jung et al. benyttet

en kvalitativ forskningstilnærming med intervju (n=14) som forskningsmetode, er ikke funnene statistisk signifikante. Vi ønsker dermed å se på om deres funn kan generaliseres til en større populasjon ved å gjennomføre en kvantitativ studie, for å teste om ChatGPT har lavere tillit sammenlignet med Google for systemtillitsdimensjonene.

Google bidrar med en bred tilgang til en rekke ulike kilder sammenlignet med ChatGPT, hvor brukeren spesifikt må etterspørre kilder. Dette gjør at brukere kan oppfatte Google som mer funksjonell fordi verktøyet gir tilgang til hele internettet, i tillegg til å tilby en rekke ulike tilleggselementer slik som kart, oversettelse, akademiske artikler etc. Google er slik funksjonell fordi den har et bredt spekter av tjenester for brukernes ulike behov. ChatGPT tilbyr som tidligere nevnt mulighet til bilde- og tekstgenerering, men har ikke de andre tilleggselementene som Google.

Det at Google presenterer mange kilder som brukeren selv kan velge mellom, kan bidra til at Google oppleves som mer hjelpsom enn ChatGPT. Samtidig gir ChatGPT ofte et direkte og konkret svar på brukerens spørsmål, noe som eliminerer behovet for å søke gjennom kilder for å finne svaret. Dette kan føre til at enkelte brukere oppfatter ChatGPT som mer hjelpsom enn Google. Likevel kan risikoen for feilaktige eller misvisende svar fra ChatGPT trekke i motsatt retning og redusere opplevelsen av hjelpsomhet sammenlignet med Google.

Sammenlignet med Google, kan ChatGPT skåre lavere på pålitelighet fordi plattformen kan generere feilaktige svar. Google lar brukeren navigere og verifisere informasjonen selv, noe som kan øke grad av pålitelighet. Det er også verdt å nevne at Google har helt fersk informasjon. ChatGPT kan til en viss grad presentere oppdaterte artikler og sanntidsdata, men det kan være treningsdataene ikke nødvendigvis inkluderer alle de nyeste hendelsene. Videre lærer ChatGPT underveis, og vil derfor ikke ha en forutsigbar konstant adferd. Dette kan illustreres med at man kan spørre ChatGPT det samme spørsmålet flere ganger og få ulike svar hver gang, noe som kan svekke pålitelighet. Dette støtter argumentasjonen om at ChatGPT kan skåre lavere på pålitelighet sammenlignet med Google, som kan gi samme svar ved hvert søk.

Overordnet gir denne argumentasjonen grunnlag for å teste følgende hypotese:

H₁₋₁: Brukerne har lavere tillit til ChatGPT sammenlignet med Google, og vil skåre ChatGPT lavere på a) funksjonalitet, b) hjelpsomhet og c) pålitelighet.

Nullhypotesen er:

H_{1-Null}: Det er ingen signifikant forskjell i brukernes tillit til ChatGPT sammenlignet med Google, og det er ingen forskjell i oppfattet funksjonalitet, hjelpsomhet og pålitelighet.

ChatGPT sammenlignet med Wikipedia for systemtillit

I likhet med argumentasjonen for ChatGPT sammenlignet med Google tar vi Jung et al. (2024) sin studie i betraktning for å utforme hypoteser om tillit til ChatGPT sammenlignet med Wikipedia. Som tidligere nevnt hadde deltakerne i denne studien mer tillit til Wikipedia enn ChatGPT, noe som i større grad skyldes plattformens åpenhet om hvor informasjonen kommer fra. Den samme argumentasjonen som for forrige hypotesesett kan benyttes her: Jung et al. sin studie er ny, og det er dermed lite trolig at tillitsnivået for de ulike informasjonsverktøyene har endret seg betraktelig.

Det kan argumenteres for at Wikipedia er mer funksjonell enn ChatGPT fordi plattformen har en enkel oppbygning som er i en standardisert form, med tydelige inndelinger slik som innledning, metode og referanser. Dette gjør at brukergrensesnittet er kjent for brukeren, og gir en tydelig oversikt over innholdet på siden. I motsetning til dette, leverer ChatGPT ofte informasjon i samtalebasert form, noe som kan være lite effektivt for omfattende tema brukeren ønsker innsikt i.

Grad av hjelpsomhet kan oppfattes som høyere hos Wikipedia fordi verktøyet gir informasjon som er basert på dokumenterte kilder. Dette gjør at brukerne kan benytte verktøyet om de skal finne materiale til studier eller akademisk forskning. ChatGPT kan gi gode og enkle forklaringer, men kan tenkes å være mindre hjelpsom når en bruker ønsker å dykke ned i et akademisk tema, fordi svarene ikke alltid er faktuelle. Likevel er det mulig å argumentere for at ChatGPT kan oppleves som hjelpsom når det gjelder andre kontekster, som for eksempel hverdagslige spørsmål som hva man skal lage til middag, pakkelister etc. Grad av hjelpsomhet kan dermed variere ut fra situasjonene de to verktøyene skal benyttes i.

Wikipedia kan skåre høyere enn ChatGPT på systemtillitsdimensjonen *pålitelighet*, fordi plattformen har et omfattende nettverk av interne lenker og eksterne referanser. Wikipedia tydeliggjør for brukerne hvor informasjonen er hentet fra gjennom en referanseliste, som gjør det mulig for brukerne å verifisere informasjonen selv. ChatGPT gir derimot ikke brukeren kildehenvisning for hver prompt (spørsmål fra bruker), men brukeren kan spesifikt etterspørre

hvor kildene er hentet fra. Likevel kan ikke ChatGPT garantere at kildene er korrekte, og verktøyet tilbyr heller ikke samme mulighet til å navigere gjennom hyperkoblinger til eksterne artikler slik som Wikipedia.

Et annet argument som styrker hypotesen om at Wikipedia er mer pålitelig enn ChatGPT, er at Wikipedia har strenge regler for at informasjonen skal være korrekt. Artikler som ikke har gode kilder, blir merket og kan fjernes (Ayers, et al., 2008). For ChatGPT er ikke dette et alternativ, da verktøyet genererer ny informasjon for hver prompt. Videre viser Wikipedia hvilke kilder som benyttes, hvem som redigerer artikler og når artiklene sist ble oppdatert. Dette er alle faktorer som kan bidra til å øke plattformens åpenhet og dermed pålitelighet. Det er ikke mulig å innhente slik informasjon fra ChatGPT, noe som trekker i retning av at Wikipedia kan ansees som mer pålitelig.

Overordnet gir denne argumentasjonen grunnlag for å teste følgende hypotese:

H₁₋₂: Brukerne har lavere tillit til ChatGPT sammenlignet med Wikipedia, og vil skåre ChatGPT lavere på a) funksjonalitet, b) hjelpsomhet og c) pålitelighet.

Nullhypotesen er følgende:

H_{1-2null}: Det er ingen signifikant forskjell i brukernes tillit til ChatGPT sammenlignet med Wikipedia, og det er ingen forskjell i oppfattet funksjonalitet, hjelpsomhet og pålitelighet.

4.2 Hypotesesett 2: Mennesketillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia

I dette andre hypotesesettet skal vi ta for oss tillit til ChatGPT sammenlignet med Google og Wikipedia for mennesketillitsdimensjonene.

Som tidligere presentert, fant Jung et al. i sin studie at deltakerne generelt hadde mest tillit til Google og Wikipedia sammenlignet ChatGPT (Jung, et al., 2024). Vi legger denne studien til grunn for å bestemme retningen til hypotesene våre om høyere mennesketillit til ChatGPT sammenlignet med Google og Wikipedia.

Mennesketillit til ChatGPT sammenlignet med Google

Den samme argumentasjonen for at Google vil kunne oppfattes som mer funksjonell enn ChatGPT gjelder også for *evne*. Brukere kan oppfatte at Google har en bedre evne til å tilby troverdig informasjon, da verktøyet som nevnt presenterer en rekke ulike kilder, sammenlignet

med ChatGPT hvor dette må etterspørres. Det at ChatGPT kan ta feil og være partisk (Ray, 2023) kan også svekke brukerens oppfatning av dens *evne* til å levere pålitelige svar.

Videre argumenterer vi for at brukere kan oppfatte Google som mer velvillig da søkemotoren er gratis og tilgjengelig for alle. Dette indikerer at Google har en vilje til å bidra positivt til informasjonsformidling for samfunnet, uten at brukerne må betale for tilgang til alle funksjonene. I motsetning til Google, har ChatGPT en betalt versjon og en gratisversjon. Den betalte versjonen har vist seg å være mer presis (Lim, et al., 2023), bedre utføre spesifikke industrirelaterte oppgaver (Lewandowski, et al., 2024), og klare å svare på spørsmål av høyere vanskelighetsgrad (Knoedler, et al., 2024). Googles tilgjengelighet på tvers av enheter og dens gratis tilgang til alle funksjoner for brukerne bidrar til en opplevelse av likhet og inkludering, noe som kan forsterke oppfatningen av velvilje. ChatGPTs skille mellom en gratis og en betalt versjon kan derimot tolkes som en begrensning som hindrer likeverdig tilgang til alle funksjoner.

Google kan argumenteres for å ha mer integritet enn ChatGPT, da Googles algoritmer er designet for å sikre at informasjonen som presenteres er både relevant og av høy kvalitet, med fokus på å gi verdi til brukeren (Pasquinelli, 2009). ChatGPTs algoritmer, derimot, er trent på store datasett og trenes kontinuerlig basert på nye inndata. Denne kontinuerlige treningen kan gjøre det utfordrende å vurdere integriteten til ChatGPT, da svarene er avhengige av stadig oppdatering og påvirkes i stor grad av hvordan brukeren formulerer sitt prompt (Piedboeuf & Langlais, 2023).

På bakgrunn av Jung et al. (2024) sin studie og den overordnede argumentasjonen av de menneskelige tillitsdimensjonene gir det grunnlag for å teste følgende hypotese:

H₂₋₁: Brukerne har lavere tillit til ChatGPT sammenlignet med Google, og vil skåre ChatGPT lavere på a) evne, b) velvilje og c) integritet.

H_{2-1null}: Det er ingen signifikant forskjell i brukernes tillit til ChatGPT sammenlignet med Google, og det er ingen forskjell i oppfattet evne, velvilje og integritet.

Mennesketillit til ChatGPT sammenlignet med Wikipedia

Wikipedia kan tenkes å skåre høyere på tillitsdimensjonen *evne* sammenlignet med ChatGPT, fordi Wikipedia har et bredt spekter av ulike fagområder som er redigert av fageksperter eller individer med mye kunnskap innenfor det gitte temaet. Dette gjør at artiklene på Wikipedia har mer dybde og nisjefakta enn informasjonen ChatGPT tilbyr brukerne.

Etttersom Wikipedia er basert på et globalt fellesskap av frivillige individer som deler kunnskap og informasjon for å hjelpe brukerne på plattformen til å øke sitt kunnskapsnivå, kan verktøyet oppfattes som velvillig. Dette er fordi det ikke er noen økonomisk sterke insentiver eller skjulte agendaer bak informasjonsdeling og redigering, utover et ønske om å bidra til fellesskapet. ChatGPT, derimot, kan oppfattes som en teknologisk plattform uten en menneskelig fellesskapsdimensjon, noe som kan svekke oppfatningen av velvilje. Disse forskjellene danner grunnlaget for en hypotese om at Wikipedia blir oppfattet som mer velvillig enn ChatGPT.

Informasjonen på Wikipedia gjennomgår en kvalitetskontroll for å sikre at materialet er av høy standard (König, 2012), noe som kan styrke plattformens integritet og støtte hypotesen om at Wikipedia har mer integritet enn ChatGPT. ChatGPTs utdata avhenger av treningsdata, og blir ikke korrigert av mennesker, i motsetning til Wikipedia, som benytter en åpen redigeringsmodell der brukerne aktivt kan forbedre og kvalitetssikre innholdet.

På bakgrunn av Jung et.al (2024) sin studie og den overordnede argumentasjonen av de menneskebaserte tillitsdimensjonene gir det grunnlag for å teste følgende hypotese:

H₂₋₂: Brukerne har lavere tillit til ChatGPT sammenlignet med Wikipedia, og vil skåre ChatGPT lavere på a) evne, b) velvilje og c) integritet.

Nullhypotesen er følgende:

H_{2-2null}: Det er ingen signifikant forskjell i brukernes tillit til ChatGPT sammenlignet med Wikipedia, og det er ingen forskjell i oppfattet evne, velvilje og integritet.

4.3 Hypotesesett 3: ChatGPT sammenlignet med ChatGPT med en advarsel

I dette hypotesesettet skal vi undersøke hvordan tilliten til ChatGPT påvirkes av en fremheving av at ChatGPT selv synliggjør på sin plattform at den kan ta feil.

ChatGPT har etter sin lansering i november 2022 vekket stort engasjement og blitt flittig brukt av mange, inkludert studenter, konsulenter, advokater og andre grupper (Zulic, et al., 2024;

Ødegaard, 2023; Merken, 2023). Selv om bedriftslederne anerkjenner ChatGPTs potensiale, hevdes det at tilliten til verktøyet er begrenset på grunn av bekymringer om at den er partisk, unøyaktig og kan gjøre feil (Gosline, et al., 2024).

På den annen side kan brukerne oppfatte det at ChatGPT fremhever egne feil på plattformen som en transparent handling. Dette innebærer at ChatGPT til en viss grad reflekterer over egen pålitelighet, noe som er en av måtene KI-verkøty kan vise transparens (Bedué & Fritzsche, 2021). Bedué og Fritzsche fremhever at transparens er en nøkkelfaktor for å øke tillit til KI (2021), mens Gilkson og Woolley (2020) bekrefter at transparens styrker tillit til virtuelle KI-verkøty. Dette trekker i retning av at tilliten til ChatGPT vil styrkes av en advarsel om at den selv fremhever at den kan ta feil, ved at det tolkes som en transparent handling.

På bakgrunn av Bedué og Fritzsche, samt Gilkson og Woolley sine studier, som viser at transparens styrker tillit, antar vi at mennesketillit og systemtillit begge styrkes av en advarsel om at ChatGPT kan ta feil. Videre kan åpenhet om egne feil antas å styrke oppfatningen av ChatGPT sin *integritet* og *pålitelighet*, da dette er en ærlig handling (NAOB, u.d.). Det kan også antas at brukerne anser ChatGPT som mer *hjelpsom* og *velvillig* når den er åpen om egne feil. Derimot kan åpenhet om egne feil svekke oppfatningen av ChatGPT sin evne og funksjonalitet.

Samlet sett vil derfor en fremheving av feil kunne styrke de systembaserte tillitsdimensjonene; *pålitelighet* og *hjelpsomhet*, og de menneskebaserte tillitsdimensjonene; *integritet* og *velvilje*. Samtidig som fremhevingen vil kunne svekke den systembaserte tillitsdimensjonen *funksjonalitet*, og den menneskebaserte tillitsdimensjonen *evne*.

Dette fører oss til følgende hypoteser:

Systemtillit

H₃₋₁: Fremheving av at ChatGPT kan ta feil gjør at tilliten styrkes, og at brukerne skårer ChatGPT høyere på a) hjelpsomhet og b) pålitelighet, men lavere på c) funksjonalitet.

H_{3-1null}: Fremheving av at ChatGPT kan ta feil påvirker ikke tilliten, og brukerne skårer ChatGPT med og uten advarsel likt på funksjonalitet, hjelpsomhet og pålitelighet.

Mennesketillit

H₃₋₂: Fremheving av at ChatGPT kan ta feil gjør at tilliten styrkes, og at brukerne skårer ChatGPT høyere på a) velvilje og b) integritet, men lavere på c) evne.

H_{3-2null}: Fremheving av at ChatGPT kan ta feil påvirker ikke tilliten, og brukerne skårer ChatGPT med og uten advarsel likt på evne, velvilje og integritet.

4.4 Hypotesesett 4: Brukserfaring sin påvirkning på tillit

For hypotesesett 4 har vi valgt å inkludere to variabler som forklarer brukerens erfaring med et informasjonsverktøy. Vi undersøker hvordan erfaring med ChatGPT, Google og Wikipedia påvirker tillit, ved å se på variablene *brukstid* og *brukshyppighet*.

Brukstid måles ved å spørre respondentene om hvor lang tid de har brukt informasjonsverktøyet (se Vedlegg: Spørreundersøkelsen). Forskning viser at når et individ begynner å bruke en teknologi, kan tillit til teknologien påvirke om de ønsker å fortsette og bruke den (Hernández-Ortega, 2011). Grunnen til dette er at individet etter første bruk av en teknologi danner seg en mening om teknologien. Denne meningsoppfatningen kan enten bekrefte tidligere forventninger og øke tilliten, eller motsi dem og føre til lavere tillit (Karahanna, et al., 2003). Hvis brukerens oppfatning stemmer overens med forventningene og skaper et positivt tillitsnivå, vil de fortsette å bruke teknologien. Hvis ikke, kan dette føre til en negativ tillitsopplevelse og brukeren kan ende opp med å slutte og bruke verktøyet. Dette indikerer at brukstid påvirker tillit, og at dette tillitsforholdet bygger på individets første oppfatning av verktøyet.

Jo lenger tid et individ har brukt et verktøy, desto mer kjent blir individet med det, og kjennskap til et objekt (i dette tilfelle informasjonsverktøy) har vist seg å ha en positiv effekt på tillit (Gefen, et al., 2011). Dette tyder på at dersom individet kjenner til informasjonsverktøyet, kan dette øke tilliten. Økt brukstid kan også indikere at individet har etablert et positivt tillitsforhold til teknologien og har valgt å fortsette og bruke den. Dette gir grunnlag for å formulere en hypotese om at økt brukstid øker tillit.

For *brukshyppighet* måler vi hvor ofte respondentene bruker informasjonsverktøyet (vedlegg: Spørreundersøkelsen). Vi har valgt å undersøke brukshyppighet fordi det kan antas at de som bruker et verktøy regelmessig, nettopp bruker det ofte fordi de anser det som nyttig og verdifullt, noe som kan indikere høyere tillit til verktøyet. Jo mer man bruker et verktøy, desto sterkere tilknytning kan man få, og jo høyere kan byttekostnadene være. Et lignende mønster

kan observeres når vi for eksempel velger flyselskap – vi velger gjerne det vi har hatt gode erfaringer med tidligere.

Ifølge Glikson og Woolley (2020) er tilliten til virtuelle KI-er, som ChatGPT, ofte høy i starten, men reduseres med økende erfaring. Denne reduksjonen i tillit kan skyldes mangel på tillit til verktøyets pålitelighet (Glikson & Woolley, 2020). En mulig årsak er at brukeren opplever at ChatGPT gir feilaktige svar ved flere anledninger, og dermed innser at verktøyet kan gjøre feil.

Vi har ikke funnet studier som indikerer at økt bruk reduserer tilliten til Google og Wikipedia. Ettersom disse også er informasjonsverktøy, er det likevel rimelig å anta at tidligere argumentasjon om at økt brukserfaring øker tillit, også gjelder for disse plattformene.

Ettersom vi har funnet studier som viser at økt brukshyppighet kan redusere tillit til KI-er, ønsker vi å undersøke om dette også gjelder for ChatGPT. Vi vil derfor se på hvordan tilliten påvirkes av økt brukshyppighet av verktøyet. Reduksjonen i tillit kan i stor grad skyldes et fall i tilliten til verktøyets pålitelighet (Glikson & Woolley, 2020). På samme måte som for brukstid, finner vi ikke studier som tyder på at det er en negativ effekt mellom brukshyppighet og tillit for Google- og Wikipedia-brukere. Dette kan være naturlig da Google for mange er den standardiserte og foretrukne søkemotoren.

Basert på argumentasjonen ovenfor, har vi inkludert følgende hypoteser knyttet til brukserfaring, både med hensyn til brukstid og brukshyppighet:

Systemtillit

H₄₋₁: Lengre bruk av verktøyet vil ha en negativ effekt på de systembaserte tillitsdimensjonene for a) ChatGPT, men positiv for b) Google og c) Wikipedia.

H₄₋₂: Hyppigere bruk av verktøyet vil ha en negativ effekt på de systembaserte tillitsdimensjonene for a) ChatGPT, men positiv for b) Google og c) Wikipedia.

Tilhørende nullhypoteser:

H_{4-1null}: Lengre bruk av verktøyet har ingen effekt på de systembaserte tillitsdimensjonene for ChatGPT, Google og Wikipedia.

H_{4-2null}: Hyppigere bruk av verktøyet har ingen effekt på de systembaserte tillitsdimensjonene for ChatGPT, Google og Wikipedia.

Mennesketillit

H₄₋₃: Lengre bruk av verktøyet vil ha en negativ effekt på de menneskelige tillitsdimensjonene for a) ChatGPT, men positiv for b) Google og c) Wikipedia.

H₄₋₄: Hyppigere bruk av verktøyet vil ha en negativ effekt på de menneskelige tillitsdimensjonene for a) ChatGPT, men positiv for b) Google og c) Wikipedia.

Tilhørende nullhypoteser:

H_{4-3null}: Lengre bruk av verktøyet har ingen effekt på de menneskelige tillitsdimensjonene for ChatGPT, Google og Wikipedia.

H_{4-4null}: Hyppigere bruk av verktøyet har ingen effekt på de menneskelige tillitsdimensjonene for ChatGPT, Google og Wikipedia.

4.5 Hypotesesett 5: Generell tilbøyelighet til å stole på nye informasjonsverktøy sin påvirkning på tillit

Hypotesesett 5 består av hypoteser som undersøker hvordan individers generelle tilbøyelighet til å stole på nye teknologier kan påvirke deres tillit til informasjonsverktøyene. Det er vist at tilbøyelighet til å stole på et nytt verktøy har stor effekt på tillit (Karahanna, et al., 2003; Li, et al., 2008), spesielt ved første møte med en teknologi (Gefen, et al., 2011). Individer som generelt er positive til nye teknologier, har ofte høy tillitstilbøyelighet (Zhou, 2011). Dette tyder på at respondentens initiale tilbøyelighet til å stole på nye teknologier har en faktisk effekt på tilliten de utvikler til et informasjonsverktøy.

Det er vist at jo mer tilbøyelig et individ er til å stole på en teknologi, desto høyere tillit har individet til KI-systemer (Küper & Krämer, 2024). Dette tyder på at grad av tillitstilbøyelighet kan påvirke tillit til ChatGPT, som er et KI-verktøy. På den andre siden kan effekten av tilbøyelighet til å stole på nye teknologier avta med erfaring (Zahedi & Song, 2008). Dette antyder at effekten av tillitstilbøyelighet kan være svakere dersom respondenten har mye erfaring med informasjonsverktøyet.

Basert på de ovennevnte argumentene, er det naturlig å anta at generell tilbøyelighet til tillit kan ha en positiv effekt på tillit til et informasjonsverktøy. Derfor undersøkes følgende hypotese:

Systemtillit

H₅₋₁: Desto mer tilbøyelig brukeren er til å stole på nye informasjonsverktøy jo sterkere positiv effekt vil det ha på de systembaserte tillitsdimensjonene for a) ChatGPT, b) Google og c) Wikipedia.

Tilhørende nullhypotese:

H_{5-1null}: Brukerens tilbøyelig til å stole på nye informasjonsverktøy har ingen effekt på de systembaserte tillitsdimensjonene for ChatGPT, Google og Wikipedia.

Mennesketillit

H₅₋₂: Desto mer tilbøyelig brukeren er til å stole på nye informasjonsverktøy jo sterkere positiv effekt vil det ha på de menneskelige tillitsdimensjonene for a) ChatGPT, b) Google og c) Wikipedia.

Tilhørende nullhypotese:

H_{5-2null}: Brukerens tilbøyelig til å stole på nye informasjonsverktøy har ingen effekt på de menneskelige tillitsdimensjonene for ChatGPT, Google og Wikipedia.

4.6 Hypotesesett 6: Oppfattet grad av menneskelignende egenskaper sin påvirkning på tillit

I hypotesesett 6 utforsker vi hvordan opplevd grad av menneskelighet påvirker tilliten til ChatGPT, Google og Wikipedia. Med kunstig intelligens har vi fått teknologier som i større grad har menneskelignende egenskaper, som for eksempel ChatGPT (Jackson, 2023). En studie viste at ChatGPT har evnen til å være med samme emosjon som brukerens prompt med 70,7% nøyaktighet (Schaaff, et al., 2023). Videre antas det at menneskelignende egenskaper og adferd øker tillit (Glikson & Woolley, 2020). Dette gir grunn til å undersøke om brukernes oppfatning av ChatGPTs menneskelignende egenskaper påvirker tilliten til verktøyet.

Vi antar at respondentene vil vurdere ChatGPT som mer menneskelig enn Google og Wikipedia, ettersom ChatGPT er designet for å etterligne menneskelig kommunikasjon. Den genererer tekst på en naturlig måte, uten den formaliteten som kjennetegner verktøy som eksempelvis Wikipedia, som benytter et akademisk språk. ChatGPTs setningsstruktur og ordvalg kan gi et inntrykk av menneskelig kommunikasjon, og den kan også tilpasse seg konteksten. Dette kan gi brukeren et inntrykk av at ChatGPT forstår brukerens følelser, selv

om den i realiteten bare analyserer mønster i teksten. Google og Wikipedia mangler derimot denne menneskelige kommunikasjonsfunksjonen. I tillegg har hverken Wikipedia eller Google evne til å uttrykke empati og sosial intelligens. ChatGPT kan fremstå som empatisk ved å bruke uttrykk som «Jeg forstår hvordan du føler deg». Selv om dette er forhåndsprogrammert, kan det gi inntrykk av ektefølt forståelse, et trekk som kjennetegner menneskelig interaksjon.

ChatGPT er programmert til å følge sosiale normer, som å være høflig, unngå fornærmelser og svare på en hjelpsom måte. Dette samsvarer med hvordan mennesker vanligvis oppfører seg, og styrker inntrykket av at ChatGPT kan oppfattes som mer menneskelig. Den kan også tilpasse tonen etter brukerens ønske, for eksempel være uformell, noe som forsterker følelsen av menneskelig interaksjon. I tillegg kan ChatGPT huske tidligere deler av en samtale og referere til denne, noe som gir inntrykk av at maskinen har samme hukommelse som et menneske. Overordnet gir alle disse argumentene grunnlag for utforming av følgende hypotese:

H₆₋₁: Brukerne vil oppfatte ChatGPT som mer menneskelig enn a) Google og b) Wikipedia.

Nullhypotesen er følgende:

H_{6-1null}: Brukerne vil ikke oppfatte ChatGPT som mer menneskelig enn Google og Wikipedia.

På grunnlag av at vi har en hypotese om at ChatGPT oppfattes som mer menneskelig enn Google og Wikipedia, og at Glikson & Woolley (2020) viser at tilliten øker med økende grad av menneskelignende egenskaper, ønsker vi å undersøke følgende hypoteser:

Systemtillit

H₆₋₂: Økt grad av oppfattet menneskelighet vil øke systemtillit for a) ChatGPT, b) Google og c) Wikipedia.

H₆₋₃: Effekten av oppfattet menneskelighet på systemtillit vil være større for ChatGPT enn for a) Google og b) Wikipedia.

Tilhørende nullhypoteser:

H_{6-2null}: Økt grad av oppfattet menneskelighet har ingen påvirkning på systemtillit for ChatGPT, Google og Wikipedia.

H_{6-3null}: Effekten av oppfattet menneskelighet på systemtillit er lik for ChatGPT, Google og Wikipedia.

Mennesketillit

H₆₋₄: Økt grad av oppfattet menneskelighet vil øke mennesketillit for a) ChatGPT, b) Google og c) Wikipedia.

H₆₋₅: Effekten av oppfattet menneskelighet på mennesketillit vil være større for ChatGPT enn for a) Google og b) Wikipedia.

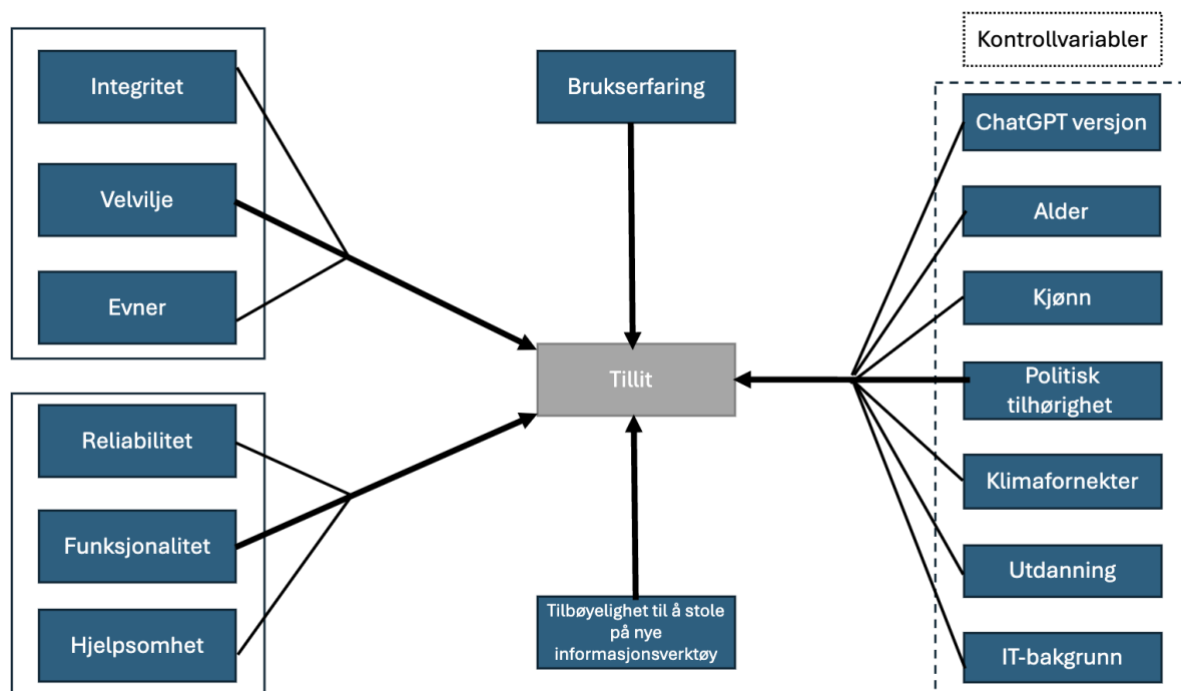
Tilhørende nullhypoteser:

H_{6-4null}: Økt grad av oppfattet menneskelighet har ingen påvirkning på mennesketillit for ChatGPT, Google og Wikipedia.

H_{6-5null}: Effekten av oppfattet menneskelighet på mennesketillit er lik for ChatGPT, Google og Wikipedia.

4.7 Forskningsmodell

Forskningsmodellen, som vist under, illustrerer hypotesene overordnet. Den viser at vi ønsker å avdekke hvordan både mennesketillitsdimensjonene (*evne, velvilje og integritet*) og systemtillitsdimensjonene (*funksjonalitet, hjelpsomhet og pålitelighet*) påvirker tillit. Modellen tar også for seg om brukserfaring påvirker tilliten til informasjonsverktøyet, både i form av hvor *lenge* og hvor *hyppig* individet benytter det. I tillegg viser modellen at vi undersøker om det er sammenheng mellom generell tilbøyelighet til å stole på nye informasjonsverktøy og tillit. I modellen har vi også tatt hensyn til kontrollvariabler da vi ønsker å se på om tillit påvirkes av andre variabler.



Figur 1: Oppgavens forskningsmodell.

5. Metode

5.1 Forskningsdesign

Forskningsdesign er den overordnede planen for å besvare et forskningsspørsmål, og inkluderer tilnærming, valg av forskningsmetode og utforming av forskningsdesignet (Saunders, et al., 2019).

Vårt forskningsdesign er forklarende, fordi vi undersøker om det er kausale sammenhenger mellom variabler (Saunders, et al., 2019), samt bekrefte eller avkrefte flere sett med hypoteser. Vår tilnærming er deduktiv, da vi søker å gå fra det generelle, teori, og til det spesifikke, empiriske data (Tranøy, 2024). En deduktiv tilnærming er ofte preget av høy grad av struktur. Formålet er å forklare eller generalisere, og det er derfor krav til god ekstern validitet. Hypotesene våre er utledet på bakgrunn av eksisterende teori, og vi skal teste disse gjennom innsamling og analyse av primærdata.

5.2 Forskningsstrategi

Forskningsstrategi beskriver den spesifikke tilnærmingen som brukes for å svare på forskningsspørsmålet (Saunders, et al., 2019). Dette inkluderer blant annet valg av forskningsmetode. Vi har valgt å benytte eksperiment. Eksperiment er en kvantitativ

forskningsmetode som søker å påvise hvordan en eller flere uavhengige variabler påvirker en avhengig variabel gjennom at forskeren manipulerer de uavhengige variablene (Svartdal & Grønmo, 2024).

Grunnen til at vi har valgt eksperiment som forskningsmetode er fordi eksperiment gjør det mulig å etablere årsakssammenhenger, og etablere kausalitet (Shadish, et al., 2002). Bruk av eksperiment fører også til at vi kan eliminere effekten av forstyrrende variabler som kan påvirke effekten på den avhengige variabelen (Campbell & Stanley, 1963). For eksempel kan sosial bias oppstå i andre forskningsmetoder, som intervju, hvor deltakerne kan gi svar de tror intervjueren ønsker. I et eksperiment, derimot, kan vi undersøke effekten av hvert informasjonsverktøy mer isolert og kontrollert.

Eksperiment som forskningsmetode er gunstig for videre forskning, da de kan kopieres av andre forskere (Montgomery, 2017). Dette er spesielt relevant fordi studien vår gir et øyeblikksbilde av dagens tillit til ChatGPT, Google og Wikipedia, noe som kan skape et behov for å gjenta studien i fremtiden, for å fange opp eventuelle endringer over tid.

Vårt design er «between-subject», som vil si at vi har målt tillit mellom fire ulike grupper som kun utsettes for én behandling (Charness , et al., 2012), som i denne studien er et informasjonsverktøy. «Between-subject»-analyse er gunstig fordi det er konservativt, ofte har høy grad av ekstern validitet, og er statistisk enkelt å gjennomføre. (Charness , et al., 2012)

En annen fordel med «between-subject» er at det ikke forekommer «carryover effects» (Plake & Wise , 2013), altså at effekten av én betingelse i eksperimentet påvirker deltakerens atferd eller respons ved senere betingelser i samme eksperiment (Brooks., 2012). Dette kan oppstå i «within-subject»-design, hvor deltakerne kan bli påvirket av tidligere erfaringer i eksperimentet som de potensielt kan «lære av» (Greenwald, 1976). Dette unngår vi ettersom våre respondenter kun utsettes for én behandling.

I tillegg er «between-subject» fordelaktig fordi vi unngår hypotesegjetting. Hypotesegjetting handler om at respondenten forsøker å forstå formålet med eksperimentet, og på grunnlag av dette endrer atferd (Lund, et al., 2002). Atferdsendringen kan medføre at hypotesene blir feilaktig bekreftet eller avkreftet, og er en risiko ved «within-subject»-design.

Dersom vi ønsket høyest mulig statistisk styrke kunne vi gjennomført en «within-subject»-analyse (Charness , et al., 2012), der vi først kartla tilliten til de tre aktørene ChatGPT, Google og Wikipedia uten en advarsel, for deretter å introdusere et nytt design med en advarsel til de

samme respondentene. Dette ville krevd betydelig mer tid og ressurser å gjennomføre enn «between-subject».

Alternativt kunne vi gjennomført eksperimentet med seks ulike eksperimentgrupper og sammenlignet tillitsnivået mellom de tre aktørene med og uten advarsel om at de kan ta feil. Dette er også en «between-subject»-analyse. Å gjennomføre et slikt eksperiment ville vært gunstig for å se om det er noen forskjeller i tillit til respondenter dersom de blir oppmerksomme på at alle informasjonsverktøyene kan ta feil, og ikke bare ChatGPT. På grunn av kostnadsnivå og tidsbegrensning var vi nødt til å velge bort dette eksperimentet. Likevel kan det være nyttig for fremtidig forskning å se på et slikt eksperiment.

I vårt eksperiment skal vi sammenligne tillitsnivået til ChatGPT med tilliten til Google, og tilliten til ChatGPT med tilliten til Wikipedia. I tillegg skal vi forske på om det er forskjeller i tillit til ChatGPT med en synliggjøring av at verktøyet kan ta feil sammenlignet med ChatGPT uten slik synliggjøring.

5.3 Utforming av eksperimentet

Eksperimentet er utført som en spørreundersøkelse og vi benyttet Qualtrics for å utforme spørreundersøkelsen (Vedlegg: Spørreundersøkelsen). Qualtrics er en plattform for data- og innsamlingsanalyser, som ofte benyttes til å lage spørreskjema. Plattformen er mulig å benytte dersom man ønsker å utforme ulike scenarioer, og kan også sørge for at respondentene randomiseres til ulike eksperimentgrupper. Manipulasjonen i vårt eksperiment er hvilken eksperimentgruppe, altså hvilket informasjonsverktøy respondentene bli randomisert til.

For å kartlegge tillitsforskjeller mellom de avhengige variablene, deler vi respondentene inn i fire ulike eksperimentgrupper. Alle gruppene skal svare på identiske spørsmål tilknyttet tillit, hvor vi både har inkludert menneskerelaterte- og systemrelaterte tillitsdimensjoner. Dette er for å utforske hvorvidt de ulike dimensjonene kan påvirke tillitsnivået respondenten har til ChatGPT, Google, Wikipedia og ChatGPT med advarsel.

Vi valgte bevisst å ikke spesifisere om deltakerne skulle forholde seg til ChatGPT 3.5 (gratisversjonen) eller ChatGPT 4 (betalt versjon). Dette ble gjort for å unngå at modellforskjeller påvirket resultatene. For sikkerhets skyld inkluderte vi et spørsmål innledningsvis for å kartlegge hvilken versjon respondentene benytter seg av, slik at vi kunne teste om dette hadde en effekt på resultatene.

Som nevnt blir deltakerne tilfeldig plassert i én av fire grupper, med jevn tilordning slik at vi får like mange respondenter i hver gruppe. Det er like stor sannsynlighet for å komme i hver av de fire gruppene. Denne randomiseringen sørger for at gruppene er tilnærmet like og at det ikke er betydelige statistiske forskjeller mellom individene i de ulike gruppene. Som nevnt, får hver av gruppene samme spørsmål, slik at det eneste som skiller gruppene er hvilket informasjonsverktøy de tildeles. På denne måten isolerer vi andre ytre variabler som kan påvirke tillit, samt at det er så få forskjeller mellom gruppene som mulig.

Alle gruppene skal besvare spørsmål om tillit i et scenario der de skal forestille seg at de søker informasjon om klimaendringer. Deltakerne skal altså ikke faktisk innhente informasjon. Et alternativ kunne vært å vise deltakerne en skjermdump av hvordan ChatGPT, Google og Wikipedia presenterer informasjon om klimaendringer. Dette alternativet ble derimot utelatt for å unngå mulige forstyrrende variabler, slik som design og layout. Dette fordi disse kunne vært med på å påvirke respondentenes oppfattede tillit til plattformene.

Klimaendringer ble valgt som tema, fordi det krever grundig og pålitelig informasjon, samtidig som det er et aktuelt og kjent tema for deltakerne. Dette fordi det stadig omtales i media, diskuteres i arbeidslivet, og er noe flere har en formening om. Derfor kan deltakerne enkelt forestille seg at de skal innhente informasjon om klimaforandringene, noe som er en av grunnene bak dette valget av tema.

Gruppen som blir eksponert for ChatGPT med advarsel advares to ganger om at ChatGPT selv informerer om at den kan ta feil. For å gjøre advarselen så naturlig som mulig, har vi tatt en skjermdump av brukergrensesnittet til ChatGPT og markert området hvor det presiseres at «*ChatGPT can make mistakes. Check important information*» (Vedlegg: Advarsel). Dette er for å synliggjøre at verktøyet ikke er ufeilbart. Gjennom denne synliggjøringen ønsker vi å avdekke om det kan påvirke tilliten respondentene har til ChatGPT.

En synliggjøring av at ChatGPT kan ta feil kan forstås som en form for «salience effect», fordi det trekker respondentenes oppmerksomhet mot en spesifikk egenskap ved teknologien, nemlig at den har potensial for feil. «Salience effect» refererer til en kognitiv skjevhet der mennesker legger uforholdsmessig mye vekt på informasjon som er spesielt fremhevet (Newristics, u.d.).

For å avdekke om det er forskjeller mellom menneske- og systemtillitsdimensjonene inkluderer eksperimentet et spørsmål om hvorvidt respondenten ser på de tre ulike informasjonsverktøyene som menneskelig eller systemlignende. Disse er hentet fra Lankton et

al. (2015) (vedlegg: Spørreundersøkelsen). Dette vil gi oss et tydelig bilde på hvilke tillitsegenskaper som er mest relevante å benytte for de ulike informasjonsverktøyene. Bruk av korrekt begrepsbruk for enten menneske eller systemtillit kan føre til mer valide resultater (Tripp, 2015).

For å måle mennesketillitsdimensjonene har vi benyttet validerte spørsmål utformet av McKnight et al. (2002b). For hver av de tre tillitsdimensjonene har vi benyttet tre spørsmål som skal fange essensen av dimensjonen, identisk med McKnight et al. sin tilnærming (vedlegg: Spørreundersøkelsen). For å måle systemtillitsdimensjonene har vi benyttet spørsmål utformet av McKnight et al. (2011). Ettersom tillit er et komplekst begrep, og det er mange faktorer som kan påvirke et individs tillitsnivå, ønsker vi å kartlegge om deltakernes tilbøyelighet til å stole på nye teknologier kan påvirke resultatet. Derfor har vi inkludert tre spørsmål fra McKnight et al. (2002b) som kartlegger deltakernes generelle tilbøyelighet til å stole på nye informasjonsteknologier.

I tillegg ønsker vi å finne ut hvor ofte deltakerne bruker de ulike verktøyene, da det kan være lav grad av kjennskap til verktøyet påvirker tillitsnivå. Følgelig har vi inkludert to spørsmål som avdekker bruk av det gitte informasjonsverktøyet (vedlegg: Spørreundersøkelsen). Disse er inspirert av Kapania et al. (2022).

For å sikre kvaliteten på innsamlede data i studier er det vanlig å inkludere en eller flere «attention checks» (Kung, et al., 2017). Dette er spesifikke spørsmål eller oppgaver designet for å evaluere om deltakerne gir tilstrekkelig oppmerksomhet til spørsmålene. Hensikten med «attention checks» er todelt. For det første bidrar de til å filtrere ut deltakere som svarer tilfeldig eller uoppmerksomt, noe som kan svekke validiteten til resultatene (Kurtz & Parrish, 2001). For eksempel kan rastløse respondenter svare uoppmerksomt med lite fokus på spørsmålene, noe som kan føre til ugyldige resultat (Johnson, 2005). Respondentene kan for eksempel trykke på samme svaralternativ for alle spørsmål for å bli raskt ferdig med undersøkelsen, eller svare i feil retning, for eksempel svare 1 (helt uenig) fremfor 5 (helt enig). Dersom slike svar forblir i datasettet, kan det føre til støy som reduserer vår evne til å identifisere reelle effekter. For det andre kan bruken av «attention checks» fungere som en indirekte motivator ved å minne deltakerne om viktigheten av å følge nøye med.

I vårt eksperiment benyttet vi et spørsmål med en spesifikk instruksjon som «attention check» (vedlegg: Spørreundersøkelsen). Vi valgte å ha denne på starten av spørreundersøkelsen for å sikre at deltakerne ikke bare svarte mekanisk. Data fra deltakere som ikke besto «attention

check» ble ekskludert fra videre analyse, i tråd med forskningsetiske retningslinjer og best praksis innen eksperimentell metode. Ved å inkludere en «attention check» styrker vi dermed både påliteligheten og gyldigheten av funnene våre. Dette er avgjørende for å trekke robuste konklusjoner om tillit mellom de ulike informasjonsverktøyene.

Vi ønsket at alle respondentene skulle ta stilling til alle spørsmålene, og valgte derfor å ha tvungen respons aktivert i Qualtrics. Dette fordi ingen av spørsmålene var å regne som sensitive eller svært vanskelig for respondentene å skulle ta stilling til.

5.4 Datainnsamling

I denne studien har vi benyttet Profilic for datainnsamling. Profilic er et teknologiselskap som muliggjør en samling av høykvalitets menneskegenerert data i stor skala (Profilic, u.d.). Selskapet har verifiserte og mangfoldige respondenter, og gjør det mulig for forskere å sikre troverdige data. Plattformen har over 200.000 aktive deltakere, og det er mer enn 35.000 forskere og 10.000 organisasjoner som benytter Profilic (Profilic, u.d.).

I Profilic blir respondentene betalt for å gjennomføre undersøkelsene, noe som kan være fordelaktig ettersom respondentene kan ha et insentiv til å svare gjennomtenkt og grundig. I tillegg kan betalingen føre til at respondentene føler seg mer engasjert til å delta, da vi viser at vi anerkjenner deres innsats og tid gjennom kompensasjonen. Deltakerne på plattformen er også ofte erfarne, noe som kan resultere i at vi får mer pålitelige svar.

Fordelen med Profilic er at man kan selektere respondenter basert på en rekke ønskede kriterier. Vi forklarer hvilke kriterier vi selekterte på ytterligere i seksjonen under. Ulempen ved bruk av Profilic er at deltakerne trolig er teknologisk kompetente, da de må ha grunnleggende teknologiske ferdigheter for å opprette en profil og delta på spørreundersøkelser. Dette kan potensielt føre til skjevheter i dataene.

5.4.1 Utvalg

Et utvalg er en liten gruppe av en større populasjon (Fink, 2003). I vitenskapelige studier er det ønskelig at utvalget er representativt, for å kunne generalisere potensielle funn til populasjonen. Et representativt utvalg reflekterer karakteristikene til hele populasjonen, for eksempel ved å ha en balansert kjønnsfordeling og jevn aldersfordeling. Vårt ønske er å generalisere til en så stor populasjon som mulig, men da vi har avgrenset utvalget vårt til å kun være amerikanere, vil vår populasjon naturligvis være USA.

I Profilic avgrenset vi utvalget til å gjelde amerikanere. Omtrent 68 millioner bruker ChatGPT i USA (Backlinko Team, 2024), noe som gjør det spesielt interessant å gjennomføre eksperimentet vårt på ChatGPT-brukere i USA.

For å sikre troverdige respondenter begrenset vi utvalget til å kun inkludere individer som (1) har høy godkjenningsgrad på Profilic, (2) har fullført minst 100 studier tidligere, (3) har engelsk som morsmål, og (4) har benyttet ChatGPT tidligere. (4) Var for å sikre at deltakerne er kjent med verktøyet da de kan havne i eksperimentgruppen med ChatGPT som informasjonsverktøy, både for ChatGPT-gruppen og for gruppen med ChatGPT som eksponeres for en advarsel. I Profilic brukte vi et kriterium som sikrer et kjønnsbalansert utvalg. Deltakerne ble tildelt 0,65 pund for deltakelsen.

5.4.2 Pilotstudie

En pilotstudie er nyttig å gjennomføre før en faktisk studie, fordi den bidrar til å sikre at studien er grundig, gjennomførbar og økonomisk forsvarlig (Arnold, et al., 2009). Hovedformålet med pilotstudier er å forhindre at man iverksetter en storskala studie uten tilstrekkelig kunnskap om metoden er velfungerende (Lowe, 2019).

For å teste eksperimentet vårt gjennomførte vi først en pilotstudie med 200 respondenter. Dette var for å teste ut designet i en mindre skala, for å forsikre oss at dette var velfungerende før vi distribuerte undersøkelsen til et større utvalg. Ved å gjennomføre denne pilotstudien, åpnet vi også opp for å få tilbakemeldinger fra de 200 respondentene. For eksempel kan det tenkes at de ville sendt en melding dersom noe var uklart i studien, noen av spørsmålene ikke var mulig å svare på eller kronglete formulert. Vi mottok ingen tilbakemeldinger fra pilotstudien, og kunne dermed distribuere studien på nytt til et større utvalg.

5.4.3 Praktisk gjennomføring

Etter pilotstudien var gjennomført åpnet vi opp samme undersøkelse i Profilic, og sendte ut eksperimentet til 790 individer. Disse respondentene var fra samme utvalg som den første runden, men vi sørget for at respondenter fra pilotrunden ikke kunne ta undersøkelsen på nytt. Samlet endte vi med et utvalg bestående av 990 respondenter.

5.5 Variabler

5.5.1 Uavhengig variabel: Grad av menneskelignende egenskaper

En uavhengig variabel i et eksperiment er en variabel som forskeren manipulerer for å undersøke om den har effekt på den avhengige variabelen (Dahlum, 2023). Vårt eksperiment

har en uavhengig variabel som måler i hvilken grad respondentene oppfatter informasjonsverktøyene for å ha menneskelignende versus systemlignende egenskaper. Dette spørsmålet ble gitt til alle respondentene i eksperimentet uavhengig av hvilken gruppe de var del av. Spørsmålet er basert på Lankton et al. (2015) sine etablerte spørsmål som undersøker individers oppfatning av objekter som mer eller mindre menneskelignende.

5.5.2 Uavhengig variabel: Generell tilbøyelighet til å stole på nye teknologier

En annen uavhengig variabel i vårt eksperiment er respondentenes generelle tilbøyelighet til å stole på nye teknologier. Dette ble målt ved å benytte spørsmål utformet av Kapania et al. (2022).

5.5.3 Uavhengig variabel: Brukserfaring

Vi måler også brukserfaring som uavhengig variabel. Med brukserfaring henviser vi både til brukstid, altså *hvor lenge* respondenten har benyttet det aktuelle informasjonsverktøyet, samt brukshyppighet som handler om *hvor ofte* respondenten benytter verktøyet. Disse spørsmålene er inspirert fra Kapania et al. (2022).

5.5.4 Uavhengig- og interaksjonsvariabel: Gruppetilhørighet

Gruppetilhørighet, altså hvilket informasjonsverktøy respondentene tildeles, er både en uavhengig variabel og en interaksjonsvariabel. Gruppetilhørighet er en uavhengig variabel fordi det er noe vi manipulerer for å undersøke forskjeller i tillit mellom de ulike informasjonsverktøyene.

En interaksjonsvariabel er to eller flere uavhengige variabler som sammen påvirker en avhengig variabel (Jaccard & Turrisi, 1997). En interaksjonseffekt oppstår når effekten av én uavhengig variabel på den avhengige variabelen endres av nivået på en annen uavhengig variabel. Gruppetilhørighet inngår som interaksjonsvariabel dersom vi gjennomfører regresjonsanalyser, hvor vi ser på samspillet mellom gruppetilhørighet og en annen uavhengig variabel, som eksempelvis brukshyppighet

5.5.5 Avhengig variabel: Tillit

I et eksperiment måler forskeren avhengig variabel for å se hvordan den påvirkes av endringer i en eller flere uavhengige variabler (Grønmo & Dahlum, 2024). Vår avhengige variabel er tillit. Dette målte vi ved å stille respondentene tre spørsmål for hver av de ulike tillitsdimensjonene for både mennesketillitsdimensjonene og systemtillitsdimensjonene (vedlegg: Spørreundersøkelsen).

Variablene tilknyttet de ulike tillitsdimensjonene ble målt ved bruk av en syvpunkts Likert-skala, med svaralternativ fra «Strongly disagree» til «Strongly agree». En syvpunktsskala er hensiktsmessig for å avdekke variasjon i respondentenes svar, samtidig som det forhindrer for mange alternativer å velge mellom. Alle spørsmålene tilknyttet tillitsdimensjonene besto av syv svaralternativer, for slik å være konsistent i spørreundersøkelsen, noe som er i samsvar med Saunders et al. (2019).

Spørsmålene tilknyttet den avhengige variabelen er alle hentet fra litteraturen og tidligere forskning. Vi benyttet spørsmål utformet av McKnight et al. (2002) for mennesketillitsdimensjonene, og spørsmål fra McKnight et al. (2011) for systemtillitsspørsmålene (vedlegg: Spørreundersøkelsen).

5.5.6 Kontrollvariabler

Vi har samlet inn data om alder, kjønn, IT-bakgrunn, utdanningsnivå og politisk tilhørighet, for å kontrollere deres innvirkning på resultatene. Dette er for å kontrollere at disse variablene ikke har en sterkere effekt på tillitsnivået til respondentene enn brukserfaring og generell tilbøyelighet til å stole på nye informasjonsverktøy. En metastudie over tidligere forskning på tillit, kunne ikke med sikkerhet fastslå om kjønn og alder hadde stor innvirkning på tillit, og oppfordret derfor at disse variablene burde bli testet grundigere opp mot tillit i andre studier (Hancock, et al., 2023). Ettersom tidligere studier er tvetydige i resultatene sine, ønsker vi å teste kontrollvariablene opp mot den avhengige variabelen vår.

Alder

Eldre mennesker stoler mer på automatisering (McBride, et al., 2011), og kan derfor tenkes å stole mer på ChatGPT, da dette er en algoritmedrevet plattform som har en rekke funksjoner som kan knyttes opp mot automatikk. I tillitsforhold mellom en ansatt og overordnet har alder hatt en homogenitet i effektstørrelsen, noe som indikerer at alder har en stabil og konsistent påvirkning på tillitsnivået mellom to parter (Hancock, et al., 2023). På den andre siden kan det tenkes eldre mennesker ikke anvender ny teknologi i like stor grad som yngre (Smith, 2014), og har mindre teknologisk kompetanse (Vaportzis, et al., 2017). Dette illustrerer spredning i forskningsfeltet når det gjelder alder sin effekt på tillit.

Kjønn

Irwin et al. har vist gjennom sin forskning at kvinner er mindre tillitsfulle enn menn (2015). Kvinner rapporterte i denne studien et generelt lavere tillitsnivå enn menn. I tillegg er det vist

at kvinner er mindre tillitsfulle enn menn i begynnelsen av et tillitsforhold, men at denne effekten snur når kvinnene først stoler på andre (Alexander, et al., 2018).

Når det gjelder tillit til teknologi, er likevel forskningen spredt (Hoff & Bashir, 2015). En studie har funnet at kvinner i større grad aksepterte KI-systemer sammenlignet med menn. En metastudie av tillit har derimot vist at forholdet mellom kjønn og generell tillit ikke var signifikant (Hancock, et al., 2023). Forskerne i denne metaanalysen understrekte likevel behovet for ytterligere testing av kjønnsforskjeller i tillit, da deres metaanalyse manglet tilstrekkelig datagrunnlag til å kunne beregne kjønnsvariablenes potensielle påvirkning.

Politisk tilhørighet

Det er vist at det er en sammenheng mellom sosial tillit (tillit til andre mennesker og institusjoner) og politisk deltakelse (Burns, et al., 2003). Dette tyder på at politisk tilhørighet korrelerer med tillitsnivå, og kan være en variabel som påvirker tillit. Mer spesifikt relatert til amerikanere, har forskning vist at tillit blant republikanere, spesielt når det gjelder tillit til vitenskapen har hatt en nedgang (Lee, 2021). Demokrater har på sin side hatt en generell økning i tillit til vitenskap de siste årene. Dette tyder på at politisk tilhørighet har en påvirkning på tillit, spesielt når det gjelder vitenskapelige tema, slik som eksempelvis klimaendringer.

Syn på klimaendringene

Kontekst har vist seg å påvirke tillit (Cho, et al., 2015), og derfor vil vi teste i vår studie om syn på klimaforandringer kan påvirke tillit. Kontekst omhandler karakteristikken i oppgaven individet skal fullføre. Oppgavens karakteristikker handler om den objektive risikoen tilknyttet oppgaven, og risikoen som potensielt kan oppstå ved sluttresultatet.

Overordnet tema for eksperimentet vårt er klimaendringer, og vi vil derfor kontrollere for om respondentens syn på dette påvirker tillitsnivået til informasjonsverktøyet de tildeles. Det kan argumenteres for at oppgaven er å regne som lavrisiko, da respondentene kun får et imaginært oppdrag, og det er ingen konsekvenser av utfallet. Dette er fordi respondentene ikke faktisk skal innhente informasjon, men kun visualisere at de gjør dette. Det kan likevel tenkes at respondentens subjektive syn på klimaendringene kan påvirke tilliten de har, selv om oppgaven i seg selv er å anse som lavrisiko.

Utdanning og IT-bakgrunn

En rekke forskning har funnet at det er en klar positiv sammenheng mellom utdanningsnivå og tillit, hvor økt grad av utdanning har en positiv effekt (Hooghe, et al., 2012; Borgonovi, 2012;

Charron & Rothstein , 2016). Det er også vist at utdanning påvirker bruk av generativ KI, hvor de med lengre utdanning har større sannsynlighet til å benytte KI-en (Korzyński, et al., 2024). Dette kan indikere at individer med høy utdanning, også er mer tilbøyelige til å ta i bruk KI, noe som kan skyldes at de stoler på teknologien.

Vi spør også respondentene om de har IT-bakgrunn. Med IT-bakgrunn viser vi til individer som har utdanning i informasjonsteknologi, for eksempel innenfor datamaskiner, eller programmering (Utdanning.no, 2024). Disse individene kan tenkes å ha høy grad av teknologisk kompetanse og kunnskap om KI, noe som potensielt kan påvirke tilliten. Dette er fordi KI er en del av det teknologiske fagfeltet, og ofte inngår i pensum i utdanningsløpet (NTNU, u.d.).

ChatGPT 3.5 og ChatGPT 4

Grad av tillit til ChatGPT har vist seg å påvirke individers intensjoner om å bruke chatboten, men også faktisk bruk av den (Choudhury & Shamszare, 2023). Vi finner begrenset med forskning som skiller mellom ChatGPT 3.5 og ChatGPT 4, og hvorvidt de ulike versjonene påvirker tillit. Tidligere forskning finner derimot at det er store forskjeller mellom de to versjonenes funksjonalitet og nøyaktighet (Lim, et al., 2023). Med tanke på at de to versjonene har ulik funksjonalitet, hvorav gratisversjonen har færre funksjoner og er bygd på et mindre treningssett (Al-Ashwal, et al., 2023), ønsker vi å utforske om det er en forskjell i tillit til de to versjonene. Når et individ bruker penger på å få tilgang til en annen versjon, kan det tyde på at individet i større grad har tillit til KI, og stor grad av intensjon om å bruke KI-chatboten.

5.6 Databehandling og analysering av data

I kvantitativ forskning handler dataanalyse om å behandle dataene som er samlet inn, og deretter benytte de til å analysere og tolke funnene (Saunders, et al., 2019). I den neste seksjonen skal vi beskrive hvordan vi har bearbeidet dataene, og hvilke analyser som er anvendt.

For å behandle dataene fra Qualtrics, benyttet vi programmeringsverktøyet R. Etter vi hadde avsluttet eksperimentet i Profilic og Qualtrics, eksporterte vi dataene fra Qualtrics til Microsoft Excel og så over i R. Alle endringer i datasettet gjorde vi i R, noe som øker påliteligheten til dataene.

Før vi startet å analysere dataen kontrollerte vi datasettet for potensielle avvik. Totalt hadde vi 990 respondenter. Vi fjernet 7 respondenter som ikke besto «attention check», i tillegg til å

fjerne 11 respondenter som ikke hadde kjennskap til ChatGPT. Vi fjernet respondentene som ikke besto «attention check» fordi dette kan tyde på at de enten ikke har vært oppmerksomme under gjennomføringen av eksperimentet, eller at de har forhastet seg gjennom, slik at svarene deres ikke er pålitelige. Ved å utelukke disse fra analysen, sikrer vi høyere datakvalitet og unngår at useriøse svar påvirker resultatene.

Vi fjernet også deltakerne som ikke hadde kjennskap til ChatGPT. Disse ble fjernet for å unngå feilkilder i datasettet. Det var nødvendig å filtrere dem ut, da flere av spørsmålene spesifikt omhandlet ChatGPT, som for eksempel hvor menneskelignende eller systemlignende respondenten oppfattet verktøyet. Vi endte totalt med 972 respondenter.

5.6.1 Normalfordeling

Utvalget vårt er på 972 respondenter fordelt på fire grupper. Det er vanlig å anta normalfordeling dersom utvalget består av mer enn 30 individer (Krithikadatta, 2014). Normalfordeling er en sannsynlighetsfordeling som beskriver hvordan data fordeles symmetrisk rundt et gjennomsnitt, med størst sannsynlighet for verdier nær gjennomsnittet og gradvis lavere sannsynlighet for verdier lenger unna (Altman & Bland, 1995). De fire eksperimentgruppene vil bestå av rundt 240 respondenter hver, noe som gir grunnlag for å anta at fordelingen i alle gruppene er tilnærmet normal.

5.6.2 Statistiske tester

Enveis variansanalyse ANOVA

Den følgende analysen inkluderer en rekke statistiske tester for å kartlegge forskjeller i tillit mellom de ulike eksperimentgruppene. Vi skal benytte ANOVA, som er en enveis variansanalyse som muliggjør en sammenligning av middelveidene til to eller flere grupper samtidig (Miller, 1997). Målet med ANOVA er å kartlegge om det er en signifikant forskjell mellom gruppene, ved å se på variasjonen innenfor gruppene sammenlignet med variasjonen mellom gruppene (Tabachnick & Fidell, 2007). Testen baserer seg på F-statistikk, som estimerer forholdet mellom gruppens systematiske variasjon og tilfeldig variasjon (Henson, 2015).

Tukey HSD

For å identifisere hvilke av eksperimentgruppene som skiller seg signifikant fra hverandre, skal vi benytte en Tukey HSD-test. Tukey HSD er en post-hoc-test som utføres etter en ANOVA-analyse for å teste om det er forskjeller i middelveidier mellom parvise grupper. Testen korrigerer for multiple sammenligninger ved å bruke en justert kritisk verdi (Abdi & Williams,

2010). Testen er en konservativ test fordi alle sammenligninger mellom par av grupper vurderes ved å bruke den samme utvalgsfordelingen som også brukes for den største forskjellen mellom gruppene (Abdi & Williams, 2010). Tukey HSD-testen gir en oversikt over hvilke grupper som er signifikant forskjellige fra hverandre (Nanda, et al., 2021), og er derfor fordelaktig når vi skal sammenligne forskjeller mellom ChatGPT-gruppen, Wikipedia-gruppen, Google-gruppen og gruppen med ChatGPT og advarsel. Ulempen med Tukey HSD er at testen ikke alltid fanger små forskjeller mellom gruppene. I slik tilfeller kan det være gunstig å benytte seg av t-tester.

T-test

For å sammenligne ChatGPT spesifikt med de andre gruppene skal vi benytte t-test. En t-test er en statistisk test som sammenligner gjennomsnittet av to grupper (Kim, 2015). T-tester er vanlig å benytte når man har gjennomført et eksperiment og ønsker å sammenligne to grupper opp mot hverandre. Det er vanlig å skille mellom to typer t-tester: (1) uavhengig t-test og (2) parret t-test (XU, et al., 2017). Uavhengig t-test benyttes når gruppene er uavhengige av hverandre, i motsetning til parret hvor de er avhengige.

Et eksempel på uavhengig t-test er Welch t-test («Welch Two Sample t-test»). Welch t-test kan brukes til å sammenligne gjennomsnittet av to uavhengige grupper når variansen mellom gruppene er ulik (Rihaldijiran, 2024). Denne testen er gunstig for hypotesetesting, og anbefales å bruke når man sammenligner to og to grupper, slik som i vårt tilfelle.

Lineær regresjonsanalyse

For å analysere forholdet mellom tillit og de nevnte uavhengige variablene skal vi benytte regresjonsanalyse. Det er en av de mest anvendte statistiske metodene for å analysere flerfaktordata, og benyttes når man ønsker å forske på forholdet mellom en avhengig variabel og et sett med uavhengige variabler (Montgomery, et al., 2012). Man skiller mellom enkel lineær regresjon og multipl lineær regresjon. Enkel regresjon analyserer sammenhengen mellom en uavhengig variabel og en avhengig variabel, mens sistnevnte inkluderer flere uavhengige variabler for å forklare variasjon i den avhengige variabelen (Hanley, 2016).

Målet med en regresjonsanalyse er å (1) finne den beste regresjonslinjen som forklarer hvordan den avhengige variabelen endrer seg i forhold til de uavhengige variablene, og (2) å predikere fremtidige observasjoner eller verdier (Su, et al., 2012). Overordnet er regresjonsanalyser nyttige fordi de bidrar til å forstå sammenhenger mellom avhengig og uavhengig variabel, og kan brukes til å predikere verdien på avhengig variabel basert på verdien til den uavhengige

variabelen. Analysen er også nyttig fordi den viser hvilke variabler som har signifikant effekt på den avhengige variabelen.

Ekvivalenstest

En ekvivalenstest er en tilnærming til signifikant-testing av nullhypoteser (Lakens, et al., 2018), som brukes for å undersøke om forskjellen mellom to grupper er så liten at den kan regnes som ubetydelig i praksis. I stedet for å teste om forskjellen mellom to grupper er lik null, tester man om det er en forskjell innenfor en forhåndsdefinert ekvivalensmargin - en grense som angir hva som er akseptabelt for å regnes som ubetydelig. Dersom forskjellen er innenfor denne ekvivalensmarginen, kan gruppene anses som ekvivalente (Walker & Nowacki, 2010). Ekvivalenstester er derfor nyttige for å skille mellom statistisk signifikans og praktisk signifikans (Lakens, et al., 2018).

Signifikansnivå

Vi ønsker å avdekke om funnene våre er realistiske og ikke tilfeldige. For å oppnå dette, sammenligner vi sannsynligheten med et konfidensintervall. Det er vanlig å benytte en kritisk verdi på 5% for samfunnsvitenskapelige eksperimenter (Boos & Stefanski, 2011). Følgelig bruker vi et 95% konfidensintervall for å bekrefte om funnene våre er statistisk signifikante eller ikke. Dersom resultatene har en p-verdi som er høyere enn 5%, er ikke funnene å regne som signifikante, og vi beholder nullhypotesen. Ved å benytte et 95% konfidensintervall, er det 5% sannsynlighet for type I-feil. Type I-feil (falsk positiv) forekommer hvis man forkaster en nullhypotese som er sann (Banerjee, et al., 2009).

5.7 Forskningsetikk og personvern

Forskningsetikk handler om de etiske prinsippene og retningslinjene som skal følges for å sikre at forskning gjennomføres på en ansvarlig og rettferdig måte (Gregory, 2003). Det handler om å beskytte respondentene, sikre integritet i forskningen, og fremme tillit til resultatene.

For å sikre at alle respondentene var klar over studiens hensikt, presenterte vi dette i forkant av eksperimentet. Vi inkluderte også et samtykkeskjema hvor vi tydeliggjorde at deltakelsen var frivillig, og at de kunne trekke seg fra eksperimentet når som helst. I tillegg understrekte vi at alle svar behandles anonymt, og at vi ikke innhenter personlige opplysninger som kan spores tilbake til deltakerne.

5.8 Reliabilitet og validitet

I enhver studie er det essensielt å vurdere reliabilitet og validitet ettersom dette reflekterer hvorvidt funnene i studien er pålitelige og gyldige. I denne delen av oppgaven skal vi ta for oss begge begrepene og begrunne hvorvidt våre resultater kan brukes som grunnlag for videre forskning.

5.8.1 Reliabilitet

Reliabilitet omhandler studiens pålitelighet (Fleiss, 1999), og kan forstås som studiens interne konsistens. Det handler om hvorvidt man ville oppnådd like resultater gitt at man gjennomførte eksperimentet på et annet tidspunkt, i andre omgivelser eller med et annet utvalg. Dersom vår studie ble gjennomført av andre i en annen kontekst og fikk samme resultater, kan man si at den interne konsistensen er høy.

For å styrke reliabiliteten er det viktig å beskrive forskningsstrategien og designet grundig, og påpeke hvilke konkrete valg og antakelser som blir tatt. Høy reliabilitet indikerer at resultatene ikke er tilfeldige.

Det er en rekke trusler mot reliabilitet, blant annet ulike former for bias (Staff, 2016). For det første kan det være at deltakerne ikke har svart optimalt. Dette innebærer at vi ikke hadde mulighet til å kontrollere deltakerne under gjennomføringen av eksperimentet, og at vi derfor ikke med sikkerhet kan vite om de var opplagte og fokuserte. Det kan for eksempel tenkes at det var forstyrrelser i omgivelsene, at deltakerne var sultne eller trøtte, eller at de hadde fokuset et annet sted under deltakelsen. Dette er alle faktorer som kan ha påvirket svarene til hver enkelt respondent.

Respondentenes fysiske og psykiske tilstand under gjennomføringen av eksperimentet er en potensiell feilkilde som ligger utenfor vår kontroll (Saunders, et al., 2019). For å redusere denne risikoen begrenset vi imidlertid utvalgmengden på Profilic til å kun inkludere «troverdige» respondenter. Dette reduserer sannsynligheten for upålitelige svar. I tillegg brukte majoriteten av deltakerne 4 minutter på å fullføre eksperimentet, som også var den forventede gjennomføringstiden. Dette tyder på at respondentene har gjennomført uten vesentlig forstyrrelser.

For det andre kan *deltaker bias* ha oppstått, som refererer til at deltakerne svarer på en sosialt ønskelig måte eller slik de tror vi som forskere ønsker at de skal svare (Saunders, et al., 2019). Sannsynligheten for dette er imidlertid lav ettersom eksperimentet er anonymt. Videre kan det

være at vi som forskere har utformet eksperimentet slik at det har forekommet feil, for eksempel ved at spørsmålene har vært uklare eller misvisende (Saunders, et al., 2019). For å redusere denne risikoen har vi basert alle spørsmålene på tidligere studier og forskning, samt gjennomført en pilotstudie for å kartlegge eventuelle misforståelser ved spørsmålene.

Det kan også ha oppstått *forsker bias*. Det handler om at våre egne forventninger eller forhåndsoppfatninger påvirker hvordan forskningen utføres, analyseres eller tolkes (Saunders, et al., 2019). For å unngå dette har vi sørget for tilfeldig fordeling av deltakerne i de ulike gruppene, analysert dataene i fellesskap, samt vært åpne om hvordan vi har hentet inn og analysert dataen.

En ytterligere trussel mot reliabilitet er risikoen for selektiv rapportering, der forskere kun bruker deler av datasettet for å støtte hypotesene eller kun fremhever funn som gir signifikante utslag (Hesselberg, 2021). For å unngå dette har vi benyttet hele datasettet, avgjort på forhånd hvilke analyser som skal gjennomføres og rapportert alle relevante funn relatert til hypotesene våre.

5.8.2 Validitet

Det er mulig å skille mellom tre ulike typer validitet: begrepsvaliditet, intern validitet og ekstern validitet. I det følgende skal vi ta for oss hver av de enkelte i relasjon til vår studie.

Begrepsvaliditet

Begrepsvaliditet omhandler hvorvidt vi faktisk måler det begrepet vi ønsker å måle (Cronbach & Meehl, 1955). I vår studie er formålet å kartlegge tillit, som er et komplekst og dynamisk begrep. For å sikre en grundig dekning av tillitsbegrepet, har vi inkludert tre spørsmål for hver av de ulike tillitsdimensjonene, både for mennesketillit og systemtillit. Alle spørsmålene våre er tilknyttet de ulike tillitsdimensjonene og hentet fra tidligere forskning. Ved å benytte etablerte spørsmål som har blitt brukt og validert i tidligere studier, øker vi begrepsvaliditeten, ettersom disse spørsmålene allerede har vist seg å måle tillit på en pålitelig måte.

Samtidig er tillit som nevnt et omfattende begrep, og det er mulig at andre spørsmål kunne vært enda mer passende for å måle tillit i vår kontekst. Det finnes mange ulike måter å måle tillit på, noe som ga oss et bredt mulighetsrom da vi skulle utforme eksperimentet. Vi har derfor valgt å avgrense oss til tre spørsmål per tillitsdimensjon, noe som potensielt kan begrense begrepsvaliditeten. Likevel er spørsmålene som vi har valgt, som nevnt, anerkjente mål fra tidligere forskning og kan derfor anses som valide indikatorer på tillit.

Intern validitet

Intern validitet omhandler hvorvidt studiens design, gjennomføring og analyse er pålitelig (Andrade, 2018), samt om det kan etableres en årsakssammenheng mellom variablene.

Bruken av randomiserte grupper i kontrollerte omgivelser øker eksperimentets interne validitet (Campbell & Stanley, 1963). Derfor har vi som nevnt benyttet randomisering for å øke den interne validiteten. Randomisering reduserer påvirkningen av konfunderende variabler, som kjønn, alder og erfaring, og sikrer at forskjeller mellom gruppene ikke skyldes andre faktorer enn eksperimentets manipulasjon. I tillegg forhindrer randomiseringen at respondentenes egne bevisste eller ubevisste valg påvirker gruppetilhørigheten, og bidrar til at gruppene er sammenlignbare. Dette gjør det mulig å trekke kausale konklusjoner.

En potensiell trussel mot den interne validiteten er deltakernes forestillingsevne. Hvis deltakerne ikke klarer å visualisere situasjonen de presenteres for, kan dette svekke resultatene. Vi har imidlertid valgt en relativt konkret oppgave der deltakerne skal se for seg å innhente informasjon om klimaendringene. Dette gjør oppgaven tydelig og skaper en mer realistisk og levende situasjon.

Forstyrrelser i respondentens omgivelser kan påvirke fokus, og føre til at deltakerne svarer annerledes enn om de var i et uforstyrret miljø. For å unngå unøyaktige svar som følge av tidspress og stress, hadde respondentene ubegrenset med tid til å svare, og alle som ikke besto «attention check» ble ekskludert fra analysen. Undersøkelsen tok gjennomsnittlig 4 minutter å fullføre, og var dermed verken utmattende eller belastende for respondentene. Bruk av kompensasjon sikret også at respondentene var påkoblet og motiverte.

En svakhet ved eksperimenter kan være at respondentene ikke vet hva de svarer på, eller misforstår spørsmålene (Saunders, et al., 2019). For å forhindre dette gjennomførte vi en pilotstudie. Det var ingen som meldte om vanskelige eller merkelig formulerte spørsmål, noe som tyder på at respondentene vet hva de svarte på.

Vi kartlegger alder, IT-bakgrunn, kjønn, utdanning, syn på klimaforandringene som menneskeskapte eller ikke, og politisk tilhørighet for å avdekke om disse faktorene er med på å påvirke tillitsnivå. Ettersom demografiske og kulturelle faktorer kan påvirke tillit, er disse spørsmålene med på å styrke validiteten av resultatene.

Ekstern validitet

Ekstern validitet handler om resultatene kan generaliseres til populasjonen (Steckler & McLeroy, 2008). Jo flere respondenter, desto høyere sannsynlighet for at funnene er statistisk generaliserbare. Det er essensielt at utvalget er representativt for at funnene skal kunne generaliseres. I vår studie har vi benyttet data fra et utvalg på 972 respondenter, som er jevnt fordelt med hensyn til kjønn, alder og utdanningsnivå. Dette øker den eksterne validiteten i vår studie.

En faktor som begrenser den eksterne validiteten i studien vår, er valg av eksperiment som metode. Eksperimentet ble gjennomført via Proflic og Qualtrics, og deltakerne var dermed ikke i en naturlig setting. Dette kan ha påvirket hvordan deltakerne opptrådte og svarte, ettersom de var klar over at de deltok i et eksperiment.

I en reell situasjon, hvor deltakerne faktisk skulle hente inn informasjon om klimaforandringene og anvende denne i praksis, for eksempel til en presentasjon, kunne svarene vært annerledes. Dette skyldes at oppgaven da ville vært knyttet til en høyere opplevd risiko og konkrete konsekvenser av deres handlinger. I vårt eksperiment var oppgaven kun et imaginært scenario, der deltakerne ikke faktisk skulle hente inn informasjon. Dermed er det en risiko for at studiens resultater ikke fullt ut kan generaliseres til virkelige situasjoner.

Den eksterne validiteten i vår studie kan også være begrenset av at utvalget kun består av amerikanske respondenter. Kulturelle forskjeller eller respondentenes omgivelser kan ha påvirket svarene. For eksempel kan det være at amerikanere har en annen forståelse av KI-verktøy, enten på grunn av manglende opplæring i hvordan slike verktøy bør brukes eller fordi de allerede er godt informerte om deres funksjonalitet og begrensninger. Dette kan skille seg fra andre land, hvis kunnskapsnivå og erfaring med KI varierer. Følgelig er det usikkert hvorvidt resultatene kan generaliseres til andre kulturer og nasjoner.

6. Resultater

I denne seksjonen skal vi presentere funnene fra analysen av dataene. Vi begynner innledningsvis med å presentere deskriptiv statistikk. Deretter presenterer vi funnene.

6.1 Deskriptiv statistikk

Deskriptiv statistikk innebærer å oppsummere data på en strukturert måte ved å beskrive forholdet mellom variabler i et utvalg (Kaur, et al., 2018). Beregning av deskriptiv statistikk er et essensielt første steg i forskningsprosessen og bør alltid utføres før man går videre med statistiske analyser og sammenligninger. I vår studie har vi kartlagt alder, kjønn, utdanningsnivå, IT-bakgrunn og politisk tilhørighet.

Etter en bearbeiding av datasettet, som utdypet i metoddelen, ble det totalt inkludert 972 respondenter i analysen. Vi forventer lite skjevhet mellom gruppene, da respondentene ble randomisert og det er et relativt stort utvalg (Brysbært, 2019). Det var 243 i ChatGPT-gruppen, 244 i Wikipedia-gruppen, 244 i Google-gruppen og 241 i ChatGPT-gruppen. Grunnet det høye antallet respondenter i hver gruppe, kan vi videre anta balanserte grupper med små skjevheter på tvers av gruppene.

6.1.2 Alder og kjønn

Kjønnsfordelingen i utvalget besto av 477 kvinner, 464 menn, 16 ikke-binære og 15 individer som ikke ønsket å oppgi kjønn (vedlegg 1). Ettersom vi på forhånd hadde presisert i Profilic at vi ønsket like mange menn og kvinner i utvalget, var dette en naturlig og forventet kjønnsfordeling. I senere analyse hvor vi ser på mulige kjønnsforskjeller mellom de ulike gruppenes tillit, har vi valgt å utelukke ikke-binære og de som ikke oppga kjønn. Dette fordi vi ønsker å undersøke om eksisterende forskning på kjønnsforskjeller mellom menn og kvinner kan bekreftes eller avkreftes i vår studie. Det er også en liten andel av utvalget som identifiserer seg som ikke-binære eller ikke ønsket å svare, noe som gjør det vanskeligere å trekke meningsfulle konklusjoner om disse to gruppene dersom vi hadde inkludert de i analysen. Vår hensikt har ikke vært å ekskludere eller ignorere en minoritet, men heller forenkle analysene.

Gjennomsnittsalderen til utvalget på tvers av gruppene var på 40,85 år. For de ulike gruppene var gjennomsnittlig alder på 41,2 år, 42,1 år, 40,7 år og 39,3 år for henholdsvis ChatGPT-gruppen, Wikipedia-gruppen, Google-gruppen og gruppen med ChatGPT med advarsel. Den yngste i utvalget var 18 år, og den eldste var 80 år. Vedlegg 2 viser aldersfordelingen til utvalget.

6.1.3 Politisk tilhørighet

Utvalget var delt når det gjaldt politisk tilhørighet, med 244 republikanere, 446 demokrater, 278 som hverken identifiserte seg med noen av partiene, og 4 som ikke ønsket å oppgi politisk tilhørighet. Prosentvis utgjør dette henholdsvis 25% republikanere, 46% demokrater, 29% for ingen av partiene og 0,4% som ikke ville oppgi svar på dette spørsmålet. Dette betyr at nesten halvparten av respondentene er demokrater.

6.1.4 Utdanning og IT-bakgrunn

Respondentene ble spurt om utdanningsnivå, samt om de hadde en bakgrunn innenfor IT. Det var kun 4 respondenter som ikke hadde fullført «high school», 247 som hadde fullført «high school», 339 som var utdannet fra «college», 341 som hadde en universitetsgrad og 36 stykker som hadde en doktorgrad. 5 respondenter ønsket ikke å oppgi svar. Overordnet ser vi at de fleste i utvalget har et høyere utdanningsnivå («college» eller høyere). Det var 301 individer i utvalget som hadde IT-bakgrunn og 671 som hadde en annen bakgrunn. Følgelig er det en stor overvekt av høyt utdannede og få med IT-bakgrunn.

6.1.5 Klimafornektelse

Vi spurte respondentene om deres syn på klimaforandringene, og hvorvidt de mener at forandringene skyldes menneskelig aktivitet eller ikke. Flertallet mente klimaforandringene er menneskeskapt (857 individer) mot at de ikke er det (115 individer). Dette viser til at majoriteten av utvalget ikke er klimafornektere.

6.1.6 Tilbøyelighet til å stole på nye informasjonsverktøy

Generelt skåret utvalget vårt 4,3 på en skala fra 1 til 7, som målte tilbøyelighet til å stole på nye teknologier (vedlegg 3). Lave verdier på skalaen indikerer at respondenten i liten grad er tilbøyelig til å stole på nye teknologier, i motsetning til høye verdier. En overordnet gjennomsnittsverdi på 4,3 tyder på en generelt høy tilbøyelighet til å stole på nye verktøy. For de ulike gruppene var gjennomsnittsverdiene på henholdsvis 4,19; 4,28; 4,52 og 4,19 for ChatGPT, Wikipedia, Google og ChatGPT med en advarsel. Dette viser at det er marginale forskjeller på tvers av gruppene, og at respondentene derfor er relativt tilbøyelige til å stole på nye teknologier. Gjennomsnitt, standardavvik og median for alle gruppene er vist i vedlegg 3.

6.1.7 Brukerfaring

For å kartlegge respondentenes brukerfaring med de ulike informasjonsverktøyene, spurte vi om *hvor lenge* de har benyttet verktøyet og *hvor hyppig* de bruker det. Brukstid henviser til *hvor lenge* verktøyene er brukt, og ble målt på en fempunkts-Likert skala hvor 1 henviser til

«aldri brukt», 2 er «mindre enn 1 måned», 3 er «mellom 1 til 6 måneder», 4 er «6 måneder til 2 år» og 5 er «mer enn 2 år». Gjennomsnittlig verdi for de ulike gruppene var 3,53 for ChatGPT, 5 for Google, 3,62 for ChatGPT med advarsel og 4,89 for Wikipedia (vedlegg 4). Dette viser at respondentene har brukt Google og Wikipedia lengre enn ChatGPT, noe som er naturlig da ChatGPT ble lansert for 2 år siden. ChatGPT er derfor det nyeste verktøyet av de tre, og det kan tenkes respondentene ikke benyttet ChatGPT med én gang etter lansering. Det er marginale forskjeller mellom Google og Wikipedia, og vi ser at medianen er på 5 for begge gruppene.

Brukshyppighet henviser til *hvor ofte* verktøyene brukes, og ble målt på en fempunkts-Likert skala hvor 1 henviser til «daglig», 2 til «ukentlig», 3 til «månedlig», 4 til «årlig», 5 til «sjeldnere» og 6 til «aldri». Lave verdier gjenspeiler derfor hyppig bruk, i motsetning til høye verdier som reflekterer sjeldnere bruk. Google benyttes hyppigst, med en gjennomsnittsverdi på 1,13 og Wikipedia brukes mest sjeldent (2,68) (vedlegg 5). For ChatGPT-gruppene er gjennomsnittsverdiene på henholdsvis 2,45 og 2,53 uten og med advarsel.

Medianen for Google er på 1, og for de andre gruppene er medianen på 2. Dette viser til at Google i høy grad benyttes daglig, og Wikipedia nærmest månedlig. Grunnen til at Google skårer høyere her kan være fordi Google er et mer generelt informasjonsverktøy i form av at mange bruker Google for å finne andre kilder, som for eksempel Wikipedia. En grunn til at Wikipedia skårer så lavt kan være at individer hovedsakelig bruker Wikipedia til akademiske formål (Selwyn & Gorard, 2016), og ikke hverdagslige slik som de kanskje benytter Google til. Det er for eksempel mer naturlig å spørre Google om dagligdagse emner slik som en oppskrift, en quiz, værmeldingen etc. På den andre siden er det trolig mer vanlig å benytte Wikipedia for å få innsikt i et tema, slik som klimaforandringene eller en kjent historiker. ChatGPT kan muligens forstås som et verktøy som er mellom disse gruppene, da det både kan benyttes for dagligdagse tema, men også akademiske.

6.1.8 ChatGPT 3.5 og ChatGPT 4

Innledningsvis spurte vi deltakerne om de benyttet gratisversjonen av ChatGPT eller den betalte. 93% av respondentene benyttet gratisversjonen i motsetning til 7% som har versjonen man selv må betale for. Dette gjenspeiler at det er få i utvalget som har tilgang til, eller benytter seg av, den nyeste versjonen av chatboten. Grunnet den store skjevfordelingen i antall som benytter gratisversjonen sammenlignet med versjonen bak betalingsmur, kommer vi ikke til å gjennomføre statistiske analyser for å undersøke om dette påvirker tillitsnivå. Dette fordi

testene ikke vil ha statistisk styrke, da det er en såpass skjev fordeling av antallet som benytter de to versjonene.

6.2 Diskusjon av hypotesesettene

I det følgende skal vi ta for oss de ulike hypotesesettene utformet tidligere i oppgaven, og drøfte funnene opp mot relevant teori. Gjennomgående vil vi foreta en helhetlig diskusjon av tilliten til ChatGPT i forhold til Google og Wikipedia.

6.2.1 Hypotesesett 1: Systemtillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia

I hypotesesett 1 skal vi presentere våre analyser av tillit til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med advarsel for systemtillitsdimensjonene.

Tabell 1 viser det gjennomsnittlige tillitsnivået for systemtillitsdimensjonene samlet for de ulike gruppene. Her ser vi at Google skårer høyest på tillitsnivå, og Wikipedia skårer lavest. Det er generelt marginale forskjeller i tillit mellom de ulike gruppene, men vi observerer at ChatGPT uten advarsel i gjennomsnitt skårer høyere på systemtillit enn ChatGPT med advarsel. Dette gir indikasjoner på at en advarsel om at verktøyet tar feil kan være med på å svekke tilliten brukerne har til verktøyet.

Gruppe	Gjennomsnittlig systemtillit	Minimum	Maksimum	Standardavvik
ChatGPT	5,22	1	7	1,15
Google	5,5	1	7	1,1
Wikipedia	5,11	1	7	0,96
Advarsel	5,05	1	7	1,15

Tabell 1: Oversikt over gjennomsnittlig systemtillit for de ulike gruppene. «Advarsel» tilsvarer gruppen som ble eksponert for ChatGPT med en advarsel.

Under vises en oversikt over resultatene for t-testene for de ulike systemtillitsdimensjonene mellom ChatGPT og de tre andre gruppene: Google, Wikipedia og ChatGPT med advarsel («Advarsel»).

Funksjonalitet	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	5,609	-0,368	-3,791	475,840	0,000*
	Google	5,977				
ChatGPT vs Wikipedia	ChatGPT	5,609	0,135	1,325	484,500	0,186
	Wikipedia	5,474				
ChatGPT vs Advarsel	ChatGPT	5,609	0,135	1,404	470,240	0,161
	Advarsel	5,474				

Tabell 2: Resultater t-test for funksjonalitet.

Hjelpsomhet	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	5,320	0,004	0,036	483,450	0,971
	Google	5,316				
ChatGPT vs Wikipedia	ChatGPT	5,320	0,314	2,869	484,960	0,004*
	Wikipedia	5,006				
ChatGPT vs Advarsel	ChatGPT	5,320	0,068	0,679	464,580	0,497
	Advarsel	5,252				

Tabell 3: Resultater t-test for hjelpsomhet.

Pålitelighet	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	4,733	-0,467	-3,089	484,980	0,000*
	Google	5,200				
ChatGPT vs Wikipedia	ChatGPT	4,733	0,072	0,555	480,920	0,579
	Wikipedia	4,661				
ChatGPT vs Advarsel	ChatGPT	4,733	0,140	1,196	477,100	0,232
	Advarsel	4,593				

Tabell 4: Resultater t-test for pålitelighet.

Vi gjennomførte en ANOVA-analyse for å undersøke om forskjellene i systemtillit mellom alle gruppene er signifikante (vedlegg 6). Resultatene viser at det er signifikante forskjeller mellom gruppene, med p-verdi på tilnærmet lik null.

For å kartlegge hvilke grupper som skiller seg signifikant fra hverandre, kjørte vi en Tukey HSD-test (vedlegg 7). Resultatene fra Tukey HSD-testen viser at det er signifikante forskjeller mellom (1) ChatGPT-gruppen og Google-gruppen, (2) Wikipedia og Google, og (3) ChatGPT med advarsel og Google. Overordnet viser dette at respondentene har ulik systemtillit til informasjonsverktøyene. For vårt forskningsspørsmål er det kun relevant å se nærmere på den signifikante forskjellen mellom ChatGPT og Google. Forskjeller mellom de andre gruppene er ikke relevant i denne studien.

Systemtillit til ChatGPT sammenlignet med Google

Tilliten til Google er høyere enn ChatGPT for systemtillit, og forskjellen mellom gruppene er 0,2773 (p-verdi = 0,0267). Dette samsvarer med en tidligere kvalitativ studie som viste at deltakerne stolte mer på Google enn ChatGPT (Jung, et al., 2024). Våre funn bekrefter dette da vi finner signifikant forskjell i tillitsnivå mellom de to gruppene. Selv om vi har funnet at brukerne har høyere tillit til Google enn ChatGPT, ønsker vi også å undersøke om det er forskjeller i tillit for hver av de ulike systemtillitsdimensjonene.

Funksjonalitet

Vi gjennomførte en Welch t-test for å måle forskjellene i oppfattet funksjonalitet mellom ChatGPT-gruppen og Google-gruppen (tabell 2).

Resultatet viser en signifikant forskjell i gjennomsnitt mellom de to gruppene (p-verdi < 0,05), hvor Google skårer 0,37 høyere enn ChatGPT på funksjonalitet. Dette er i tråd med Jung et al. (2024) som fant at brukerne har høyere tillit til Google enn for ChatGPT. En høyere grad av oppfattet funksjonalitet for Google enn ChatGPT er trolig årsaken til den høyere overordnede tilliten til Google. Til tross for at ChatGPT skårer lavere enn Google på opplevd funksjonalitet, ser vi likevel av gjennomsnittene på henholdsvis 5,609 og 5,97 (på en skala fra 1 til 7), at brukerne opplever at begge har en høy funksjonalitet.

Hjelpsomhet

Vi gjennomførte også en Welch t-test for å analysere om det er signifikante forskjeller i oppfattet hjelpsomhet mellom ChatGPT-gruppen og Google-gruppen (tabell 3).

Her viste analysen at det ikke var noen signifikant forskjell mellom gruppene. Gjennomsnittsverdiene for hjelpsomhet for Google og ChatGPT var tilnærmet like med verdier på henholdsvis 5,315 og 5,319. Dette indikerer at det er en svært lik vurdering av informasjonsverktøyenes hjelpsomhet. En grunn til dette kan være at begge verktøyene dekker brukernes behov, men på ulike måter.

Både Google og ChatGPT er hjelpsomme ved at de gir rask tilgang til informasjon. Google gir raske søkeresultater til et bredt spekter av relevante kilder, noe som er hjelpsomt for brukeren hvis de skal utforske et tema og innhente informasjon, slik som for eksempel i vårt eksperiment. ChatGPT, på sin side, gir brukeren direkte svar i samtaleform, og hjelper brukeren med å spare tid da den oppsummerer informasjon, og kan enkelt forklare komplekse konsepter. Overordnet gir begge verktøyene dermed brukeren hjelp til å finne og forstå informasjon, hvor Google lar brukeren selv navigere mellom kilder, mens ChatGPT gir informasjon direkte.

På bakgrunn av at vi fikk et tilnærmet likt gjennomsnitt for ChatGPT og Google på oppfattet hjelpsomhet, ønsket vi å sjekke om denne likheten er signifikant. Vi gjennomførte derfor en ekvivalenstest (TOST) for å vurdere om forskjellen i hjelpsomhet mellom ChatGPT- og Google-gruppen kunne anses som praktisk ubetydelig (vedlegg 8). Vi definerte derfor ekvivalensmarginen til $\pm 0,249$ (Lakens, 2017; Meyners, 2012).

Resultatene fra ekvivalenstesten indikerer at forskjellen mellom gruppene med hensyn til oppfattet hjelpsomhet, er minimal og innenfor den definerte ekvivalensgrensen (-0.249 ; 0.249). Den observerte middelveidiforskjellen (0.004) er svært liten, og TOST-testen bekreftet at 90 % konfidensintervallet ($-0.182, 0.19$) ligger innenfor de definerte grensene for praktisk ekvivalens. Dette tyder på at gruppene kan betraktes som ekvivalente. Videre viste NHST-resultatene at 95 % konfidensintervallet ($-0.218, 0.226$) inkluderte null, noe som betyr at det ikke er noen signifikant forskjell mellom gruppene. Samlet sett støtter analysen at ChatGPT-gruppen og Google-gruppen kan anses som likeverdige med hensyn til hvor hjelpsomme de oppfattes for å være av brukerne.

Dette funnet er interessant og indikerer at selv om brukerne har høyere overordnet systemtillit til Google sammenlignet med ChatGPT, så ser det likevel ut til at de oppfattes som like hjelpsomme.

Pålitelighet

I likhet med tillitsdimensjonene funksjonalitet og hjelpsomhet, gjennomførte vi en t-test for å avdekke om det var signifikante forskjeller mellom Google og ChatGPT i hvor pålitelig brukerne oppfatter verktøyene (tabell 4).

Funnene fra denne analysen viser at gjennomsnittlig vurdering av påliteligheten var høyere for Google (5.20) sammenlignet med ChatGPT (4.73). Det var en signifikant forskjell i gjennomsnittsverdi mellom de to informasjonsverktøyene, hvor Google skårer 0,47 høyere. Resultatene er signifikante som bekreftes av at p-verdien er tilnærmet lik null. Dette betyr at brukerne oppfatter Google som mer pålitelig enn ChatGPT.

Oppsummerende finner vi i vår analyse at tillit til Google er signifikant høyere enn for ChatGPT for systemtillitsdimensjonene samlet. Dette skyldes hovedsakelig at Google oppfattes som å være mer funksjonell og pålitelig enn ChatGPT. Begge oppfattes som like hjelpsomme. Av den grunn kan vi beholde H_{1-1a} og H_{1-1c} , men ikke H_{1-1b} .

Systemtillit til ChatGPT sammenlignet med Wikipedia

I likhet med Google, gjennomførte vi t-test for å kartlegge om det var forskjeller i systemtillit mellom ChatGPT og Wikipedia. T-testen viste forskjeller i samlet systemtillit for de to informasjonsverktøyene, med 5,22 i gjennomsnittlig tillitsnivå for ChatGPT og 5,05 for Wikipedia. Dette indikerer at brukerne har høyere systemtillit til ChatGPT sammenlignet med Wikipedia. Imidlertid var ikke forskjellene statistisk signifikant på 5%-nivå, da p-verdien var 0,0968 (vedlegg 1). P-verdien er på tilnærmet 0,01, noe som betyr at dersom vi hadde valgt et 90% konfidensintervall, ville forskjellene vært signifikante.

Funksjonalitet, hjelpsomhet og pålitelighet

På bakgrunn av at vi på et 90% konfidensintervall ville fått signifikante forskjeller i systemtillit til ChatGPT sammenlignet med Wikipedia (vedlegg 9), ønsket vi å undersøke om det er signifikante forskjeller mellom gruppene for de tre systemtillitsdimensjonene.

Vi gjennomførte derfor en Welch t-test, for hver av de tre dimensjonene for ChatGPT og Wikipedia. T-testene viste at det ikke var signifikante forskjeller i funksjonalitet og pålitelighet mellom ChatGPT og Wikipedia. Derimot viste t-testen at det var signifikante forskjeller i hvor hjelpsomme brukerne oppfattet de to informasjonsverktøyene (tabell 3).

Fra analysen ser vi at ChatGPT blir oppfattet som mer hjelpsom enn Wikipedia, med en forskjell på omtrent 0,31 i gjennomsnittsverdi (p-verdi= 0,0043). En grunn til at ChatGPT kan

bli oppfattet som mer hjelpsom er fordi verktøyet gir brukeren skreddersydde svar basert på brukerens prompt, det vil si stilte spørsmål eller instruks på plattformen. I tillegg kan brukeren få justere svar basert på kontekst og egne preferanser eller nivå av forståelse. I informasjonssøk om klimaforandringer kan dette være gunstig om brukeren ønsker et raskt og spesifikt svar, noe som kan forklare hvorfor ChatGPT skårer høyere på hjelpsomhet.

Wikipedia på sin side tilbyr statiske artikler som ikke kan tilpasse seg individuelle behov. Brukeren må selv selektere informasjonen som er relevant, i artikler som ofte er lange. I tillegg kan ChatGPT bli ansett som mer hjelpsom fordi verktøyet lar brukeren stille oppfølgingsspørsmål om noe er uklart. ChatGPT kan også komme med presiseringer eller eksempler, noe som gjør det enklere for brukeren å forstå et tema.

Videre kan ChatGPT også tilpasse språket og formulere informasjonsmaterialet for at brukeren lettere skal forstå. På denne måten har ChatGPT høyere grad av personalisering enn Wikipedia, noe som kan øke brukerens tillit (Glikson & Woolley, 2020), og gjøre at ChatGPT oppfattes som mer hjelpsom.

Wikipedia kan ofte ha tyngre og mer akademisk språk med tekniske forklaringer, som ikke alltid er like brukervennlig. Wikipedia har heller ikke informasjon om alle mulige tema, i motsetning til ChatGPT, som kan svare på alt brukeren lurer på.

En annen grunn til at ChatGPT kan bli oppfattet som mer hjelpsomt enn Wikipedia kan være fordi KI-chatboten leverer informasjon raskt, og det krever mindre innsats fra brukeren å finne og forstå det de trenger. På Wikipedia må brukeren selv sortere og vurdere informasjon fra ulike deler av en eller flere artikler, noe som krever mer tid og innsats.

Oppsummerende viser våre analyser at det ikke er signifikante funn som gir støtte til hypotesen om at ChatGPT skårer lavere på systemtillit enn Wikipedia. Analysene gir heller ikke signifikante utslag for systemtillitsdimensjonene funksjonalitet og pålitelighet. Likevel er det signifikante forskjeller i hjelpsomhet, hvor brukerne oppfatter ChatGPT som mer hjelpsom enn Wikipedia. Av den grunn beholder vi $H_{1-2null}$.

6.2.2 Hypotesesett 2: Mennesketillitsdimensjonene for ChatGPT sammenlignet med Google og Wikipedia

I hypotesesett 2 skal vi presentere analyser av tillit til ChatGPT sammenlignet med Google og Wikipedia for mennesketillitsdimensjonene.

Tabell 5 viser gjennomsnittlig tillitsnivå for mennesketillitsdimensjonene samlet for de ulike gruppene. Her ser vi at Google skårer høyest på tillitsnivå (4,87) og Wikipedia skårer lavest (4,75). Det er generelt svært marginale forskjeller i tillit mellom de ulike gruppene, noe som indikerer at mennesketillit er relativt likt på tvers av gruppene.

Gruppe	Gjennomsnittlig mennesketillit	Minimum	Maksimum	Standardavvik
ChatGPT	4,85	1	7	1,2
Google	4,87	1	7	1,26
Wikipedia	4,83	1	7	1,09
Advarsel	4,75	1	7	1,33

Tabell 5: Oversikt over gjennomsnittlig mennesketillit for de ulike gruppene. «Advarsel» tilsvare gruppen som ble eksponert for ChatGPT med en advarsel.

Under vises en oversikt over resultatene for t-testene for de ulike mennesketillitsdimensjonene mellom ChatGPT og de tre andre gruppene:

Evne	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	5,219	-0,188	-1,606	484,700	0,109
	Google	5,407				
ChatGPT vs Wikipedia	ChatGPT	5,219	0,181	1,470	481,950	0,142
	Wikipedia	5,038				
ChatGPT vs Advarsel	ChatGPT	5,219	0,169	1,512	476,880	0,131
	Advarsel	5,050				

Tabell 6: Resultater t-test for evne.

Velvilje	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	4,357	0,097	0,768	478,220	0,443
	Google	4,260				
ChatGPT vs Wikipedia	ChatGPT	4,357	-0,061	-0,506	484,130	0,613
	Wikipedia	4,418				
ChatGPT vs Advarsel	ChatGPT	4,357	-0,084	-0,742	478,980	0,458
	Advarsel	4,441				

Tabell 7: Resultater t-test for velvilje.

Integritet	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Google	ChatGPT	4,909	-0,083	-0,706	484,080	0,480
	Google	4,992				
ChatGPT vs Wikipedia	ChatGPT	4,909	0,055	0,458	478,270	0,647
	Wikipedia	4,854				
ChatGPT vs Advarsel	ChatGPT	4,909	-0,025	-0,227	472,310	0,820
	Advarsel	4,934				

Tabell 8: Resultater t-test for integritet.

Google-gruppen og Wikipedia-gruppen har de høyeste standardavvikene, noe som betyr at det er større variasjon i svarene til disse to gruppene sammenlignet med de to andre. Dette tyder på at det er mer homogene svar for de to ulike ChatGPT-gruppene i spørsmålene tilknyttet mennesketillit.

Forskjellen i gjennomsnittlig tillitsnivå for menneskedimensjonene er betydelig mindre enn for systemtillitsdimensjonene, noe som indikerer at det er større sprik i svarene på sistnevnte. Dette støtter Lankton et al. (2015) sitt syn på at man bør benytte systemtillitsdimensjoner når man måler tillit til systemer, spesielt ettersom både Google og ChatGPT blir sett på som mer systemlignende i vårt eksperiment. Dette kommer vi tilbake til i hypotesesett 6.

I lys av at vi fikk små forskjeller i gjennomsnittlig tillitsnivå for mennesketillitsdimensjonene, var det ikke overraskende at ANOVA-analysen ikke viste signifikante forskjeller mellom de fire gruppene. P-verdien var på 0,741, og indikerer at variasjonen mellom gruppene trolig skyldes tilfeldigheter. Dette funnet ble bekreftet av en Tukey HSD-test, som viser at det ikke er signifikante forskjeller når man sammenligner to og to av gruppene. En implikasjon av dette er at respondentene vurderer mennesketillit konsistent på tvers av verktøyene.

Oppfattet evne, velvilje og integritet for ChatGPT sammenlignet med Google og Wikipedia

Til tross for ingen signifikante forskjeller mellom gruppene på samlet mennesketillit gjennomførte vi en ANOVA-analyse og Tukey HSD-test for å undersøke om det er forskjeller mellom gruppene for de ulike tillitsdimensjonene separat.

I ANOVA-analysen over forskjeller i evne mellom de ulike verktøyene, viste analysen at det er signifikante forskjeller mellom to eller flere av gruppene. En Tukey HSD-test bekreftet at dette i all hovedsak skyldtes forskjeller mellom (1) ChatGPT-gruppen med advarsel og Google-

gruppen, og (2) Wikipedia-gruppen og Google-gruppen. Da vår studie ser på forskjeller mellom ChatGPT og de andre gruppene, velger vi å ikke gå nærmere inn på disse funnene.

Det var ingen andre signifikante forskjeller da vi sammenlignet de andre gruppene, noe som indikerer at oppfattet evne for de ulike verktøyene er relativt likt. Dette samsvarer også med de relativt like gjennomsnittsskårene vi fikk for mennesketillit samlet.

For velvilje og integritet viste ANOVA-analysen at det ikke er signifikante forskjeller for de ulike verktøyene. Tukey HSD-testen bekrefter dette funnet, noe som betyr at respondentene vurderer velvilje og integritet konsistent uavhengig av informasjonsverktøy.

Oppsummert, har vi ingen signifikante utslag for mennesketillit eller de tilhørende tillitsdimensjonene. Dermed beholder vi nullhypotesene $H_{2-1null}$ og $H_{2-2null}$.

6.2.3 Hypotesesett 3: ChatGPT sammenlignet med ChatGPT med advarsel

I hypotesesett 3 skal vi gjennomføre analyser for å teste om tillit til ChatGPT påvirkes av en advarsel om at den kan ta feil. Vi skal først ta for oss påvirkningen av advarsel for systemtillit for deretter å se på mennesketillit.

Tabell 1 viste at ChatGPT uten advarsel hadde et gjennomsnittlig tillitsnivå på 5,22 for systemtillitsdimensjonene samlet, mens det for ChatGPT med advarsel var 5,11. Standardavviket i ChatGPT-gruppen med advarsel var på 0,96, og for ChatGPT-gruppen uten advarsel var standardavviket på 1,15. Dette viser at det var mindre variasjon i hvor tillitsverdige respondentene oppfattet ChatGPT med en advarsel, til tross for at differansen i standardavvik er lav (0,19). I lys av at vi måler tilliten på en skala på 1 til 7, er derfor ikke denne forskjellen særlig betydelig i praktisk forstand. Ved gjennomføring av Tukey HSD-test fant vi at ChatGPT-gruppen med advarsel har 0,11 lavere systemtillit enn for ChatGPT-gruppen uten advarsel. Imidlertid var denne forskjellen langt fra signifikant (p-verdi = 0,66).

Videre gjennomførte vi t-tester for systemtillit samlet, og de tilhørende tillitsdimensjonene. Resultatene fra t-testene mellom ChatGPT og ChatGPT med en advarsel ga null utslag, noe som indikerer at det ikke er forskjeller i systemtillit mellom gruppene. Videre viste funnene at det ikke er ulikheter mellom hvor funksjonell, pålitelig og hjelpsom ChatGPT oppfattes som både med og uten advarsel. Vi gjennomførte tilsvarende tester for mennesketillit, hvor det heller ikke her var signifikante utslag.

Funnene gir altså grunnlag for å argumentere for at en advarsel om at ChatGPT kan ta feil, ikke har påvirkning på tillit til verktøyet, både for systemtillitsdimensjonene og mennesketillitsdimensjonene. Dermed må vi beholde nullhypotesene $H_{3-1null}$ og $H_{3-2null}$.

Disse funnene var overraskende. ChatGPT fremhever selv at den kan ta feil på sin plattform, noe vi antok brukerne ville tolke som en transparent handling (Bedué & Fritzsche, 2021; Glikson & Woolley, 2020). Det kan være flere grunner til at vi ikke fikk dette utslaget.

Først og fremst kan det skyldes at advarselen ikke var tydelig nok. Det kan tenkes at utfallet hadde blitt annerledes dersom vi heller utformet advarselen slik som for cookies på nettsider. Dersom brukeren aktivt måtte krysse av for at den har fått med seg informasjonen, kunne dette gitt et annet utfall.

For det andre kan det tenkes at brukerne har en «baseline-tillit» til ChatGPT som ikke endres ved en synliggjøring av at verktøyet kan ta feil. «Baseline-tillit» er det forhåndsbestemte nivået av tillit et individ har i møte med en teknologi, og er ikke basert på tidligere erfaringer eller kjennskap til teknologien (Wang & Redmiles, 2016).

Selv med en advarsel, kan denne «baseline-tilliten» opprettholdes hvis brukeren allerede har dannet seg en oppfatning av systemet som kompetent, og har tillit til det. I tillegg kan det tenkes at brukerne er vant med at ChatGPT kan ta feil, da teknologier ikke alltid er perfekte (Ryseff, et al., 2024). En advarsel om feil påvirker derfor ikke nødvendigvis hvordan individer bruker eller vurderer ChatGPT, hvis de allerede har en kjennskap til og en forventning om at verktøyet kan ta feil.

Det kan også være at advarselen, som argumentert for tidligere, øker tilliten til noen brukere fordi systemet oppfattes som mer transparent og troverdig (Zerilli, et al., 2022). For enkelte brukere kan effekten være motsatt, men gitt at begge effektene oppstår, kan dette føre til at gjennomsnittet likevel er tilnærmet likt som for ChatGPT-gruppen, fordi de to effektene nøytraliserer hverandre.

6.2.4 Hypotesesett 4: Brukserfaring sin påvirkning på tillit

I hypotesesett 4 skal vi analysere hvorvidt tilliten til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med advarsel påvirkes av erfaringen brukerne har med informasjonsverktøyet. For å undersøke dette, ser vi på de to uavhengige variablene *brukstid* og *brukshyppighet*, som begge gir en indikasjon på brukernes brukererfaring med verktøyet.

Brukstid og systemtillit

Først skal vi analysere om brukstid påvirker systemtillit til de ulike informasjonsverktøyene, og hvorvidt denne effekten varierer mellom gruppene.

Dette undersøkte vi ved å kjøre en lineær regresjonsanalyse (vedlegg 10). Resultatet fra analysen viser at brukstid («use») påvirker systemtillit positivt. Den positive effekten betyr at jo lengre en respondent har brukt informasjonsverktøyet, desto høyere er systemtillit (p -verdi $< 0,001$). For ChatGPT (referansegruppen) øker systemtilliten med 0,311 for hver økning i brukstid. Videre viser modellen at interaksjonene mellom brukstid og de andre gruppene (Google, Wikipedia, ChatGPT med advarsel) ikke er signifikante.

Grunnen til at tilliten øker med tidsperioden et individ har brukt verktøyet, kan skyldes flere grunner. For det første kan det tenkes at jo mer et individ bruker et verktøy, jo mer kjent blir individet med dette, noe som kan redusere usikkerhet rundt bruk, og øke tillit. En fortrolighet med et verktøy kan skape en følelse av kontroll for brukeren, noe som kan øke tilliten til verktøyet. Med økt brukslengde kan også brukeren få flere positive opplevelser med verktøyet, som kan styrke tilliten over tid. Det kan også føre til at de positive opplevelsene dominerer potensielle negative opplevelser dersom verktøyet tar feil.

Videre kan det være at langvarig bruk av et system fører til at individet utvikler vaner, og automatisk stoler på verktøyet fordi den kritiske vurderingen faller. Det kan også være at brukeren går i bekräftelsesfellen (Oswald & Grosjean, 2005), ved at brukeren over tid kun velger å fokusere på situasjoner der verktøyet har vært nyttig eller riktig, og ignorerer feil eller svakheter. Dette kan føre til en selvforsterkende syklus av tillit.

En annen grunn til at det er en positiv sammenheng mellom disse variablene kan være at mengde tid et individ investerer i et verktøy, kan øke sannsynligheten for at de føler seg praktisk eller emosjonelt tilknyttet verktøyet. Dette er en form for «tilknytnings bias», som handler om at individet rettferdiggjør egen bruk ved å stole mer på verktøyet (Back, 2010). Gitt at brukeren ikke oppdager feil med informasjonsverktøyet, kan dette også forsterke tilliten.

Overordnet har regresjonsmodellen lav justert forklaringskraft, $R^2 = 3,65\%$. Dette betyr at det er andre variabler som i større grad er med på å forklare variansen i systemtillit for de andre gruppene.

Selv om brukstid ikke forklarer forskjeller i tillit mellom alle gruppene, ønsket vi å utforske om brukstid positivt påvirker systemtilliten for hver av gruppene individuelt. Dette undersøkte

vi ved å gjennomføre en lineære regresjonsmodeller for hver av gruppene og deretter sammenligne resultatene. Grunnen til at vi ønsket å gjøre dette, var for å se i hvor stor grad brukstid påvirker de ulike gruppene, og om effekten av denne variabelen var tilnærmet lik på tvers av gruppene.

Resultatet av regresjonene viste at brukstid ikke hadde signifikant påvirkning på systemtillit for Google og Wikipedia. Derimot var det signifikante utslag for ChatGPT-gruppene både med og uten advarsel med p-verdier tilnærmet lik 0,001 for begge.

Resultatene viser at respondenter som har benyttet ChatGPT i lengre tid, har høyere systemtillit (vedlegg 11). For hver økning i hvor lenge respondenten har brukt verktøyet, øker systemtillit med 0,31 (p-verdi= 0,0011). Den justerte forklaringskraften til modellen er 3,93%, som viser at modellen forklarer variasjonen i systemtillit med tilnærmet 4%. Overordnet betyr dette at brukstid positivt påvirker systemtillit til ChatGPT, noe som indikerer at brukererfaring med verktøyet kan øke tilliten til det. Dermed beholder vi $H_{4-1null}$.

Oppsummerende ser vi at lengre bruk av verktøyet vil ha en positiv effekt på de systembaserte tillitsdimensjonene for ChatGPT, både med og uten advarsel.

Brukstid og mennesketillit

Vi utførte tilsvarende analyser av påvirkning av brukstid for mennesketillit (vedlegg 14). Tilsvarende som for systemtillit var det ikke signifikante forskjeller i brukstid sin påvirkning på mennesketillit mellom de ulike gruppene. Likevel viste resultatet fra regresjonsmodellen at påvirkningen av brukstid på mennesketillit er tilnærmet signifikant (p-verdi =0.0643) (vedlegg 2). Da dette funnet ikke er langt fra signifikant, velger vi å se på hva resultatene impliserer.

Koeffisienten viser at for hver økning i hvor lenge brukeren har brukt verktøyet, øker mennesketilliten med 0,18. Regresjonsmodellens justerte forklaringskraft var på 0,005, som tyder på at modellen ikke forklarer mye av variasjonen i mennesketillit.

For å undersøke om bruk av informasjonsverktøyet hadde effekt på mennesketillit i hver av de enkelte gruppene, gjennomførte vi også her lineære regresjonsanalyser for alle gruppene. Funnene viste at det ikke var signifikante effekter i noen av gruppene foruten gruppen som ble eksponert for ChatGPT med advarsel. Her viste regresjonsanalysen at brukstid var positivt assosiert med mennesketillit, hvor hver økning i brukstid økte tillit med 0,25 (p-verdi < 0,01).

Oppsummerende ser vi at lengre bruk av verktøyet vil ha en positiv effekt på de menneskebaserte tillitsdimensjonene for ChatGPT med advarsel, men ingen signifikant effekt

for ChatGPT uten advarsel, Google og Wikipedia. Dermed beholder vi nullhullhypotesen, H_4 .
3null.

Brukshyppighet og systemtillit

I det følgende skal vi analysere hvorvidt systemtillit til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med advarsel påvirkes av brukshyppighet.

Som sist, benyttet vi lineær regresjonsanalyse (vedlegg 12). Skalaen her var motsatt fra brukstid, noe som resulterer i at tolkningen blir motsatt. Resultatet fra analysen viser at den uavhengige variabelen brukshyppighet («frequency») påvirker systemtillit negativt. Dette betyr at jo sjeldnere en respondent bruker informasjonsverktøyet, desto lavere er systemtilliten. Effekten er signifikant med en p-verdi tilnærmet lik 0,001. For hver reduksjon i variabelen, faller systemtilliten med 0,34.

Dette funnet indikerer at brukshyppighet har en overordnet signifikant effekt på systemtillit. Dette antyder at sjeldnere bruk av informasjonsverktøyene er assosiert med lavere tillit til verktøyet, og motsatt at økt bruk er assosiert med høyere tillit. Interaksjonene mellom brukshyppighet og de ulike gruppene er ikke signifikante.

Modellen har en justert forklaringskraft på 0,142, som tilsvarer 14,2%, og fanger opp en del av variasjonen i systemtillit. Sammenlignet med brukstid, har modellen for brukshyppighet en større forklaringskraft. Dette indikerer at hvor hyppig et individ benytter verktøyet har større påvirkning på tillit sammenlignet med hvor lenge individet har benyttet verktøyet. Dette kan være fordi økt eksponering overfor verktøyet innebærer flere muligheter for både positive og negative opplevelser.

Vi kjørte regresjonsmodeller for hver av gruppene. Vi finner at brukshyppighet ikke har signifikant påvirkning på systemtillit for Google. Derimot er det sterke signifikante utslag for de tre andre gruppene med p-verdier tilnærmet lik null. Wikipedia har den største reduksjonen i tillit med sjeldnere bruk (-0,41). ChatGPT med advarsel har en reduksjon på 0,297 og for ChatGPT uten advarsel er reduksjonen på 0,34.

Resultatene viser at respondenter som sjeldnere benytter ChatGPT, har lavere systemtillit (vedlegg 13). Modellens justerte forklaringskraft er på 13,06%. For hver enhetsreduksjon i brukshyppighet, er det en reduksjon i systemtillit med 0,34. Dette indikerer at jo sjeldnere en respondent har brukt verktøyet, desto lavere tillit.

Overordnet viser funnet at jo hyppigere et individ bruker verktøyet, jo høyere systemtillit. Dette gjelder for begge ChatGPT-gruppene og Wikipedia. Dermed beholder vi H_{4-2c} , men forkaster H_{4-2a} og H_{4-2b} .

Brukshyppighet og mennesketillit

Vi utførte tilsvarende multiple lineære regresjonsanalyser for mennesketillit. I likhet med funnene for systemtillit finner vi ikke signifikante forskjeller mellom de ulike gruppene.

Da vi gjennomførte lineære regresjoner for hver av gruppene, for å se om brukshyppighet påvirker mennesketilliten, fikk vi til forskjell fra den tidligere analysen av mennesketillit og brukstid, signifikante utslag for alle gruppene, inkludert Google. For Google var p-verdien på 0,015, mens for de tre andre gruppene var p-verdien tilnærmet lik null. Analysen indikerer dermed at respondentene som bruker ChatGPT, Google og Wikipedia sjeldnere, også har lavere mennesketillit til verktøyene. Dette betyr at jo oftere en respondent har brukt verktøyet, desto høyere tillit har den. Dermed forkaster vi H_{4-4a} , men beholder H_{4-4b} og H_{4-4c} .

6.2.5 Hypotesesett 5: Generell tilbøyelighet til å stole på nye informasjonsverktøy sin påvirkning på tillit

I hypotesesett 5 analyserer vi hvordan generell tilbøyelighet til å stole på nye informasjonsverktøy påvirker tillit til informasjonsverktøyene.

Generell tilbøyelighet til å stole på nye informasjonsverktøy og systemtillit

Vi gjennomførte en lineær regresjonsanalyse av hvordan generell tilbøyelighet til å stole på nye informasjonsverktøy og gruppetilhørighet, påvirker systemtillit for de ulike verktøyene (vedlegg 15).

Funnene viser at hver økning i tilbøyelighet til å stole på nye informasjonsverktøy («dispo_combined»), øker tilliten til ChatGPT med 0,42. Systemtilliten øker betydelig mindre for Wikipedia med 0,20 per økning i tilbøyelighet til tillit når vi hensyntar interaksjonseffekten. Resultatene viser interaksjonseffekten er sterkt signifikant negativ (p-verdi < 0,01). Dette betyr tilbøyelighet til å stole på nye informasjonsverktøy sin effekt på systemtillit er svakere for Wikipedia-brukere enn for ChatGPT-brukere.

En grunn til effekten er sterkere for ChatGPT-gruppen kan være at ChatGPT er et nyere verktøy enn Wikipedia, noe som gjør at respondentene i større grad må basere tilliten på en generell disposisjon til å stole på teknologi fremfor tidligere erfaringer. Wikipedia har vært tilgjengelig

siden 2001 (Henriksen, 2024) og kan slik ha opparbeidet seg en etablert tillit som ikke er like avhengig av brukernes generelle disposisjon til å stole på nye verktøy som for ChatGPT.

Funnene viser at tilbøyelighet til tillit har en svak effekt på systemtillit for Google sammenlignet med ChatGPT. Dette resultatet var marginalt ikke-signifikant (p-verdi= 0,0665). Funnet kan derfor likevel være en indikasjon på at tilbøyelighet til tillit kan påvirke forskjeller i systemtillit mellom de to gruppene.

Overordnet viser analysen at tilbøyelighet til å stole på nye informasjonsverktøy har en positiv og signifikant påvirkning på systemtillit, men styrken av denne sammenhengen varierer mellom gruppene. Effekten er større for ChatGPT sammenlignet med Wikipedia. For ChatGPT sammenlignet med Google er forskjellen i effekt nesten signifikant. Modellens justerte R^2 -verdi er 0,225, noe som betyr at modellen forklarer 22,5 % av variasjonen i systemtillit. Dermed har vi grunnlag til å forkaste nullhypotesen og beholde hele H_{5-1} .

Det kan være flere grunner til at individets generelle tilbøyelighet til å stole på nye informasjonsverktøy påvirker tillit til informasjonsverktøy. En grunn kan være at individer med høy tilbøyelighet til tillit har en mer positiv innstilling til en teknologi sin kompetanse, noe som gjør at de har et generelt høyere forhånds nivå av tillit. De som er mer tilbøyelige til å stole på teknologier, kan også være mindre kritiske og bruke mindre tid på å evaluere om et informasjonsverktøy er pålitelig.

En annen grunn til at individer med høyere grad av tillitstilbøyning stoler mer på et verktøy, kan være fordi disse individene er mer tilbøyelige til å bruke verktøyet og dermed erfare positive opplevelser med dette, noe som kan øke tillit. På den andre siden kan individer med lav grad av tillitstilbøyelighet unngår å bruke verktøyene, og slik ikke oppdager de positive sidene med verktøyet.

Individer med lav tilbøyelighet til tillit kan være mer skeptiske til både nye og etablerte verktøy, og kanskje i større grad behøver bevis eller bekreftelse på at systemet er å stole på, før de faktisk har tillit til systemet. Disse individene kan også være mer fokusert på å avdekke feil eller svakheter i systemer, noe som gjør det vanskeligere å ha tillit til systemene.

Generell tilbøyelighet til å stole på nye informasjonsverktøy og mennesketillit

Vi gjorde tilsvarende analyser for å undersøke påvirkning av generell tilbøyelighet til å stole på nye informasjonsverktøy og mennesketillit (vedlegg 16). I likhet med analysene for systemtillit viser regresjonsanalysen at tilbøyelighet til tillit har en positiv og signifikant effekt

på mennesketillit, noe som gjelder for alle gruppene (p -verdi $< 0,001$) (se vedlegg 3). Modellen viser at for hver økning i tilbøyelighet til tillit, øker tillit med 0,46. Dette betyr at respondenter som generelt har høyere tilbøyelighet til tillit, også stoler mer på informasjonsverktøy.

Analysene for mennesketillit viser at Wikipedia-gruppen hadde en signifikant svakere effekt av tilbøyelighet til tillit på mennesketillit, sammenlignet med ChatGPT. Denne effekten hadde en p -verdi på 0,0013, og indikerer at tillit til Wikipedia er mindre påvirket av individuelle nivåer av tilbøyelighet til tillit enn ChatGPT. Det var ingen signifikante forskjeller mellom Google- og ChatGPT-gruppen.

Modellen er overordnet brukbar, da den har en justert forklaringskraft på 23,11%, noe som betyr at generell tilbøyelighet til tillit forklarer nesten 25% av variasjonen i mennesketillit. Her henviser vi til argumentasjon fra avsnittet om systemtillit angående grunner til hvorfor tillitstilbøyelighet kan ha en effekt på overordnet tillit.

Vi finner altså, gjennom våre analyser, at generell tilbøyelighet til å stole på nye informasjonsverktøy har en positiv påvirkning på mennesketillit for alle gruppene. Dermed forkaster vi nullhypotesen og beholder hele H_{5-2} .

6.2.6 Hypotesesett 6: Oppfattet grad av menneskelignende egenskaper sin påvirkning på tillit

Den første hypotesen i hypotesesett 6 handler om at brukerne vil oppfatte ChatGPT som mer menneskelig enn de andre på grunnlag av at ChatGPT kan tenkes å ha menneskelignende egenskaper fordi den er et kunstig intelligent verktøy.

Alle respondentene ($n=972$) svarte på spørsmål om hvor systemlignende versus menneskelignende de oppfattet ChatGPT, Wikipedia og Google. Totalt skåret respondentene alle de tre informasjonsverktøyene lavere enn 3, på en skala fra 1 til 5, hvor 1 tilsvarte «mye mer teknologiske egenskaper» og 5 tilsvarte «mye mer menneskelignende egenskaper». Overraskende nok fikk ChatGPT den laveste skåren på henholdsvis 1,88, mens Google og Wikipedia fikk henholdsvis 2,13 og 2,99. ChatGPT oppfattes derfor som mer systemlignende enn Google og Wikipedia. Dermed forkaster vi H_{6-1} , og beholder nullhypotesen $H_{6-1null}$.

Gruppe	Grad av menneskelignende egenskaper	Minimum	Maksimum
ChatGPT	1,88	1	5
Google	2,13	1	5
Wikipedia	2,99	1	5

Tabell 9: Oversikt over hvor systemlignende versus menneskelignende respondentene oppfattet de ulike informasjonsverktøyene.

Vi gjennomførte en lineær regresjonsanalyse for hver av informasjonsverktøyene for å teste om det er sammenheng mellom oppfattet grad av menneskelighet og tillit til verktøyet, både for systemtillit og mennesketillit.

Grad av menneskelignende egenskaper og systemtillit

For Google og Wikipedia finner vi ingen signifikant effekt av oppfattet grad av menneskelignende egenskaper på systemtillit, med p-verdier på henholdsvis 0,62 og 0,67. Analysene viste derimot at det var en sterkt signifikant effekt for ChatGPT-gruppen (p-verdi < 0,001) (vedlegg 17). Av den grunn beholder vi H_{6-2a} , men ikke H_{6-2b} og H_{6-2c} , da vi finner at økt grad av oppfattet menneskelighet øker systemtillit, men kun for ChatGPT og ikke for Google og Wikipedia. Vi forkaster derfor H_{6-3} .

Resultatene fra regresjonen viser at oppfatningen av grad av menneskelignende egenskaper for ChatGPT har en positiv og signifikant påvirkning på systemtillit. Dersom vi ikke tar hensyn til påvirkningen av grad av menneskelighet har respondentene i ChatGPT-gruppen et grunnleggende systemtillitsnivå på 4,76. Dersom respondenten oppfatter ChatGPT som rent teknologisk, øker systemtillitsnivået med 0,241. Dette gir et samlet forventet tillitsnivå på 5 (fra 1 til 7). Denne effekten er på et 0,001-signifikansnivå, noe som tyder på at funnet er sterkt signifikant.

Sammenhengen mellom grad av oppfattet menneskelighet og tillitsgrad er statistisk signifikant, men regresjonsmodellen har en justert forklaringskraft på 0,041, noe som tyder på at modellen forklarer 4.1 % av variasjonen i respondentenes tillitsnivå. Dette indikerer at det er andre faktorer som i større grad er med på å forklare tillit til ChatGPT, fremfor variasjon i oppfatningen av verktøyet som systemlignende eller menneskelignende.

Grad av menneskelignende egenskaper og mennesketillit

Vi fikk ingen signifikante utslag for Google og Wikipedia på grad av menneskelighet sin påvirkning på mennesketillit, med p-verdier på henholdsvis 0,44 og 0,52. Det var derimot en sterkt signifikant sammenheng mellom oppfattet grad av menneskelighet og mennesketillit for ChatGPT (vedlegg 18).

Resultatene fra regresjonen viser at det er en statistisk signifikant effekt på oppfatningen av verktøyet som menneskelignende eller systemlignende på overordnet mennesketillit. For hver enhets økning i oppfatningen av ChatGPT som *mer* menneskelig, øker mennesketillit med 0,261. Denne effekten er statistisk signifikant på 1%-nivå, noe som indikerer at hvor menneskelig respondenten oppfatter verktøyet, har stor påvirkning på hvor velvillig, kompetent og hvor stor integritet ChatGPT oppfattes å ha.

Modellen har en justert forklaringskraft på 0,045, noe som betyr at den bidrar til å forklare 4,5% av variasjonen i mennesketillit til ChatGPT. Disse funnene er tilsvarende som regresjonsanalysen med systemtillit. Dette indikerer at grad av oppfattet systemlignende eller menneskelignende egenskaper ikke er ulikt mellom mennesketillit og systemtillit. Dette virker rimelig da begge tillitsformene måler tillit, og den eneste forskjellen er hvilke spørsmål som anvendes for å kartlegge disse tillitsformene.

Vi beholder dermed H_{6-4a} , men ikke H_{6-4b} og H_{6-4c} , da vi finner at økt grad av oppfattet menneskelighet øker mennesketillit, men at dette kun gjelder for ChatGPT og ikke for Google og Wikipedia. Vi forkaster H_{6-5} .

6.2.7 Demovariabler sin påvirkning på tillit til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med advarsel

Demovariabler og systemtillit

For å avdekke om demografiske variabler, slik som kjønn, alder, utdanningsnivå, IT-bakgrunn og politisk tilhørighet, påvirker overordnet systemtillit gjennomførte vi en multippel lineær regresjonsanalyse for hver av de enkelte gruppene. Vi finner at det ikke er utslag på noen av de demografiske variablene for ChatGPT-gruppen, Wikipedia-gruppen eller gruppen som ble eksponert for ChatGPT med en advarsel. Regresjonsmodellen for Google viste derimot at kjønn påvirker overordnet systemtillitsnivå (p-verdi= 0,032) (vedlegg 20). Her viser resultatene at menn har lavere overordnet systemtillit til Google enn kvinner, ettersom kvinner er vår referansekategori.

Demovariabler og mennesketillit

Vi gjennomførte også en multippel lineær regresjonsanalyse for å undersøke om de demografiske variablene: kjønn, alder, utdanningsnivå, IT-bakgrunn og politisk tilhørighet, påvirker overordnet mennesketillit for de ulike informasjonsverktøyene.

Resultatene viser at det ikke er utslag på noen av de demografiske variablene for ChatGPT-gruppen, Google-gruppen, Wikipedia-gruppen eller gruppen som ble eksponert for ChatGPT med en advarsel (vedlegg 19). Derimot fikk vi et svakt utslag for alder (-0,0069), da vi kjørte en lineær regresjon med mennesketillit som avhengig variabel og demovariablene som uavhengige variabler. Her viser resultatene at effekten av alder på mennesketillit er svak (-0,0069), men likevel signifikant negativ (p-verdi < 0,05) (vedlegg 5). Dette indikerer at desto eldre respondentene er, jo lavere mennesketillit har de til informasjonsverktøy. Modellen har en justert forklaringskraft på 0,23%, noe som indikerer at modellen ikke bidrar til å forklare mye av variasjonen i mennesketillit, og at det er andre faktorer som har større påvirkning.

Syn på klimaforandringene

Til tross for lite utslag for demovariablene, fikk vi signifikante funn som viser at klimaskeptikere har lavere tillit til informasjonsverktøy enn de som ikke er det. I lys av at vårt overordnet tema for eksperimentet var klimaforandringer, la vi til et spørsmål om hvorvidt respondentene tror på at klimaforandringene er menneskeskapte. Dette var for å kartlegge hvorvidt skeptisisme til at klimaforandringene er menneskeskapte har en påvirkning på tillit til de ulike informasjonsverktøyene.

For å teste dette benyttet vi en lineær regresjonsanalyse for de ulike gruppene. Spørsmålet vi stilte respondentene i hver av gruppene var: «Tror du klimaforandringer skyldes menneskelig aktivitet?». Verdien 1 tilsvarer at respondenten mener klimaendringene er menneskeskapte, og verdien 2 at de mener klimaforandringene ikke er menneskeskapte.

Regresjonsmodellen for ChatGPT viser at det er en sterkt signifikant negativ sammenheng (p-verdi tilnærmet lik null) mellom systemtillit og respondentens syn på om klimaforandringene er menneskeskapte eller ikke (vedlegg 21). Respondentene som ikke tror klimaforandringene er menneskeskapte har 0,88 lavere systemtillit til ChatGPT enn de som mener klimaforandringene skyldes menneskelig aktivitet (se vedlegg 6). Regresjonsmodellens forklaringskraft er på 0,0547, noe som betyr at modellen forklarer 5,47% av variansen i systemtillit (se vedlegg 6).

Vi finner tilsvarende for Google og Wikipedia, hvor det også er en sterkt signifikant negativ sammenheng mellom systemtillit og respondentens syn på om klimaforandringene er menneskeskapte.

Overordnet er disse funnene interessante, da de viser at dersom respondentene mener klimaforandringene er menneskeskapte, har de mer tillit til informasjonsverktøyet. En forklarende faktor kan være at disse respondentene generelt har høyere tilbøyelighet til å stole på vitenskapelige kilder og faktaorienterte fremstillinger, noe som ofte er karakteristisk for moderne informasjonsverktøy som Google, ChatGPT og Wikipedia. En annen grunn kan være at vårt utvalg generelt var høyt utdannet og hovedsakelig bestående av demokrater. Det er vist at demokrater med høy utdanning er mer bekymret for klimaforandringene enn republikanere med høy utdanning (Hamilton, 2010).

For ChatGPT-gruppen med en advarsel viste regresjonsmodellen at respondentens syn på klimaendringene ikke hadde statistisk signifikant betydning for systemtilliten. Dette funnet kan være overraskende ettersom denne variabelen hadde signifikant betydning for alle de andre gruppene. Det kan være flere grunner til at det ikke ble utslag for denne gruppen. En grunn kan være fordi denne gruppen ble eksponert for en advarsel om at ChatGPT kan ta feil, noe som kan ha redusert tilliten generelt eller skapt en viss nøytralitet i svarene, som gjør at sammenhengen mellom variablene svekkes. Videre kan det tenkes at selve advarselen førte til mer ensartede svar på spørsmålene samlet sett, noe som reduserer variasjonen i dataene (se vedlegg 7).

Vi analyserte også om tro på at klimaforandringene er menneskeskapte eller ikke påvirker mennesketillit. Slik som for systemtillit, fant vi ikke signifikante forskjeller mellom gruppene. Vi fant likevel at respondentene som mener klimaforandringene ikke er menneskeskapte har signifikant lavere mennesketillit enn de som mener motsatt. Denne effekten er sterk signifikant (p -verdi $< 0,001$). Til forskjell for analysen av systemtillit, fikk vi derimot signifikant utslag for ChatGPT-gruppen med advarsel.

7. Oppsummerende diskusjon

I denne seksjonen skal vi diskutere resultatene fra analysen mer overordnet. Formålet med denne oppgaven har vært å besvare følgende forskningsspørsmål:

Hvordan er tillit til ChatGPT sammenlignet med Google og Wikipedia, og vil tilliten påvirkes av en advarsel om at ChatGPT selv informerer om at den kan ta feil?

Bakgrunnen for forskningsspørsmålet er den økende bruken av ChatGPT og om brukerne stoler blindt på et verktøy som selv advarer om at det kan gjøre feil. Det har blitt reist økende bekymring knyttet til bias og feilinformasjon i ChatGPT, og konsekvensene av for høy tillit kan derfor være betydelige, spesielt dersom verktøyet kan spre unøyaktig informasjon.

Formålet med vår studie har derfor vært å kartlegge i hvilken grad tillitsnivået til ChatGPT skiller seg fra eller er tilsvarende tilliten til de etablerte informasjonsverktøyene Google og Wikipedia. Disse har vært på markedet i over 20 år og anses derfor som etablerte aktører sammenlignet med ChatGPT, noe som gir et solid sammenligningsgrunnlag. Ved å sammenligne tillitsnivået mellom disse tre aktørene, ønsket vi å utforske om ChatGPT kan representere fremtidens informasjonsverktøy og potensielt utkonkurrere de mer tradisjonelle og godt etablerte verktøyene.

Den andre delen av forskningsspørsmålet omhandler effekten av en synliggjøring av at ChatGPT selv fremhever at den kan ta feil. Vi ønsker å undersøke hvorvidt en slik advarsel påvirker individers tillit til verktøyet, eller om den har liten praktisk betydning for tillitsnivået.

For å besvare forskningsspørsmålet gjennomførte vi et eksperiment med fire grupper. Hver gruppe ble bedt om å se for seg at de skulle innhente informasjon om klimaforandringene, før de svarte på spørsmål om tillit til verktøyet de var tildelt: ChatGPT, Google, Wikipedia eller ChatGPT med en advarsel.

Våre resultater viser at tilliten til ChatGPT er lavere enn tilliten til Google. Dette kan skyldes at Google oppfattes som mer funksjonell og pålitelig av respondentene, selv om begge verktøyene vurderes som like hjelpsomme i informasjonssøk. Funnet indikerer at ChatGPT for øyeblikket anses som mindre tillitsverdige enn Google. Samtidig var forskjellene i tillitsnivå mellom disse to tjenestene relativt små, noe som kan tyde på at denne dynamikken kan endre seg i fremtiden. ChatGPT representerer allerede en utfordrer til Google, da begge verktøyene hovedsakelig fungerer som søkemotorer og brukes til informasjonssøk.

Vi finner videre at ChatGPT oppfattes som mer hjelpsom enn Wikipedia, men at det ikke er noen forskjell i hvordan brukerne vurderer funksjonalitet og pålitelighet mellom de to verktøyene. Overraskende ble ChatGPT ansett som mer tillitsverdig enn Wikipedia, noe som står i kontrast til funnene fra en tidligere kvalitativ studie som viste det motsatte (Jung, et al., 2024). Samme studie fant imidlertid at en synliggjøring av hvor ChatGPT henter informasjonen fra, potensielt øker tilliten til verktøyet sammenlignet med andre informasjonsverktøy. Dette antyder at dersom ChatGPT implementerer en slik funksjonalitet, kan dette øke reliabiliteten til verktøyet, som igjen kan øke generelt tillitsnivå til tjenesten. Likevel er det viktig å påpeke at denne tidligere kvalitative studien hadde et svært begrenset utvalg (n=14), som kan ha påvirket generaliserbarheten, siden utvalget trolig ikke er representativt.

Ingen signifikante forskjeller ble funnet mellom ChatGPT og ChatGPT med en advarsel. Vi hadde forventet at en advarsel om at ChatGPT fremhever egne feil ville redusere tilliten, men det er mulig at advarselen ikke ga utslag fordi vi ikke synliggjorde den nok. Advarselen var utformet på samme måte som den eksisterende meldingen i ChatGPTs grensesnitt, noe som kan ha medført at respondentene ikke tenkte nevneverdig over denne, hvis de allerede er kjent med meldingen. En annen form for advarsel, eksempelvis ved å tvinge respondentene til å aktivt godta advarselen, eller bekrefte at de var klar over at verktøyet tar feil, kunne ha gitt et annet resultat.

Våre funn viser at det ikke er signifikante utslag for mennesketillit eller de tilhørende tillitsdimensjonene; evne, velvilje og integritet for sammenligningen mellom de ulike gruppene. Dette er en indikasjon på at systemtillitsdimensjonene er bedre skikket til å måle tilliten til ChatGPT, Google og Wikipedia, enn mennesketillitsdimensjonene. Denne indikasjonen samsvarer med teorien til Lankton et al (2015), som argumenterer for at bruk av menneskelige eller systembaserte tillitsmål kan gi ulike empirisk resultater. Våre funn støtter denne teorien, ved at vi finner signifikante forskjeller når vi benytter systembaserte tillitsdimensjoner, men ikke når vi benytter mennesketillitsdimensjoner.

Brukstid og brukshyppighet, som begge måler individenes brukserfaring med verktøyene, påvirker tillit til alle verktøyene positivt. Resultatene viser at jo lengre et individ har benyttet verktøyet, og jo oftere det brukes, desto høyere tillit har individet til verktøyet. Dette gjelder på tvers av alle gruppene, noe som betyr at ChatGPT ikke skiller seg vesentlig ut sammenlignet med Google og Wikipedia når det gjelder hvordan brukserfaring påvirker tillit. Samlet sett har brukserfaring en positiv effekt på bruk av informasjonsverktøy.

Generell tilbøyelighet til å stole på ny teknologi påvirker også tillitsnivå på tvers av alle gruppene. De som er generelt mer tilbøyelige til å stole på nye teknologier, viser høyere tillit til informasjonsverktøyene. Dette funnet antyder at personlighetsfaktorer spiller en viktig rolle på tillit til spesifikke verktøy.

Et interessant funn er økt grad av oppfattede menneskelignende egenskaper ved verktøyet øker tilliten kun til ChatGPT. Dette var et interessant funn, siden forskningen er delt på når man bør benytte mennesketillitsdimensjoner eller systemtillitsdimensjoner for å måle tillit. Overraskende nok oppfattes ChatGPT som det minst menneskelignende verktøyet til tross for at virtuelle KI-verktøy ofte tilskrives menneskelignende egenskaper (Glikson & Woolley, 2020).

For kontrollvariablene fant vi kun at respondentene syn på klimaforandringene som menneskeskapte eller ikke, påvirker tillitsnivået. Overordnet for alle gruppene viser de som mener klimaendringene skyldes menneskelig aktivitet høyere tillit enn de som mener motsatt. Dette funnet understreker viktigheten av kontekst og situasjon for vurdering av tillit. Funnene våre indikerer at individets personlige mening om et gitt tema kan ha påvirkning på tilliten.

Det er også viktig å påpeke at tillit er dynamisk og kan endre seg over tid (Chang, et al., 2005) (Chang, et al., 2005). Selv om vi fant høy tillit for enkelte tillitsdimensjoner i eksperimentet, kan dette endres etter hvert som brukerne får mer informasjon eller erfaring med verktøyene.

Samlet sett gir våre funn en indikasjon på at tilliten til ChatGPT er lavere enn tilliten til Google, men høyere enn tilliten til Wikipedia. Videre viser resultatene at brukserfaring, tilbøyelighet til å stole på teknologi og individuelle oppfatninger om verktøyenes menneskelighet spiller en viktig rolle i tillitsvurderinger.

8. Konklusjon

I den siste delen av oppgaven skal vi besvare forskningsspørsmålet presentert innledningsvis. Videre skal vi gå inn på oppgavens begrensinger, før vi presenterer praktiske implikasjoner av funnene våre. Avslutningsvis kommer vi med innspill til videre forskning.

Hensikten med denne oppgaven har vært å undersøke hvordan tillitsnivået til ChatGPT er sammenlignet med Google og Wikipedia, samt om en advarsel om at ChatGPT kan gjøre feil påvirker tilliten til verktøyet. For å kunne besvare dette forskningsspørsmålet, har vi først

presentert relevant teori knyttet til tillitsbegrepet, inkludert både menneskelige tillitsdimensjoner og systembaserte tillitsdimensjoner.

Resultatene våre viser at Google oppfattes som mer pålitelig og funksjonell enn ChatGPT, men at forskjellen i tillit mellom de to verktøyene er relativt liten. Dette antyder at ChatGPT allerede utfordrer etablerte aktører som Google, men enda ikke har oppnådd tilsvarende tillitsnivå. Samtidig oppfattes ChatGPT som mer hjelpsom enn Wikipedia, noe som indikerer at brukerne ser en verdi i tjenesten til tross for at tillitsnivået ikke er like høyt som for Google.

Et interessant funn er at en advarsel om at ChatGPT kan gjøre feil ikke påvirket respondentenes tillit signifikant. Dette kan skyldes at advarselen ikke var utformet tydelig nok, eller at respondentene allerede var kjent med denne informasjonen fra tidligere erfaringer med verktøyet. Alternativt kan en mer eksplisitt advarsel, der brukerne aktivt må bekrefte at de har forstått risikoen for at plattformen kan generere feilinformasjon, hatt en større effekt.

Videre finner vi at brukserfaring spiller en betydelig rolle i tillitsvurderinger. Både hyppig bruk og langvarig erfaring med verktøyene øker tilliten, noe som gjelder for alle informasjonsverktøyene i studien. Dette viser at erfaring med verktøyet bidrar til å styrke tilliten over tid. Vi ser også at individers generelle tilbøyelighet til å stole på ny teknologi har en positiv effekt på tillitsnivået.

Et annet sentralt funn er sammenhengen mellom opplevde menneskelignende egenskaper ved verktøyene og tillit. Jo mer menneskelig verktøyet oppfattes som, desto høyere tillit har brukerne til det. Overraskende nok ble ChatGPT oppfattet som verktøyet med minst menneskelignende egenskaper, til tross for at tidligere forskning har vist at KI-verktøy ofte tilskrives slike egenskaper.

Når det gjelder kontekst og situasjon, fant vi at individers oppfatning av klimaforandringer som menneskeskapt eller ikke, påvirket tilliten til informasjonsverktøyene. Respondenter som mente klimaendringene skyldes menneskelig aktivitet, hadde generelt høyere tillit til verktøyene enn de som mente det motsatte. Dette funnet illustrerer hvordan kontekst og personlige meninger spiller en viktig rolle i tillitsvurderinger.

Samlet sett bidrar våre funn til å belyse hvordan tillit til ChatGPT er i forhold til etablerte informasjonsverktøy som Google og Wikipedia. Selv om ChatGPT har utfordret de tradisjonelle aktørene, er det fremdeles forskjeller i hvordan verktøyene oppfattes, særlig med hensyn til funksjonalitet, hjelpsomhet og pålitelighet. Samtidig gir resultatene innsikt i hvilke

faktorer som påvirker tillitsnivået, inkludert brukserfaring, generell tilbøyelighet til å stole på nye teknologier, og opplevde menneskelige egenskaper. Dette er viktige momenter å vurdere når vi ser på ChatGPTs potensiale som fremtidens informasjonsverktøy.

8.1 Begrensninger

I det følgende skal vi se på oppgavens begrensninger, da disse kan påvirke endelig utfall og ha betydning for funnene. For det første finner vi at en synliggjøring av at ChatGPT kan ta feil, ikke har noen effekt på tillit. Dette kan skyldes at vi i utformingen av eksperimentet ikke har klart å tydeliggjøre denne advarselen godt nok, slik at det ikke oppfattes noen reelle forskjeller mellom de to ChatGPT-gruppene i eksperimentet. Dersom vi hadde valgt å benytte en annen form for advarsel som var enda mer fremtredende, kan det være dette ville gitt utslag. Dette er dog kun spekulasjoner, men det kan tenkes at en sterkere synliggjøring ville gitt større differanse mellom de to gruppene.

En annen begrensning er at vi kun har målt *tillitsfulle overbevisninger*, som forklart i teoridelen. Med dette mener vi at vi kun har målt respondentenes oppfatninger av tillit til verktøyene, men ikke faktisk intensjon eller atferd tilknyttet bruken av dette. Ettersom eksperimentet vårt var fiktivt og uten at svarene hadde konsekvenser, kan det tenkes at respondentene hadde svart annerledes dersom de faktisk skulle innhente informasjon om klimaforandringene på de respektive plattformene og benyttet denne informasjonen. Vår studie er et lab-eksperiment, og kan derfor sees på som stilisert, noe som kan svekke den overordnede generaliserbarheten. Dette er fordi respondentene kan ha andre meninger og oppfatninger i naturlige settinger. I en mer reell setting kan det for eksempel være at respondentene hadde hatt et annet tillitsnivå til ChatGPT fordi individers tillitsoverbevisning ikke nødvendigvis korrelerer med tillitsfull atferd.

Knyttet opp til dette er det også en svakhet at vi har benyttet spørreundersøkelse på nett i vår studie. Eksperimentet foregår dermed ikke i en naturlig setting, og det kan ha påvirket hvordan respondentene vurderte de ulike verktøyene. Dersom eksperimentet skulle vært mer reelt kunne vi gjennomført undersøkelsen fysisk og gitt deltakerne i oppgave å innhente informasjon om klimaforandringene, for deretter å fremføre dette foran et ekspertpanel. Parallelt med dette kunne vi også latt deltakerne selv velge hvilket verktøy de skulle benyttet for denne oppgaven, da dette valget ville reflektert faktisk tillitsfull atferd, og ikke kun tillitsfull overbevisning slik som i vårt eksperiment. Grunnet tid- og kostnadsbegrensninger var ikke dette et mulig alternativ for oss.

Det kan også tenkes at det er andre tillitsmål som kunne vært aktuelle å benytte for å måle tillit, og at disse kan være mer passende enn de vi har benyttet. Likevel anser vi våre spørsmål som relativt valide, da de er anvendt i tidligere studier og forskning. I tillegg har vi både inkludert spørsmål som kartlegger de menneskelige tillitsdimensjonene og de systembaserte tillitsdimensjonene. Til tross for dette, ser vi det mulig at andre spørsmål kunne vært benyttet, og kunne bidratt til å illustrere forskjeller i tillit. Dette svekker studiens begrepsvaliditet.

En annen begrensning ved vår studie er at vi har benyttet et utvalg som kun består av amerikanere. Dette svekker generaliserbarheten til andre nasjonaliteter og deler av verden, da det kan tenkes at amerikanere har et annet syn på KI eller informasjonsverktøy enn andre kulturer. Videre er også vårt utvalg hentet fra Profilic, som er en profesjonell datainnsamlingsplattform. Individene på Profilic har ofte mye erfaring med undersøkelser, og kan ha høy teknologisk kompetanse da de er profesjonelle. Dette kan være en svakhet, da den teknologiske kompetansen kan være en variabel som systematisk har påvirket resultatene. Våre respondenter hadde generelt et høyt utdanningsnivå, noe som systematisk også kan ha påvirket resultatene.

Tidsaspektet er også en begrensning ved oppgaven. Vi har utført studien på et semester, noe som ikke har gjort det mulig å gjennomføre en mer omfattende studie, slik som eksempelvis nevnt ovenfor. Vi kunne også sett nærmere på hvordan tillit endres over tid, noe som ville vært spennende da ChatGPT fremdeles er et relativt nytt verktøy, som stadig flere bruker og som ofte kommer med nye funksjoner. ChatGPT har en ekstrem fart i utviklingen, og er i konstant endring (Nyst, 2024). En rask og konstant utvikling vil trolig føre til en endring i tillitsnivået til chatboten. Det kan også tenkes at tillitsforskjellene mellom ChatGPT og de to andre verktøyene vil endres, da Google og Wikipedia per dags dato ikke har samme endringsakselerasjon som ChatGPT.

Et viktig aspekt å trekke frem til slutt er også at tillit er et dynamisk begrep og kan endres over tid. Tilliten til ChatGPT og de andre verktøyene vil derfor trolig ikke være konstant i tiden fremover, spesielt med tanke på at EU sin AI Act skal tre i kraft. Den hyppige utviklingen i KI-markedet kan også påvirke tillit til KI-tjenester, slik som ChatGPT.

8.2 Teoretiske implikasjoner

I vår oppgave har vi undersøkt om det er signifikante forskjeller i tillit til tre av markedets mest brukte og kjente verktøy for informasjonsinnhenting, samt om synliggjøring av at ChatGPT kan ta feil påvirker tillitsnivået til plattformen. Funnene bidrar til ny innsikt i litteraturen da de

gir en dypere forståelse av dagens tillitssituasjon, hvor vi ser en endring fra funn i tidligere forskning. En tidligere studie konkluderte med at individer stoler mest på Google, deretter Wikipedia og til slutt ChatGPT (Jung, et al., 2024). Vår studie støtter dette ved å bekrefte at Google fortsatt oppfattes som det mest tillitsfulle verktøyet. Samtidig motstrider studien vår også denne forskningen, ettersom våre funn indikerer at brukere i dag stoler mer på ChatGPT enn Wikipedia. Dette funnet er signifikant og sannsynligvis generaliserbart, ettersom vår studie hadde et representativt utvalg på 972 respondenter, sammenlignet med den tidligere studien som var kvalitativ og baserte seg på et lite utvalg (n=14). Dette bidrar til mer robuste og generaliserbare funn, samtidig som det understreker at tillit er dynamisk og kan endre seg over tid. Funnene tyder på at oppslutningen og tillitsnivået til ChatGPT øker, siden plattformen har forbigått Wikipedia.

Våre funn indikerer ikke at en synliggjøring av at ChatGPT kan ta feil påvirker tilliten til plattformen i betydelig grad. Dette gir en viktig innsikt i litteraturen ved å antyde at individer (1) ignorerer advarsler, (2) allerede vet at ChatGPT kan ta feil, men ikke lar dette påvirke sin tillit, eller (3) oppfatter advarselen som irrelevant. Selv om vi ikke kan fastslå at en advarsel har effekt, viste våre data at ChatGPT med advarsel hadde noe lavere skår på overordnet tillitsnivå for både mennesketillit og systemtillit sammenlignet med ChatGPT uten advarsel. Dette funnet var derimot ikke statistisk signifikant, noe som peker på behovet for ytterligere studier på dette feltet. Studien vår kan derfor fungere som et grunnlag eller en inspirasjon for fremtidig forskning.

Når det gjelder systemtillit, viste vår studie større spredning i svarene for systemtillitsdimensjonene, med flere utslag mellom gruppene på disse spørsmålene. Både Google og ChatGPT ble oppfattet som mer systemlignende enn menneskelignende, noe som er i tråd med tidligere forskning som viser at systemtillit er mer relevant for teknologiske verktøy (Lankton, et al., 2015). Dette styrker eksisterende funn i litteraturen og understreker at det er empirisk utslagsgivende å benytte korrekte tillitsdimensjoner for å fange opp variasjoner i tillit til ulike teknologier.

ChatGPT ble av våre respondenter ansett som mer systemlignende enn menneskelignende, noe som kan utfordre eksisterende forskning som hevder at virtuelle KI-er ofte tilskrives menneskelignende egenskaper og funksjoner (Pelau, et al., 2021; Chandra, et al., 2022; Tian, et al., 2017). Dette funnet viser at til tross for implementeringen av menneskelignende

funksjoner, som stemmebaserte interaksjoner, oppfatter brukerne fortsatt ChatGPT primært som et system. Om denne oppfatningen vil vedvare, krever videre forskning

8.3 Praktiske implikasjoner

Vi har funnet at Google oppfattes som mer tillitsfull enn ChatGPT, hovedsakelig fordi Google er vurderes som mer funksjonell og pålitelig. Denne informasjonen kan være relevant for OpenAI dersom de ønsker å øke sin oppslutning. For å kapre større markedsandeler, spesielt fra Google-brukere, bør OpenAI fokusere på å forbedre ChatGPTs funksjonalitet og pålitelighet.

OpenAI kan potensielt øke den oppfattede funksjonaliteten ved å tilby flere tjenester, slik som Google gjør, for eksempel kartfunksjoner, akademiske søk eller bilde- og videospesifikke søk. For å forbedre opplevd pålitelighet kan de øke transparensen ved å konsekvent vise hvor algoritmen henter informasjon fra (von Eschenbach, 2021). Dette kan potensielt øke tillitsnivået til ChatGPT og bidra til at flere brukere velger denne plattformen fremfor Google.

For Google indikerer funnene våre at de for øyeblikket fortsatt har en høyere grad av tillit blant respondentene, men forskjellen sammenlignet med ChatGPT er marginal. Med ChatGPTs stadige utvikling kan tillitsnivået til KI-chatboten endres og potensielt øke i fremtiden. Dette indikerer at Google bør kaste seg på KI-bølgen, og fortsette å implementere KI-løsninger i sine tjenester, for å møte den økende konkurransen i markedet. Videreutvikling av deres KI-chatbot, Gemini, er viktig for at Google skal opprettholde sin posisjon som en ledende aktør for informasjonsinnhenting.

For Wikipedia viser funnene våre at ChatGPT oppfattes som mer hjelpsom. Dette kan påvirke Wikipedia på to måter. For det første kan økt bruk av ChatGPT føre til at innhold generert av KI blir inkludert i Wikipedia, noe som kan føre til en større likhet mellom plattformene og potensielt redusere forskjellene i tillitsnivå. For det andre kan forskjellene i tillit øke dersom ChatGPT fortsetter sin raske vekst og implementerer nye funksjoner mens Wikipedia forblir uendret. Dette kan føre til et økende tillitsgap mellom de to plattformene.

Våre funn viser også at en advarsel om at ChatGPT kan ta feil, ikke har en målbar påvirkning på brukernes tillitsnivå. En praktisk implikasjon av dette er at advarselen slik den er i dag, ikke ser ut til å effektivt påvirke brukernes bevissthet om plattformens begrensninger. Dette er viktig fordi plattformen har vist seg å være partisk eller ha skjevheter, og i enkelte tilfeller generere feilaktig eller skadelig informasjon (Ray, 2023). Dersom brukerne ikke er klar over

begrensningene, kan det føre til ukritisk bruk og feilinformasjon, noe som kan ha alvorlige konsekvenser. Eksempelvis kan det medføre en skjev forståelse av sensitive temaer.

Dersom flere individer blir kjent med ChatGPT og bruker verktøyet oftere, kan dette øke tilliten til plattformen. OpenAI kan dra nytte av dette funnet ved å gjøre ChatGPT mer tilgjengelig, for eksempel ved å utvikle utvidelser for populære nettlesere som Chrome og Safari, eller gjennom samarbeid med plattformer som Slack. Et annet forslag er å tilby gratis tilgang til ChatGPT 4.0 for utdanningsinstitusjoner.

OpenAI kan også arbeide med å tilføre ChatGPT flere menneskelignende egenskaper for å styrke tilliten. Dette kan inkludere videreutvikling av stemmefunksjon og implementering av hukommelse på tvers av samtaletråder (Knudsen, 2024). Videre kan de gjøre plattformen mer empatisk ved å programmere verktøyet til å uttrykke mer sympati og følelser i svarene, slik at den viser mer forståelse for brukerens situasjon. De kan også videreutvikle ChatGPT til å inkludere en visuell avatar som kan uttrykke følelser. Slike endringer kan bidra til at ChatGPT oppfattes som mer menneskelignende, noe som kan styrke både tillit og brukervennlighet

8.4 Anbefalinger til videre forskning

For videre forskning anbefaler vi å undersøke om funnene er gyldige og kan overføres til situasjoner som kartlegger tillitsfull atferd, og ikke kun tillitsfull overbevisning. Det ville vært interessant å se på hvorvidt våre funn er aktuelle i situasjoner hvor individer må ta et aktivt valg mellom de ulike informasjonsverktøyene. Dette fordi det kan tenkes at resultatene ville vært annerledes i et slikt tilfelle, da konsekvenser tilknyttet valget kan påvirke individets atferd. Et eksperiment som gjennomføres i en reell kontekst er derfor å anbefale for videre forskning.

Det ville også vært interessant å undersøke hvordan skårene på de ulike tillitsdimensjonene påvirker type bruk av ChatGPT. For eksempel om individer som skårer plattformen høyest på velvilje, bruker ChatGPT mer til samtaler enn de som skårer den lavere på denne tillitsdimensjonen.

Vi anbefaler å forske videre på om bruken av ChatGPT i ulike kontekster har påvirkning på tillit. Med dette viser vi til at tillit til ChatGPT ved informasjonsinnhenting kan skille seg fra tillit til ChatGPT i andre brukssituasjoner. Eksempelvis medisinsk rådgivning, reisetips, produksjon av tekst eller oversettelse av språk.

En annen interessant problemstilling for videre forskning er hvordan ChatGPTs evne til å tilpasse seg brukeren påvirker tillit. I enkelte situasjoner kan brukeren påvirke ChatGPT til å

støtte en bestemt oppfatning eller justere svarene sine etter brukerens ønsker. Dette reiser spørsmål om hvorvidt slik tilpasning undergraver tilliten til ChatGPT som en objektiv informasjonskilde, noe som vil være relevant for fremtidige studier å undersøke.

Våre funn gir et øyeblikksbilde av tillitsnivået individer har til ChatGPT sammenlignet med Google, Wikipedia og ChatGPT med en advarsel. Studien vår danner dermed et viktig grunnlag for å forstå dagens tillitsnivå til ChatGPT, og kan fungere som et sammenligningsgrunnlag for fremtidige studier som undersøker tillit til ChatGPT og andre informasjonsverktøy.

Med den raskt akselererende utviklingen av KI-verktøy, ChatGPT inkludert, forventes det at dette påvirker hvordan vi oppfatter disse verktøyene og tilliten til dem. Et eksempel er OpenAIs lansering av en ny funksjon 13. desember 2024, som gjør det mulig for brukere å filme seg selv og omgivelsene samtidig som de har en løpende dialog med chatboten (Wiggers, 2024). I tillegg kan tillitsnivået påvirkes av nye lover, reguleringer eller økt kunnskap om KI-verktøy i samfunnet. Vi foreslår derfor å gjenta studien vår om noen år, for å undersøke om tilliten til de ulike verktøyene har endret seg. Dette spesielt i lys av den pågående KI-revolusjonen (Makridakis, 2017). Det ville også vært interessant å kartlegge tillitsnivået til ulike KI-chatboter for å avdekke om det er store forskjeller mellom de forskjellige variantene. For eksempel kunne man sammenligne tillit til Copilot og Gemini, som også er store aktører på KI-markedet.

Vi anbefaler å gjennomføre en kvalitativ studie for å utforske hvorfor individer har ulik tillit til de forskjellige verktøyene, samt få en dypere forståelse av hva som ligger til grunn for at man velger å stole på et informasjonsverktøy. En slik tilnærming vil kunne supplere våre funn med mer nyansert innsikt i årsakene til variasjonene i tillitsnivå.

Litteraturen peker på flere faktorer som kan bidra til å øke tilliten til teknologier og KI. Eksempler inkluderer hvordan KI presenteres visuelt eller interaktivt (Glikson & Woolley, 2020), reduksjon av oppfattet risiko ved bruk (Li, et al., 2008), forbedret brukervennlighet (Xu, et al., 2014) og økt opplevelse av maskinintelligens (Glikson & Woolley, 2020). Gitt at disse faktorene kan ha en betydelig innvirkning på tillit, vil det være verdifullt å avdekke hvilke spesifikke tiltak som kan implementeres for å styrke tilliten til informasjonsverktøy som ChatGPT

Avslutningsvis anbefaler vi å undersøke hvordan ulike former for advarsler påvirker brukeres tillit til ChatGPT og andre KI-baserte søkemotorer, spesielt når det gjelder fakta sjekk av generert innhold. Våre resultater viste ingen effekt av den standardiserte advarselen vi benyttet, noe som kan tyde på at andre måter å synliggjøre systemets feilmargin på bør utforskes. Ved å teste ulike typer advarsler, kan man få innsikt i hvilke tilnæringer som faktisk fanger brukernes oppmerksomhet og har en reell effekt på tillitsnivået.

Et mulig tiltak kan være å merke alt KI-generert innhold tydelig. Dette innebærer at innhold som er produsert av KI får en merkelapp eller varsel som gjør brukerne bevisste på at materialet ikke er menneskeskapt og potensielt kan inneholde feil. En lignende praksis er allerede etablert i Norge, der innhold som er redigert eller manipulert på sosiale medier må merkes i henhold til lovpålagte krav (Grant, 2021). Å avdekke hvilke typer advarsler eller merking som har effekt, er avgjørende for å bevisstgjøre brukere om behovet for å kvalitetssikre KI-generert innhold og dermed redusere risikoen for feilinformasjon.

9. Litteraturliste

- Abdi, H. & Williams, L. J., 2010. Tukey's Honestly Significant Difference (HSD) Test. I: *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage, pp. 1-5.
- Al-Ashwal, F. Y. et al., 2023. Evaluating the Sensitivity, Specificity, and Accuracy of ChatGPT-3.5, ChatGPT-4, Bing AI, and Bard Against Conventional Drug-Drug Interactions Clinical Tools. *Drug, Healthcare and Patient Safety*, 20 September, pp. 137-147.
- Aldas-Manzano, J., Ruiz-Mafe, C., Sanz-Blas, S. & Lassala-Navarré, C., 2011. Internet banking loyalty: evaluating the role of trust, satisfaction, perceived risk and frequency of use. *The Service Industries Journal* , 9 September, pp. 1165-1190 .
- Alexander, V., Blinder, C. & Zak, P. J., 2018. Why trust an algorithm? Performance, cognition, and neurophysiology. *Computers in Human Behavior*, Desember, pp. 279-288.
- Altman, D. G. & Bland , J. M., 1995. Statistics notes: the normal distribution. *BMJ*, 4 Februar, p. 298.
- Andrade, C., 2018. Internal, External, and Ecological Validity in Research Design, Conduct, and Evaluation. *Indian Journal of Psychological Medicine*, September-Oktober, p. 498–499.
- Anon., n.d. *An Introduction to Generative AI Vs. AI*. [Online]
Available at: <https://monetate.com/resources/blog/generative-ai-vs-ai/>
- Armstrong, K., 2023. *ChatGPT: US lawyer admits using AI for case research*. [Online]
Available at: <https://www.bbc.com/news/world-us-canada-65735769>
[Accessed 6 Desember 2024].
- Arnold, D. M. et al., 2009. The design and interpretation of pilot trials in clinical research in critical care. *Critical Care Medicine*, Januar.
- Ayers, P., Matthew, C. & Yates, B., 2008. *How Wikipedia Works: And how You Can be a Part of it*. s.l.:No Starch Press.
- Azaria, A., Azoulay, R. & Reches, S., 2024. ChatGPT is a Remarkable Tool—For Experts. *Data Intelligence*, p. 240–296.
- Önköl, D. et al., 2009. The Relative Influence of Advice From HumanExperts and Statistical Methods onForecast Adjustments. *Journal of Behavioral Decision Making*, Februar, p. 390–409.
- Back, I. H., 2010. Commitment bias: mistaken partner selection or ancient wisdom?. *Evolution and Human Behavior*, Januar, pp. 22-28.

Backlinko Team, 2024. *ChatGPT / OpenAI Statistics: How Many People Use ChatGPT?*. [Internet] Available at: <https://backlinko.com/chatgpt-stats>

[Funnet 24 Oktober 2024].

Bailyn, E., 2024. *Top Generative AI Chatbots by Market Share – December 2024*. [Online]

Available at: <https://firstpagesage.com/reports/top-generative-ai-chatbots/>

[Accessed 11 Desember 2024].

Baker, A. L., Phillips, E. K., Ullman, D. & Keebler, J. R., 2018. Toward an Understanding of Trust Repair in Human-Robot Interaction: Current Research and Future Directions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 16 November, pp. 1-30.

Banavar, D. G., 2016. *Learning to trust artificial intelligence systems*, Somers, NY 10589, USA.: IBM Corporation.

Banerjee, . A. et al., 2009. Hypothesis testing, type I and type II errors. *Industrial Psychiatry Journal*, Juli - Desember, pp. 127-131.

Basu, C. & Singhal, M., 2016. Trust Dynamics in Human Autonomous Vehicle Interaction: A Review of Trust Models. *Association for the Advancement of Artificial Intelligence*, pp. 85-91.

Bedué, P. & Fritzsche, A., 2021. Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *Esmerald Insight*, 11 April, pp. 530-549.

Benbasat, I. & Wang, W., 2005. Trust In and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, Februar.

Bhattacharjee, A., 2002. Individual Trust in Online Firms: Scale Development and Initial Test. *Journal of Management Information Systems*, Mai, pp. 211-242.

Boos, D. & Stefanski, L. A., 2011. P-Value Precision and Reproducibility. *The American Statistician*, November, pp. 213-221 .

Borgonovi, F., 2012. The relationship between education and levels of trust and tolerance in Europe. *The British Journal of Sociology*, 8 Mars, pp. 146-167.

Bravo, V. J. & Young, M. F., 2011. The Impact of a Collaborative Wikipedia Assignment on Teaching, Learning, and Student Perceptions in a Teacher Education Program. *Canadian Journal of Learning and Technology*.

Brewer, M. B., 2008. Depersonalized trust and ingroup cooperation. I: J. I. Krueger, red. *Rationality and social responsibility: Essays in honor of Robyn Mason Dawes*. s.l.:Psychology Press, p. 215–232.

Bristol, H. et al., 2024. *Adopting AI at speed and scale: The 4IR push to stay competitive*. [Internett]
Available at: <https://www.mckinsey.com/capabilities/operations/our-insights/adopting-ai-at-speed-and-scale-the-4ir-push-to-stay-competitive>

[Funnet 14 Desember 2024].

Brooks., J. L., 2012. Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, p. 600–614.

Brysbaert, M., 2019. How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis with Reference Table. *Journal of Cognition*, Juli.

Buchanan, J. & Hickman, W., 2024. Do people trust humans more than ChatGPT?. *Journal of Behavioral and Experimental Economics*, 31 Mai, pp. 1-6.

Bughin, J. et al., 2018. *Notes from the AI frontier: Modeling the impact of AI on the world economy*. [Internett]

Available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy#/>

[Funnet 21 Oktober 2024].

Burns, N., Kinder, D. & Rahn, W., 2003. *Social Trust and Democratic Politics*, Chicago Illinois: Annual Meeting of the Mid-West Political Science Association.

Burr, J., 2023. *9 ways we use AI in our products*. [Internett]

Available at: <https://blog.google/technology/ai/9-ways-we-use-ai-in-our-products/#:~:text=Google%20Cloud%20has%20built%20AI,at%20scale%2C%20and%20many%20other>

[Funnet 14 Desember 2024].

Cadario, R., C. L. & Morewedge, C. K., 2021. Understanding, explaining, and utilizing medical artificial intelligence. *Nature Human Behaviour*, 28 Juni, pp. 1636-1642.

Cahour, B. & Forzy, J.-F., 2009. Does projection into use improve trust and exploration? An example with a cruise control system. *Safety Science*, November, pp. 1260-1270.

Calvin, K. & Karsh, B.-T., 2009. A Systematic Review of Patient Acceptance of Consumer Health Information Technology. *Journal of the American Medical Informatics Association*, Juli, p. 550–560.

Campbell, D. T. & Stanley, J. C., 1963. *Experimental and Quasi-Experimental Designs for Research*. U.S.A: Houghton Mifflin Company .

Carroll., J. M., 2005. Chapter 2: Human Computer Interaction (HCI). I: C. Ghaoui, red. *Encyclopedia of Human Computer Interaction*. s.l.:Information Science Reference , pp. 21-42.

Cesarini, D. et al., 2008. Heritability of cooperative behavior in the trust game. 11 Mars, pp. 3721-3726.

Chakraborty, A., 2024. *Enhancing AI accuracy by nudging users to catch generative AI errors*. [Online]

Available at: <https://fortune.com/2024/09/04/accenture-study-enhance-how-to-enhance-ai-accuracy/>

Chandra, S., Shirish, A. & Srivastava, S. C., 2022. To Be or Not to Be ...Human? Theorizing the Role of Human-Like Competencies in Conversational Artificial Intelligence Agents. *Journal of Management Information Systems*, 11 Desember, pp. 969-1005 .

Chandra, S., Ward, P. R. & Mohammadnezhad, M., 2018. Trust and Communication in a Doctor-Patient Relationship: A Literature Review.. *Journal of Healthcare Communications*, Januar.

Chang, E., Thomson, P., Dillon, T. & Hussain, F., 2005. The Fuzzy and Dynamic Nature of Trust. I: *Trust, Privacy, and Security in Digital Business*. s.l.:s.n.

Charness , G., Gneezy , U. & Kuhn, M. A., 2012. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, Januar, pp. 1-8.

Charron, N. & Rothstein , B., 2016. Does education lead to higher generalized trust? The importance of quality of government. *International Journal of Educational Development*, September, pp. 59-73.

ChatNorskAI, 2024. *Hva er ChatGPT: En kort introduksjon til kunstig intelligens-samtaleroboten*. [Online]

Available at: <https://chatnorskai.com/hva-er-chatgpt/>

[Accessed 5 Desember 2024].

Choi, J. K. & Ji, Y. G., 2015. Views 575 CrossRef citations to date 13 Altmetric Introduction Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, 5 Oktober, pp. 692-702.

Cho, J.-H., Chan, K. & Adali, S., 2015. A Survey on Trust Modeling. *ACM Computing Surveys (CSUR)*, 12 Oktober, pp. 1-40.

Chopra, K. & Wallace, W. A., 2003. Trust in Electronic Environments. *Proceedings of the 36th Hawaii International Conference on System Sciences*, pp. 1-10.

Choudhury, A. & Shamszare, H., 2023. Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. *J Med Internet Res* , 14 Juni.

Collins, L., 2007. Research Design and Methods. *Encyclopedia of Gerontology*, p. 419-429.

Cook, K., 2001. *Trust in Society*. s.l.:Russell Sage Foundation.

Cronbach, L. J. & Meehl, P. E., 1955. CONSTRUCT VALIDITY IN PSYCHOLOGICAL TESTS. *Psychological Bulletin*, Juli, pp. 281-302.

Dahlum, S., 2023. *uavhengig variabel (forskningsmetode)*. [Internett]

Available at: https://snl.no/uavhengig_variabel_-_forskningsmetode#:~:text=En%20uavhengig%20variabel%20forutsettes%20%C3%A5,vil%20da%20v%20%C3%A6re%20uavhengig%20variabel.

[Funnet 9 Desember 2024].

Den europeiske kommisjonen, 2024. *AI Act enters into force*. [Online]

Available at: https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_en

[Accessed 17 Oktober 2024].

Denning, P., Horning, J., Parnas, D. & Weinstein, L., 2005. Wikipedia Risks. *COMMUNICATIONS OF THE ACM*, Desember, p. 152.

Denning, P. J., 2023. The Smallness of Large Language Models. pp. 24-27.

Dietvorst, B. J., Simmons, J. P. & Massey, C., 2015. Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. *Journal of Experimental Psychology: General*, p. 114 –126 .

Dietvorst, B. J., Simmons, J. P. & Massey, C., 2018. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science* , Mars, p. 1155–1170.

DigitalNorway, u.d. *Hva er og hvordan bruke ChatGPT?*. [Internett]

Available at: <https://digitalnorway.com/kurs/kunstig-intelligens-sprakmodell-og-chatbot-hva-er-chatgpt-og-hvordan-fungerer-det/>

Dimitriadis, S. & Kyrezis, N., 2010. Linking Trust to Use Intention for Technology-Enabled Bank Channels: The Role of Trusting Intentions. *Psychology and Marketing*, Juli, pp. 799-820.

Elevkanalen, u.d. *Elevkanalen*. [Internett]

Available at:

<https://www.elevkanalen.no/samling/medielabben/produkt/1947/del/15699/artikkel/118999>

[Funnet 6 November 2024].

European Commission, 2024. *European Commission*. [Internett]

Available at: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

European Parliament , 2024. *EU AI Act: first regulation on artificial intelligence*. [Internett]

Available at: <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

- Evans, A. M. & Krueger, J. I., 2009. The Psychology (and Economics) of Trust. *Social and Personality Psychology Compass* , pp. 1003-1017.
- Eysenbach, G., 2008. *Credibility of Health Information and Digital Media: New Perspectives and Implications for Youth*, Cambridge, MA. : MacArthur Foundation Series on Digital Media and Learning. .
- Fallis, D., 2008. Toward an Epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology* , Juli.
- Fallows, D., 2005. Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve. *Pew Internet & American Life Project* , 23 Januar.
- Fink, A., 2003. *How to Sample in Surveys*. 2. utgave red. s.l.:SAGE Publications, Inc. .
- Fleiss, J. L., 1999. *The Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons, Inc..
- Foddy, M., Platow, M. J. & Yamagishi, T., 2009. Group-based trust in strangers: The role of stereotypes and expectations.. *Psychological Science*, pp. 419-422.
- Freitag, M. & Bauer , P. C., 2016. Personality traits and the propensity to trust friends and strangers. *The Social Science Journal*, Desember, pp. 467-476.
- Gambetta, D., 2000. Can We Trust Trust?. *Trust: Making and Breaking Cooperative Relations*, pp. 213-237.
- Gefen, D., Karahanna, E. & Straub, D. W., 2011. Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, Vol.27, No.1, 14 Mai, pp. 51-90.
- Generative AI explained in 2 minutes*. 2023. [Film] Directed by KI-Campus. s.l.: s.n.
- Ghazizadeh, M., Lee, J. D. & Boyle, L. N., 2012. Extending the Technology Acceptance Model to assess automation. *Cognition Technology and Work* , Mars, pp. 39-49.
- Giles, J., 2005. Internet encyclopaedias go head to head. *Nature*, pp. 900-901.
- Glaeser , E., Liabson, D., Scheinkman, J. A. & Soutter, C. L., 2000. Measuring Trust. *Quarterly Journal of Economics*, Februar, pp. 811-846.
- Glikson, E. & Woolley, A. W., 2020. Human trust in artificial intelligence: Review of empirical research.. *Academy of Management Annals* , April, pp. 627-660.

Google DeepMind, u.d. *Gemini 2.0: Built for the agentic era*. [Internett]

Available at: <https://deepmind.google/technologies/gemini/>

[Funnet 14 Desember 2024].

Google, u.d. *Automatically generating and ranking results*. [Internett]

Available at: <https://www.google.com/intl/en/search/howsearchworks/how-search-works/ranking-results/>

[Funnet 8 Desember 2024].

Gosline, R. R. et al., 2024. *Nudge Users to Catch Generative AI Errors*. [Online]

Available at: <https://sloanreview.mit.edu/article/nudge-users-to-catch-generative-ai-errors/>

Grabner-Kräuter, S., 2008. Consumer acceptance of internet banking: the influence of internet trust. *Emerald Insight*, pp. 483-504.

Grant, K., 2021. *Influencers react to Norway photo edit law: 'Welcome honesty' or a 'shortcut'?*.

[Internett]

Available at: <https://www.bbc.com/news/newsbeat-57721080>

[Funnet 12 Desember 2024].

Grønmo, S. & Dahlum, S., 2024. *variabel (forskning)*. [Internett]

Available at: [https://snl.no/variabel_-](https://snl.no/variabel_-forskning#:~:text=En%20avhengig%20variabel%20kan%20p%C3%A5virkes,valg%20vil%20v%C3%A6re%20avhengig%20variabel.)

[forskning#:~:text=En%20avhengig%20variabel%20kan%20p%C3%A5virkes,valg%20vil%20v%C3%A6re%20avhengig%20variabel.](https://snl.no/variabel_-forskning#:~:text=En%20avhengig%20variabel%20kan%20p%C3%A5virkes,valg%20vil%20v%C3%A6re%20avhengig%20variabel.)

[Funnet 9 Desember 2024].

Greenwald, A. G., 1976. Within-subjects designs: To use or not to use?. *Psychological Bulletin*, p. 314–320.

Gregory, I., 2003. *Ethics in Research*. s.l.:Bloomsbury Academic.

Hamilton, L. C., 2010. Education, politics and opinions about climate change evidence for interaction effects. *Climate Change*, 14 Desember, p. 231–242.

Hancock, P. A., Billings, D. R. & Schaefer, K. E., 2011. Can You Trust Your Robot?. *Ergonomics in Design The Quarterly of Human Factors Applications*, September, pp. 24-29.

Hancock, P. A. et al., 2023. How and why humans trust: A meta-analysis and elaborated model. *Frontiers in Psychology*, 27 Mars.

Hanley, J. A., 2016. Simple and multiple linear regression: sample size considerations. *Journal of Clinical Epidemiology*, November, pp. 112-119.

- Hansen, A., 2023. *Hva gjør chatgpt unikt*. [Internett]
Available at: <https://chatgpt.no/hva-gjor-chatgpt-unikt>
- Haraldsen, F., 2021. *probabilisme*. [Online]
Available at: https://snl.no/probabilisme_-_filosofi#:~:text=Probabilisme%20er%20en%20oppfatning%20om,n%C3%B8ye%20seg%20med%20sannsynlige%20overbevisninger.
[Accessed 24 Oktober 2024].
- Hardin, R., 2006. *Trust*. s.l.:Polity Press.
- Henson , R., 2015. Analysis of Variance (ANOVA).. I: *Brain Mapping: An Encyclopedic Reference*. s.l.:Elsevier, pp. 477-481.
- Hernández-Ortega, B., 2011. The role of post-use trust in the acceptance of a technology: Drivers and consequences. *Technovation*, November, pp. 523-538.
- Hesselberg, J.-O., 2021. *Det syndes i vitenskapen*. [Internett]
Available at: <https://www.forskningsetikk.no/ressurser/magasinet/2021-2/arlig-talt/>
[Funnet 9 Desember 2024].
- Hochstein, R. E., Harmeling, C. M. & Perko, T., 2023. Toward a theory of consumer digital trust: Meta-analytic evidence of its role in the effectiveness of user-generated content. *Journal of the Academy of Marketing Science*, 30 September.
- Hoff, K. & Bashir, M., 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors The Journal of the Human Factors and Ergonomics Society*, Mai, pp. 407-434.
- Hooghe, M., Marien , S. & de Vroome , T., 2012. The cognitive basis of trust. The relation between education, cognitive ability, and generalized and political trust. *Intelligence*, Desember, pp. 604-613.
- Horrigan, J. & Rainie, L., 2006. *The Internet's Growing Role in Life's Major Moments*, s.l.: Pew Internet & American Life Project.
- Howe, D. C., 2000. Trust Online. *Communications of the ACM*, November, pp. 34-40.
- Huang, J., Shi, S., Chen, Y. & Chow, W. S., 2015. How do students trust Wikipedia? An examination across genders. *Information Technology & People*, 24 Oktober, pp. 750-773.
- Huang, Y. & Sundar, S. S., 2020. Do We Trust the Crowd? Effects of Crowdsourcing on Perceived Credibility of Online Health Information. *Health Communication* , September, pp. 93-102.

- Hurley, D., 2024. Using Early Responses to Wikipedia and Google to Consider ChatGPT. *Information Experience and Information Literacy*, 1 Februar, p. 15–26.
- Irwin, K., Edwards, K. & Tamburello, J. A., 2015. Gender, trust and cooperation in environmental social dilemmas. *Social Science Research*, Mars, pp. 328-342.
- Jaccard, J. & Turrisi, R., 1997. 1. Introduction. I: J. Jaccard & R. Turrisi, red. *Interaction Effects in Multiple Regression*. s.l.:Sage.
- Jackson, A., 2023. *Technology magazine*. [Online]
Available at: <https://technologymagazine.com/articles/chatgpt-turns-one-how-ai-chatbot-has-changed-the-tech-world>
[Accessed 22 Oktober 2024].
- Jambak, T., 2024. *KI i skolen: Mellom hybris og håp*. [Internett]
Available at: <https://www.utdanningsforbundet.no/nyheter/2024/ki-i-skolen-mellom-hybris-og-hap/>
[Funnet 14 Desember 2024].
- Jang, K. et al., 1998. Heritability of facet-level traits in a cross-cultural twin sample: support for a hierarchical model of personality. *Journal of Personality and Social Psychology*, pp. 1556-1565.
- Jean, G., 2024. Social Network Analysis for Trust: Analyzing the structure of the Wikipedia editor network to identify trusted editors and communities. 23 September.
- Johnson, J. A., 2005. Ascertaining the validity of individual protocols from Web-based personality inventories. *Journal of Research in Personality*, Februar, pp. 103-129.
- Jones, S. & Marsh, S., 1997. Human-Computer-Human Interaction: Trust in CSCW. *SIGCHI*, Jui, pp. 36-40.
- Jung, Y., Jang, E., Chen, C. & Sundar, S. S., 2024. Do We Trust ChatGPT as much as Google Search and Wikipedia?. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 11-16 Mai, pp. 1-9.
- Kapania, S. et al., 2022. "Because AI is 100% right and safe": User Attitudes and Sources of AI Authority in India. *CHI Conference on Human Factors in Computing Systems*.
- Karahanna, E., David, G. & Straub, D., 2003. Inexperience and experience with online stores: The importance of TAM and Trust. *IEEE Transactions on Engineering Management*, September, pp. 307 - 321.
- Katz, H. A. & Rotter, J. B., 1969. Interpersonal Trust Scores of College Students and Their Parents. *Child Development*, Juni, pp. 657-661 .

- Kaur, P., Stoltzfus, J. & Yellapu, V., 2018. Descriptive statistics. *International Journal of Academic Medicine*, April, pp. 60-63.
- König, R., 2012. WIKIPEDIA Between lay participation and elite knowledge representation. *Information, Communication & Society*, 29 Oktober, pp. 160-177 .
- Küper, A. & Krämer, N., 2024. Psychological Traits and Appropriate Reliance: Factors Shaping Trust in AI. *International Journal of Human-Computer Interaction*, 13 Mai, pp. 1-18.
- Kelly, W., 2024. *GPT-3.5 vs. GPT-4: Biggest differences to consider*. [Internett]
Available at: <https://www.techtarget.com/searchenterpriseai/tip/GPT-35-vs-GPT-4-Biggest-differences-to-consider>
[Funnet 10 Desember 2024].
- Kelton, K., Fleischmann, K. R. & Wallace, W. A., 2008. Trust in Digital Information. *Journal of the American Society for Information Science and Technology* , pp. 363-374.
- Khamitov, M., Rajavi, K., Huang, D.-W. & Hong, Y., 2024. Consumer Trust: Meta-Analysis of 50 Years of Empirical Research. *Journal of Consumer Research*, Juni, pp. 7-18.
- Kim, D. J., Ferrin , . D. L. & Raghav, R. H., 2008. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, Januar, pp. 544-564.
- Kim, T. K., 2015. T test as a parametric statistic. *Korean Journal of Anesthesiology*, 7 September, pp. 540-546.
- Kinapel, O. M., 2024. *Ny KI-tjeneste for generering av musikk er helt spinnvill. Hør her!*. [Internett]
Available at: <https://www.teknokratiet.no/2024/04/11/ny-ki-tjeneste-for-generering-av-musikk-er-helt-spinnvill-hor-her/>
[Funnet 14 Desember 2024].
- Knight, C. & Pryke, S., 2012. Wikipedia and the University, a case study. *Teaching in Higher Education*, November, pp. 1-11.
- Knoedler, L. et al., 2024. Pure Wisdom or Potemkin Villages? A Comparison of ChatGPT 3.5 and ChatGPT 4 on USMLE Step 3 Style Questions: Quantitative Analysis. *JMIR Med Educ* .
- Knoll, D. L. & Gill, H., 2011. Antecedents of trust in supervisors, subordinates, and peers. *Journal of Managerial Psychology*, 3 Mai, pp. 313-330.
- Knudsen, C., 2024. *Selskapet bak ChatGPT henter penger – verdsettes til 1.660 milliarder*. [Internett]
Available at: <https://e24.no/boers-og-finans/i/KMjRAX/selskapet-bak-chatgpt-henter-penger->

verdsettes-til-1660-milliarder

[Funnet 14 Desember 2024].

Knudsen, E., 2024. *Blir enda mer menneskelignende: Nå skal Chat GPT få «hukommelse»*. [Internett]

Available at: <https://www.digi.no/artikler/blir-enda-mer-menneskelignende-na-skal-chat-gpt-fa-hukommelse/543693>

[Funnet 9 Desember 2024].

Komiak, S. Y. X. & Benbasat, I., 2006. The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents. *MIS Quarterly*, Desember, pp. 941-960 .

Kopperud, I. E., 2024. *Mer personlig shopping: Slik endrer AI motebransjen*. [Internett]

Available at: https://www.minmote.no/nyheter/a/zAV851/ai-mote-shopping?utm_source=vgfront-personalized&utm_content=hovedlopet_row14_pos1&utm_term=niche_personalized_front-test-standard_bandit-2-norman_chouse

[Funnet 16 Desember 2024].

Korzyński, P., Silva, S. C. e., Górska, A. M. & Mazurek, G., 2024. Trust in AI and Top Management Support in Generative-AI Adoption. *Journal of Computer Information Systems*, 12 September, pp. 1-16.

Krithikadatta, J., 2014. Normal Distribution. *Journal of Conservative Dentistry*, Februar, pp. 96-97.

Kumar, S. & Shah, N., 2018. False Information on Web and Social Media: A Survey. *Social and Information Networks*, 23 April.

Kung, F. Y., Kwok, N. & Brown, D. J., 2017. Are Attention Check Questions a Threat to Scale Validity?. *Applied Psychology* , 24 August, p. 264–283.

Kurtz , J. E. & Parrish, C. L., 2001. Semantic Response Consistency and Protocol Validity in Structured Personality Assessment: The Case of the NEO–PI–R. *Journal of Personality Assessment* , p. 315–332.

Lakens, D., 2017. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 5 Mai, p. 355–362.

Lakens, D., Scheel, A. M. & Isager, P. M., 2018. Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, Juni, pp. 259-269.

Lankton, N. K., McKnight, D. H. & Tripp, J., 2015. Technology, Humanness, and Trust: Rethinking Trust in Technology, Humanness, and Trust: Rethinking Trust in Technology. *nal of the Association for Information Systems* , 23 Oktober, pp. 880-918.

Lapidot, Y., Kark, R. & Shamir, B., 2007. The impact of situational vulnerability on the development and erosion of followers' trust in their leader. *The Leadership Quarterly*, Februar, pp. 16-34.

Larzelere, R. E. & Huston, T. L., 1980. The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage and Family*, August, pp. 595-604.

Latif, I. & Jaffry, S. W., 2013. Trust Evaluation Mechanisms for Wikipedia. *Proceedings of the IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP)*, Oktober, pp. 36-42.

Løøv, M., 2019. *Store Norske Leksikon*. [Online]

Available at: <https://snl.no/antropomorfisme>

[Accessed 15 Oktober 2024].

Lee, J. D. & See, K. A., 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, pp. 50-80.

Lee, J. J., 2021. Party Polarization and Trust in Science: What about Democrats?. *Socius*.

Lekanger, K., 2023. *Microsoft Ignite 2023: Nå har alt blitt Copilot*. [Online]

Available at: <https://www.kode24.no/artikkel/microsoft-ignite-2023-na-har-alt-blitt-copilot/80524795>

[Accessed 24 Oktober 2024].

Leonard, S., 1960. The influence of some types of power relationships and game strategies upon the development of interpersonal trust. *The Journal of Abnormal and Social Psychology*, pp. 223-230.

Lewandowski, M., Łukowicz, P., Świetlik, D. & Barańska-Rybak, W., 2024. ChatGPT-3.5 and ChatGPT-4 dermatological knowledge level based on the Specialty Certificate Examination in Dermatology. *Clinical and Experimental Dermatology*, Juli, p. 686–691.

Li, X., Hess, T. J. & Valacich, J. S., 2008. Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, 1 Mars, pp. 39-71.

Lim, S., 2009. How and Why Do College Students Use Wikipedia?. *Journal of the American Society for Information Science and Technology*, Oktober, pp. 2189-2202.

Lim, Z. W. et al., 2023. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. *eBioMedicine*, September, pp. 1-11.

- Li, X., Hess, T. J. & Valacich, J. S., 2008. Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, Mars, pp. 39-71.
- Lo, C. K., 2023. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, April.
- Logg, J. M., Minson, J. A. & Moore, D. A., 2018. *Do People Trust Algorithms More Than Companies Realize?*. [Internett]
Available at: <https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize>
- Logg, J. M., Minson, J. A. & Moore, D. A., 2019. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, Mars, pp. 90-103.
- Longoni, C., Bonezzi, A. & Morewedge, C. K., 2019. Resistance to Medical Artificial Intelligence. *Journal of Consumer Research* , pp. 629-650.
- Lorentzen, E. A., 2024. *Vil utstyre deg med KI mot rømt oppdrettslaks*. [Internett]
Available at: <https://www.hi.no/hi/nyheter/2024/februar/vil-utstyre-deg-med-ki-mot-romt-oppdrettslaks>
[Funnet 14 Desember 2024].
- Lowe, N. K., 2019. What Is a Pilot Study?. *AWHONN, the Association of Women's Health, Obstetric and Neonatal Nurses*, Issue 48, pp. 117-118.
- Lubbad, M., 2023. *GPT-4 Parameters: Unlimited guide NLP's Game-Changer*. [Online]
Available at: <https://medium.com/@mlubbad/the-ultimate-guide-to-gpt-4-parameters-everything-you-need-to-know-about-nlps-game-changer-109b8767855a>
- Lucassen, T. & Schraagen, J. M., 2010. Trust in Wikipedia: How Users Trust Information from anUnknown Source. *Proceedings of the 4th ACM Workshop on Information Credibility on the Web*, April.
- Lund, T., Kleven, T. A., Kvernbekk, T. & Christophersen, K.-A., 2002. *Innføring i forskningsmetodologi*. 4. utgave red. Oslo: Fagbokforlaget.
- Luo, Q., Puett, M. J. & Smith, M. D., 2024. A "Perspectival" Mirror of the Elephant: Investigating Language Bias on Google, ChatGPT, YouTube, and Wikipedia. *Computers and Society*, 8 Mars.
- Luyt, B., Tay, A. C. h. & Hong, C. K., 2008. Improving Wikipedia's accuracy: Is edit age a solution?. *Journal of the American Society for Information Science and Technology* , Januar, pp. 715-722.

Macdonald, S. & Williams, C., 1994. The survival of the gatekeeper. *Research Policy*, Mars, pp. 123-132.

MacKenzie, S. I., 2024. I: S. I. MacKenzie, red. *Human-Computer Interaction: An Empirical Research Perspective*. Chapter 2: The human factor: Elsevier, pp. 31-85.

Madhavan, P. & Wiegmann, D. A., 2007. Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, 16 Mai, pp. 277-301 .

Maggio, L. A. et al., 2017. Wikipedia as a gateway to biomedical research: The relative distribution and use of citations in the English Wikipedia. *PloS one*, Desember.

Makridakis, S., 2017. The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, Juni, pp. 46-60.

Malik, E., 2023. *Artificial Intelligence (AI) and ChatGPT timelines*. [Internett]

Available at: <https://www.officetimeline.com/blog/artificial-intelligence-ai-and-chatgpt-history-and-timelines>

[Funnet 15 Desember 2024].

MarketsandMarkets, u.d. *The LLM Boom: Growth Drivers and Key Players Shaping the Future*.

[Internett]

Available at: <https://www.marketsandmarkets.com/ResearchInsight/size-and-share-of-large-language-model-llm-market.asp>

Marr, B., 2023. *A Short History Of ChatGPT: How We Got To Where We Are Today*. [Online]

Available at: <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/>

[Accessed 24 Oktober 2024].

Massey, P. A., Montgomery, C. & Zhang , A., 2023. Comparison of ChatGPT–3.5, ChatGPT-4, and Orthopaedic Resident Performance on Orthopaedic Assessment Examinations. *Journal of the American Academy of Orthopaedic Surgeons*, 1 Desember.

Mathew, A. R., Hajj, A. A. & Abri, A. A., 2011. Human-Computer Interaction (HCI): An overview. *IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, pp. 99-100.

Mayer, R. C., Davis, J. H. & Schoorman, D. F., 1995. An Integrative Model of Organizational Trust. *The Academy of Management Review*, Juli, pp. 709-734.

McBride, S. E., Rogers, W. A. & Fisk, A. D., 2011. Understanding the effect of workload on automation use for younger and older adults. *Human Factors*, Desember, pp. 672 - 686.

McCarthy, C., u.d. *What are AI virtual assistants and how can businesses use them?*. [Internett] Available at: <https://www.dialpad.com/blog/ai-virtual-assistant/> [Funnet 9 Desember 2024].

McFarland, A., 2024. *ChatGPT-modellguide: Forklaring av GPT-3.5, GPT-4, GPT-4 Turbo og GPT-5*. [Online] Available at: <https://www.techopedia.com/no/chatgpt-modellguide#:~:text=ChatGPT%2Dmodeller%20forklart%3A%20GPT%2D,%2D4%20Turbo%20%26%20GPT%2D5> [Accessed 24 Oktober 2024].

McKinsey & Company, 2024. *What is generative AI?*. [Online] Available at: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>

McKnight, D. H., 2005. Trust in Information Technology. I: G. Davis, red. *The Blackwell Encyclopedia of Management*. Malden, Massachusetts, USA: Blackwell, pp. 329-331.

McKnight, D. H., Carter, M., Thatcher, J. B. & Clay, P., 2011. Trust in a specific technology: An Investigation of its Components and Measures. *ACM Transactions on Management Information Systems*, Mai, pp. 12-32.

McKnight, D. H., Choudhury, V. & Kacmar, C., 2002. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *The Journal of Strategic Information Systems*, Desember, pp. 297-323.

Mcknight, H. D., Choudhury, V. & Kacmar, C., 2002b. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, September, pp. 334-359.

Meer, D. V., 2023. *Number of ChatGPT Users and Key Stats*. [Online] Available at: <https://www.namepepper.com/chatgpt-users> [Accessed 23 Oktober 2024].

Melteig, E., 2023. *ChatGPT: Bør vi trykke på pauseknappen?*. [Online] Available at: <https://www.forskning.no/informasjonteknologi-internett-partner/chatgpt-bor-vi-trykke-pa-pauseknappen/2203846> [Accessed 11 Desember 2024].

Merken, S., 2023. *New York lawyers sanctioned for using fake ChatGPT cases in legal brief*. [Online] Available at: <https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases->

[legal-brief-2023-06-22/](#)

[Accessed 6 Desember 2024].

Merritt, S. M. & Ilgen, D. R., 2008. Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions. *Human factors*, April, pp. 194-210.

Meticulous Research, 2024. *AI Chatbots Market to be Worth \$22.6 Billion by 2031*. [Internett] Available at: <https://www.meticulousresearch.com/pressrelease/1259/ai-chatbots-market> [Funnet 14 Desember 2024].

Metzger, M. et al., 2003. *Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment*, s.l.: Annals of the International Communication Association.

Metzger, M. J. & Flanagin, A. J., 2013. Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, Desember, pp. 210-220.

Meuter, M. L., Ostrom, A. L., Bitner, M. J. & Roundtree, R., 2003. The influence of technology anxiety on consumer use and experiences with self-service technologies. *Journal of Business Research*, November, pp. 899-906.

Meyners, M., 2012. Equivalence tests – A review. *Food Quality and Preference*, Desember, pp. 231-245.

Microsoft Azure, 2024. *Hva er dyplæring?*. [Online]

Available at: <https://azure.microsoft.com/nb-no/resources/cloud-computing-dictionary/what-is-deep-learning>

[Accessed 23 Oktober 2024].

Miller, R. G., 1997. 3 One-way classification . I: *Beyond ANOVA: Basics of Applied Statistics*. s.l.:Routledge, pp. 67-159.

Moe, H. & Bjørgan , J., u.d. *Nyheter og kunstig intelligens*. [Internett]

Available at: <https://nyhetsbruk.w.uib.no/rapport-2024/nyheter-og-kunstig-intelligens/>

[Funnet 14 Desember 2024].

Montag, C., Becker, . B. & Li , B. J., 2024. On trust in humans and trust in artificial intelligence: A study with samples from Singapore and Germany extending recent research. *Computers in Human Behavior: Artificial Humans*, Desember.

Montgomery, D. C., 2017. *Design and Analysis of Experiments*. 9 red. Hoboken, New Jersey: John Wiley & Sons.

Montgomery, . D. C., Peck, E. A. & Vining , G. G., 2012. *Introduction to Linear Regression Analysis*. John Wiley & Sons Inc: s.n.

Muir, B. M., 1994. Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, p. 1905–1922.

Mukiibi, E., 2024. *Role of the public editor in enhancing trust in legacy media: case study of daily monitor newspaper in Uganda*. , s.l.: The Aga Khan University.

Munson, J. L. & Macri , M. J., 2009. Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, Desember, pp. 424-438.

Nanda, A., Mohapatra, B. B. & Mahapatra , A. P. K., 2021. Multiple comparison test by Tukey's honestly significant difference (HSD): Do the confident level control type I error. *International Journal of Applied Mathematics and Statistic*, Januar, pp. 59-65.

NAOB, n.d. *Integritet*. [Online]

Available at: <https://naob.no/ordbok/integritet>

[Accessed 5 Desember 2024].

Nærings-og fiskeridepartementet, 2023. *Kunstig intelligens*, s.l.: Nærings- og fiskeridepartementet.

Newristics, u.d. *Saliency effect*. [Internett]

Available at: <https://newristics.com/heuristics-biases/saliency-effect>

[Funnet 9 Desember 2024].

Nicolaou, A. I. & McKnight, H. D., 2006. Perceived Information Quality in Data Exchanges: Effects on Risk, Trust, and Intention to Use. *Information Systems Research*, 1 Desember.

NTB, 2024. *Vil lære av Drammen sjukehus og innføre KI i helsesektoren*. [Internett]

Available at: <https://www.nrk.no/buskerud/helsedirektoratet-vil-innfore-ki-i-helsesektoren-1.16883747>

[Funnet 14 Desember 2024].

NTNU, u.d. *Bachelor i informatikk - Trondheim: Hva lærer du?*. [Internett]

Available at:

<https://www.ntnu.no/studier/bit/laeringsmal#:~:text=Fag%20som%20webteknologi%2C%20Informatikk,1%20matematikk,1%20konkrete%20problemer%20innen%20informatikk>

[Funnet 13 Desember 2024].

Nyst, A., 2024. *History Of ChatGPT: A Timeline Of The Meteoric Rise Of Generative AI Chatbots*.

[Internett]

Available at: <https://www.searchenginejournal.com/history-of-chatgpt-timeline/488370/>

[Funnet 16 Desember 2024].

Ognyanova, K., Lazer, D., Robertson, R. E. & Wilson, C., 2020. Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power.

Harvard Kennedy School (HKS) Misinformation Review, 2 Juni, pp. 1-19.

OpenAI, 2024. *How ChatGPT and our language models are developed*. [Online]

Available at: <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed>

[Accessed 24 Oktober 2024].

Open AI, 2024. [Online]

Available at: <https://chatgpt.com/?model=gpt-4o>

[Accessed 5 12 2024].

Orgeret, K. S. & Henrik, D., 2020. *falske nyheter*. [Internett]

Available at: https://snl.no/falske_nyheter

[Funnet 4 November 2024].

Oswald, M. E. & Grosjean, S., 2005. Confirmation bias. I: R. F. Pohl, red. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. s.l.:Psychology Press, pp. 79-97.

Pasquinelli, M., 2009. Google's PageRank Algorithm: A Diagram of the Cognitive Capitalism and the Rentier of the Common Intellect. I: K. Becker & F. Stalder, red. *Deep Search: The Politics of Search Beyond Google*. London: Transaction Publishers, pp. 1-14.

Pavlou, P. A., 2016. Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. *International Journal of Electronic Commerce* , 7 Mars, pp. 69-103.

Pelau, C., Dabija, . D.-C. & Ene, I., 2021. What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry Author links open overlay panel.

Computers in Human Behavior, September.

Pichai, S., 2023. *An important next step on our AI journey*. [Internett]

Available at: <https://blog.google/technology/ai/bard-google-ai-search-updates/>

[Funnet 14 Desember 2024].

- Piedboeuf, F. & Langlais, P., 2023. Is ChatGPT the ultimate Data Augmentation Algorithm?. *Findings of the Association for Computational Linguistics: EMNLP 2023*, Desember, p. 15606–15615.
- Plake, B. S. & Wise, S. L., 2013. Dynamics of guessing behavior: Between-group versus within-group designs. *Bulletin of the Psychonomic Society*, 13 Oktober, p. 251–253.
- Poon, J. M., 2013. Effects of benevolence, integrity, and ability on trust-in-supervisor. *Employee Relations*, pp. 396-407.
- Profilic, u.d. *Quality data. From real people. For faster breakthroughs..* [Internett]
Available at: <https://www.prolific.com/>
[Funnet 15 Oktober 2024].
- Ray, P. P., 2023. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, pp. 121-154.
- Reeves, B. & Nass, C., 2003. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. I: *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. s.l.:Center for the Study of Language and Inf; Reprint edition, pp. 18-36.
- Reichheld, F. F. & Schefter, P., 2000. *E-Loyalty: Your Secret Weapon on the Web*. [Online]
Available at: <https://hbr.org/2000/07/e-loyalty-your-secret-weapon-on-the-web>
[Accessed 9 November 2024].
- Reich, T., Kaju, A. & Maglio, S. J., 2022. How to overcome algorithm aversion: Learning from mistakes. *Journal of Consumer Psychology*, Juni, pp. 283-465.
- Ren, F. & Yanwei, B., 2019. A Review on Human-Computer Interaction and Intelligent Robots. *International Journal of Information Technology & Decision Making*, Desember.
- Rieh, S. Y. & Danielson, D. R., 2008. Chapter 7: Credibility: A Multidisciplinary Framework. *Annual Review of Information Science & Technology*, Oktober, pp. 307-364.
- Rigby, B., 2024. *A Primer on ChatGPT, LLMs, and Generative AI*. [Internett]
Available at: <https://www.callcentrehelper.com/chatgpt-llms-and-generative-ai-225423.htm>
[Funnet 23 Oktober 2024].
- Rigby, B., 2024. *Video: The Difference Between ChatGPT, LLMs, and Generative AI*. [Internett]
Available at: <https://www.callcentrehelper.com/the-difference-between-chatgpt-llms-and-generative-ai-222433.htm>
[Funnet 25 Oktober 2024].

Rihaldijiran , V., 2024. *Welch's t-Test: The Reliable Way to Compare 2 Population Means with Unequal Variances*. [Internett]

Available at: <https://towardsdatascience.com/welchs-t-test-the-reliable-way-to-compare-2-population-means-with-unequal-variances-bbf7a62967bc>

[Funnet 17 Desember 2024].

Rolfe, A. et al., 2014. Interventions for improving patients' trust in doctors and groups of doctors. *Cochrane Database of Systematic Reviews*, Mars.

Rouse , M., 2024. *Natural language processing (naturlig språkbehandling)*. [Internett]

Available at: <https://www.techopedia.com/no/ordliste/naturlig-sprakbehandling>

Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C., 1998. Introduction to Special Topic Forum: Not so Different after All: A Cross-Discipline View of Trust. *The Academy of Management Review*, Juli, pp. 393-404.

Ryseff, J., De Bruhl, B. F. & Newberry, S. J., 2024. *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*. [Internett]

Available at: https://www.rand.org/pubs/research_reports/RRA2680-1.html

[Funnet 11 Desember 2024].

Salah, M., Hussam, A., Ismail, M. M. & Abde, F., 2023. Chatting with ChatGPT: decoding the mind of Chatbot users and unveiling the intricate connections between user perception, trust and stereotype perception on self-esteem and psychological well-being. 20 Juli, p. 16.

Saunders, M. N., Lewis, P. & Thornhill, A., 2019. *Research Methods for Business Students*. 8. utgave red. s.l.:Pearson Education Limited.

Schaaff, K., Reining, C. & Schlippe, T., 2023. Exploring ChatGPT's Empathic Abilities. *IU International University of Applied Sciences*, Desember 2024, p. 5.

Schniter, E., Shields, T. & Sznycer, D., 2020. Trust in humans and robots: Economically similar but emotionally different. *Journal of Economic Psychology*, Juni, pp. 1-28.

Seligman, A. B., 1997. *The Problem of Trust*. s.l.:Princeton University Press.

Selwyn, N. & Gorard , S., 2016. Students' use of Wikipedia as an academic resource — Patterns of use and perceptions of usefulness. *The Internet and Higher Education*, Januar, pp. 28-34.

Seon, n.d. *Blackbox Machine Learning*. [Online]

Available at: [https://seon.io/resources/dictionary/blackbox-machine-](https://seon.io/resources/dictionary/blackbox-machine-learning/#:~:text=A%20blackbox%20model%20means%20no,how%20the%20decisions%20were%2)

[learning/#:~:text=A%20blackbox%20model%20means%20no,how%20the%20decisions%20were%2](https://seon.io/resources/dictionary/blackbox-machine-learning/#:~:text=A%20blackbox%20model%20means%20no,how%20the%20decisions%20were%2)

Omade.

[Accessed 24 Oktober 2024].

Shadish, W. R., Cook, T. D. & Campbell, D. T., 2002. Experiments and Generalized Causal Inference. In: *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, New York: Wadsworth Publishing Co Inc, pp. 1-32.

Siegrist, M., 2000. The Influence of Trust and Perceptions of Risks and Benefits on the Acceptance of Gene Technology. *Risk Analysis*, April, pp. 155-296.

Siegrist, M. & Cvetkovich, G., 2002. Perception of Hazards: The Role of Social Trust and Knowledge. *Society for Risk Analysis*, 23 Mai, pp. 713-719.

Simpson, T. W., 2012. What Is Trust?. *Pacific Philosophical Quarterly*, 1 Oktober, pp. 550-569.

Singla, A. et al., 2024. *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value*. [Online]

Available at: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>

Smith, A., 2014. *Older Adults and Technology Use*. [Internett]

Available at: <https://www.pewresearch.org/internet/2014/04/03/older-adults-and-technology-use/>

[Funnet 18 Desember 2024].

Smith, J.-A., 2024. *Hva er generativ kunstig intelligens (GPT)*. [Online]

Available at: <https://www.evelon.no/artikler/hva-er-generativ-kunstig-intelligens-gpt>

Staff, A., 2016. *Bias*. [Internett]

Available at: <https://www.forskningsetikk.no/ressurser/fbib/uavhengighet/bias/>

[Funnet 7 Desember 2024].

Steckler, A. & McLeroy, K. R., 2008. The Importance of External Validity. *American Journal of Public Health*, Januar, pp. 9-10.

Storøy, T. O., 2024. *Forbud og muligheter: Dette må du vite om EU AI Act*. [Internett]

Available at: <https://www.sticos.no/fagstoff/forbud-og-muligheter-dette-m%C3%A5-du-vite-om-eu-ai-act>

Strümke, I., 2023. I: *Maskiner som tenker*. Oslo: Kagge Forlag AS, p. 25.

Stryker, C. & Kavlakoglu, E., 2024. *What is artificial intelligence (AI)?*. [Online]

Available at: <https://www.ibm.com/topics/artificial-intelligence>

Su, X., Yan, X. & Tsai, C.-L., 2012. Linear regression. *WIREs Computational Statistics*, Mai, pp. 227-339.

Svare, H., Gausdal, A. H. & Möllering, G., 2019. The function of ability, benevolence, and integrity-based trust in innovation networks. *Industry and Innovation*, 20 Juni, p. 585–604.

Svartdal, F. & Grønmo, S., 2024. *eksperiment*. [Internett]

Available at: <https://snl.no/eksperiment>

[Funnet 12 Oktober 2024].

Tabachnick, B. G. & Fidell, L. S., 2007. Experimental Designs Using ANOVA. I: *Experimental Designs Using ANOVA*. s.l.:Duxbury.

Tan, H. H. & Lim, A. K., 2009. Trust in Coworkers and Trust in Organizations. *The Journal of Psychology*, p. 45–66.

Tan, H. H. & Tan, C. S., 2000. Toward the Differentiation of Trust in Supervisor and Trust in Organization. *Genetic, Social, and General Psychology Monographs*, pp. 241-260.

Tanis, M. & Postmes, T., 2005. A Social Identity Approach to Trust: Interpersonal Perception, Group Membership, and Trusting Behavior. *European Journal of Social Psychology*, Mai, pp. 413-424.

Tian, Y.-h.et al., 2017. Towards human-like and transhuman perception in AI 2.0: a review. *Frontiers of Information Technology & Electronic Engineering*, 4 Februar, p. 58–67.

Tideman, A., 2022. *nevralt nettverk*. [Online]

Available at: https://snl.no/nevralt_netverk

Tideman, A. & Elster, A. C., 2024. *maskinl ring*. [Internett]

Available at: <https://snl.no/maskinl%C3%A6ring>

Tran y, K. E., 2024. *metode*. [Internett]

Available at: <https://snl.no/metode>

[Funnet 9 Desember 2024].

Tripp, J. F., 2015. Technology, Humanness and Trust: Rethinking Trust in Technology.. *Journal of the Association for Information Systems* , Desember, pp. 880-918.

Tutella, F., 2024. *Q&A: In ChatGPT we trust?*. [Internett]

Available at: <https://www.psu.edu/news/research/story/qa-chatgpt-we-trust>

[Funnet 9 Desember 2024].

Tzur, A., 2024. *Alt du trenger   vite om spr kmodeller: enkelt forklart*. [Online]

Available at: <https://aiavisen.no/sprakmodeller/>

[Accessed 24 Oktober 2024].

- UiB, 2023. *Kunstig intelligens: Visjoner, realiteter og framtidige konsekvenser*. [Online]
Available at: <https://www.uib.no/fag/dannelsemner/122306/kunstig-intelligens-visjoner-realiteter-og-framtidige-konsekvenser-vit216>
- Uspenskyi , S., 2024. *Springs*. [Internett]
Available at: <https://springsapps.com/knowledge/large-language-model-statistics-and-numbers-2024>
- Utdanning.no, 2024. *Informasjons- og datateknologi*. [Internett]
Available at: [https://utdanning.no/utdanningsoversikt/informasjons-og-datateknologi#:~:text=Informasjons%2D%20og%20datateknologi%20\(IT\),behandle%2C%20overf%C3%B8re%20og%20manipulere%20informasjon](https://utdanning.no/utdanningsoversikt/informasjons-og-datateknologi#:~:text=Informasjons%2D%20og%20datateknologi%20(IT),behandle%2C%20overf%C3%B8re%20og%20manipulere%20informasjon).
[Funnet 13 Desember 2024].
- Van Slyke, C., Belanger, F. & Comunale, C. L., 2004. Factors influencing the adoption of web-based shopping: the impact of trust. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 14 Juni, pp. 32 - 49 .
- Vance, T., Elie-Dit-Cosaque, C. & Straub, D., 2008. Examining Trust in Information Technology Artifacts : The Effects of System Quality and Culture. *Journal of Management Information Systems*, Mars, pp. 73-100.
- Vaportzis , E., Clausen, M. G. & Gow , A. J., 2017. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. *Frontiers in psychology*, 4 Oktober.
- von Eschenbach, W. J., 2021. Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology*, p. 1607–1622 .
- Walker, E. & Nowacki , A. S., 2010. Understanding Equivalence and Noninferiority Testing. *Journal of General Internal Medicine*, 21 September , p. 192–196.
- Walsh, T., 2017. *The AI Revolution*, s.l.: NSW Department of Education.
- Wang, Y. & Redmiles, D., 2016. The Diffusion of Trust and Cooperation in Teams with Individuals' Variations on Baseline Trust. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* , p. 303–318.
- Warkentin, M., Gefen, D., Pavlou , P. A. & Rose, G. M., 2010. Encouraging Citizen Adoption of e-Government by Building Trust. *Electronic Markets*, 1 Desember, pp. 157-162 .
- Welsh, T. N., Weeks, D. J., Chua, R. & Goodman, D., 2009. Perceptual-motor interaction: Some implications for HCI. I: A. Sears & J. A. Jacko, red. *Human-Computer Interaction Fundamentals*. s.l.:CRC Press, pp. 3-17.

Wiggers, K., 2024. *ChatGPT now understands real-time video, seven months after OpenAI first demoed it*. [Internett]

Available at: <https://techcrunch.com/2024/12/12/chatgpt-now-understands-real-time-video-seven-months-after-openai-first-demoed-it/>

[Funnet 14 Desember 2024].

Xu, J., Le, K., Deitermann, A. & Montague, E., 2014. How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. *Applied Ergonomics*, November, pp. 1495-1503.

XU, M. et al., 2017. The Differences and Similarities Between Two-Sample T-Test and Paired T-Test. *Shanghai Arch Psychiatry*, Juni, p. 184–188.

Yang, D., Halfaker, A., Kraut, R. & Hovy, E., 2016. Who Did What: Editor Role Identification in Wikipedia. *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 446-455.

Yu, V. G., Lasco, G. & David, C. C., 2021. Fear, mistrust, and vaccine hesitancy: Narratives of the dengue vaccine controversy in the Philippines. *Vaccine, Volume 39, Issue 35*, 16 August, pp. 4964-4972.

Zahedi, F. M. & Song, J., 2008. Dynamics of Trust Revision: Using Health Infomediaries. *Journal of Management Information Systems*, p. 225–248.

Zerilli, J., Bhatt, U. & Weller, A., 2022. How transparency modulates trust in artificial intelligence. *Patterns*, 8 April.

Zhou, T., 2011. An empirical examination of initial trust in mobile banking. *Emerald Insights*, 12 August, pp. 527-540.

Zulic, A., Reikerås, M. & Granbo, K., 2024. *1 av 5 elever bruker ChatGPT til skolearbeid*. [Online]

Available at: <https://www.nrk.no/kultur/na-sier-1-av-5-norske-elever-at-de-bruker-chatgpt-til-skolearbeid-1.16442993>

[Accessed 6 Desember 2024].

Ødegaard, M., 2023. *Bedrifter bruker ChatGPT – hvem eier resultatet?*. [Online]

Available at: <https://www.finansavisen.no/jus/2023/10/28/8051292/bedrifter-bruker-chatgpt-hvem-eier-resultatet>

[Accessed 6 Desember 2024].

Vedlegg

Vedlegg: Spørreundersøkelsen

Dette vedlegget viser deler av spørreundersøkelsen som ble distribuert. Alle gruppene mottok spørsmålene stilt under *Introduksjon* og *Demografiske variabler*. Vi har kun valgt å inkludere én av eksperimentgruppene (ChatGPT) i vedlegget, da de andre gruppene har mottok samme spørsmål, med eneste forskjell at ChatGPT var byttet ut med «Google» eller «Wikipedia» for de respektive gruppene.

Introduksjon

Welcome! Thank you for your interest in our research!

In this survey, you will be asked to answer some questions. The survey should take around 4 minutes. You will receive £0.50.

To minimize platform glitches, please complete the task in one go and carefully read all instructions. All your responses are anonymous.

Informed Consent

Please read the following conditions carefully for participation in this survey.

- You are at least 18 years of age.
- You have a general understanding of the task.
- Your participation is completely voluntary, and you can withdraw participation at any time without any negative consequences.
- You will receive £0.50 base payment after you complete the full survey. However, if you withdraw before completion, you will not be paid.
- You have complete anonymity. All your responses will be treated anonymously and can never be traced back to your identity. Any presentation of results will be at the aggregate level. Any data shared with other researchers for transparency will also be anonymized and untraceable.

If you agree to all of the above points, then confirm informed consent by selecting "I AGREE" and continue to the next page. If you disagree with any of the points, simply navigate away from this page contribution.

I AGREE

Have you heard about ChatGPT?

Yes

No

Which version of ChatGPT are you using?

ChatGPT (ChatGPT 3.5) / free version

ChatGPT Plus (ChatGPT 4o) / paid version

Attention check

Please read before proceeding.

In surveys like this, a minority of participants do not carefully read the instructions and just rush through the survey. This can drastically reduce the quality of the collected data and compromise our research. We must therefore make sure that all our participants are reading the text before responding to each question.

To confirm that you have read this general instruction carefully, please select "Always" as your answer to the question below, about how much it rains in Nebraska.

Nebraska How often does it rain in Nebraska?

- Always
 - Most of the time
 - About half the time
 - Sometimes
 - Never
-

ChatGPT-gruppe

Intro Imagine that you need to search for information about **climate change**. Visualize yourself using ChatGPT for this purpose. As you go through the following questions, keep this scenario in mind and think about how you would approach finding information about **climate change** using ChatGPT. We will now ask you a series of questions to understand your perception of ChatGPT when using it to search for information about **climate change**.

Do you believe that climate change is caused by human activity?

- Yes
 - No
-

Please rate on a scale from strongly disagree to strongly agree: ChatGPT has the functionality I need to find information about climate change.

- Strongly disagree
- Disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Agree
- Strongly agree

ChatGPT has the features required for searching for information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT has the ability to provide me with information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

Please rate on a scale from strongly disagree to strongly agree: *ChatGPT supports me in finding information about climate change.*

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT provides competent guidance on how to gain information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT provides whatever help I need when searching for information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

Please rate on a scale from strongly disagree to strongly agree: ChatGPT is a very reliable information tool for information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT does not fail me in providing information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT is extremely dependable in giving information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

Please rate on a scale from strongly disagree to strongly agree: ChatGPT is truthful in its dealings with me when providing information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT is honest when providing information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT keeps its commitments to regulatory standards and guidelines when providing information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

Please rate on a scale from strongly disagree to strongly agree: ChatGPT is competent and effective in providing information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT performs its role as an information tool about climate change very well.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT is a capable and proficient information tool for finding information about climate change.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

Please rate on a scale from strongly disagree to strongly agree: *ChatGPT acts in my best interest.*

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT does its best to help me.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

ChatGPT is interested in my well-being, not just its own.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

How frequent do you use ChatGPT?

- Daily
 - Weekly
 - Monthly
 - Yearly
 - More rarely
 - Never
-

How long have you been using ChatGPT?

- Never used
 Less than 1 month
 1-6 months
 6 months to 2 years
 More than 2 years
-

Demografiske variabler

For each of the following information tools below, please rate how strongly you associate ChatGPT, Wikipedia and Google with technological and human qualities:

	Much more technological qualities	Primarily technological, but some human qualities	Equal balance of technological and human qualities	Primarily human, but some technological qualities	Much more human qualities
ChatGPT	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wikipedia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Google	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please rate on a scale from strongly disagree to strongly agree: My typical approach is to trust new information tools until they prove to me that I should not trust them.

- Strongly disagree
 Disagree
 Somewhat disagree
 Neither agree nor disagree
 Somewhat agree
 Agree
 Strongly agree
-

I usually trust information tools until it gives me a reason not to.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

I generally give an information tool the benefit of the doubt when I first use it.

- Strongly disagree
 - Disagree
 - Somewhat disagree
 - Neither agree nor disagree
 - Somewhat agree
 - Agree
 - Strongly agree
-

What gender do you identify with?

- Female
 - Male
 - Non-binary/ third gender
 - Prefer not to answer
-

How old are you? Write in numbers.

What is your highest level of education?

- Less than high school
 - High school
 - College
 - University
 - PhD
 - Prefer not to answer
-

Do you have an education within technology?

- Yes
 - No
-

Which political party do you identify with?



- Republican
 - Democrat
 - Neither Republican nor Democrat
 - Prefer not to say
-

Vedlegg: Advarsel

Dette var advarselen den ene av ChatGPT-gruppene ble vist før eksperimentet, og én gang i løpet av eksperimentet:

WARNING! Be aware that ChatGPT can make mistakes, as stated on ChatGPT's platform.

What can I help with?

 Message ChatGPT 

 Create image  Summarize text  Make a plan  Analyze data  More

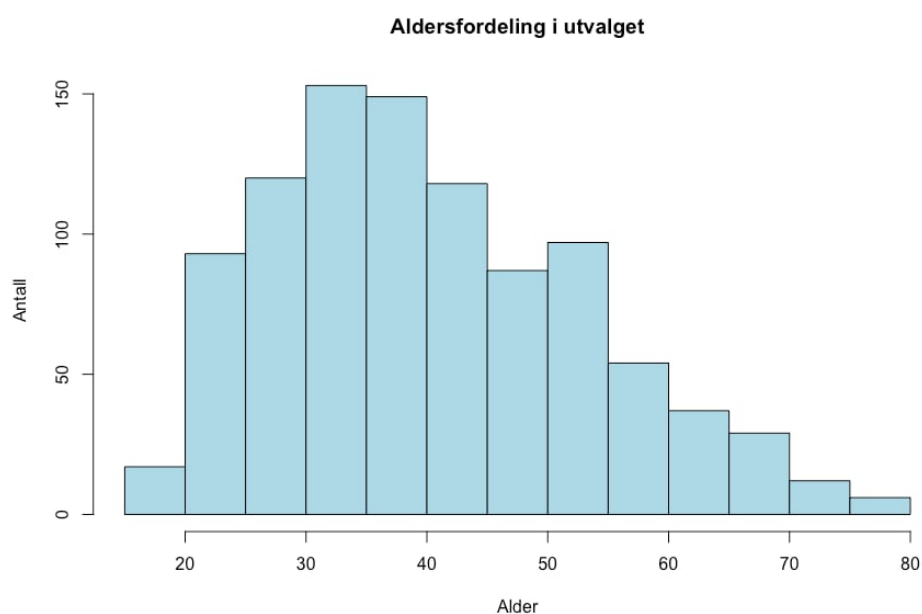
ChatGPT can make mistakes. Check important info.

Vedlegg fra R

Vedlegg 1: En oversikt over kjønnsfordeling i de ulike gruppene

Kjønn	ChatGPT	Google	Wikipedia	Advarsel
Kvinne	114	120	130	113
Mann	119	114	109	122
Ikke-binær	6	2	2	6
Ønsket ikke svare	4	5	3	3

Vedlegg 2: Histogram over aldersfordeling i utvalget



Vedlegg 3: Gjennomsnittlig generell tilbøyelighet til å stole på nye informasjonsverktøy

Gruppe	Gjennomsnitt	Median	Standardavvik
ChatGPT	4,19	4,33	1,51
Google	4,52	5,00	1,48
Wikipedia	4,28	4,67	1,51
Advarsel	4,19	4,33	1,57

Vedlegg 4: Oversikt over brukstiden respondentene har anvendt de respektive verktøyene

Gruppe	Gjennomsnitt	Median	Standardavvik
ChatGPT	3,53	4,00	0,77
Google	5,00	5,00	0,06
Wikipedia	4,89	5,00	0,50
Advarsel	3,62	4,00	0,81

Vedlegg 5: Brukshyppighet av de ulike verktøyene

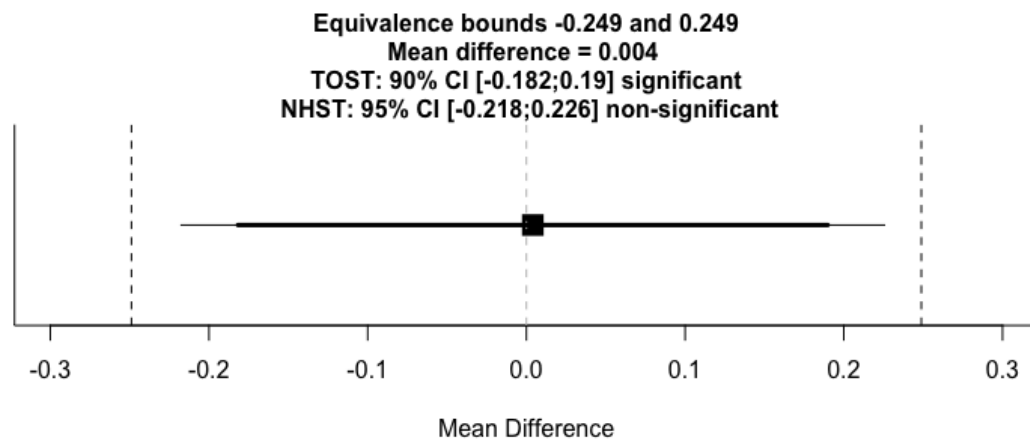
Gruppe	Gjennomsnitt	Median	Standardavvik
ChatGPT	2,45	2,00	1,22
Google	1,13	1,00	0,53
Wikipedia	2,68	2,00	1,24
Advarsel	2,53	2,00	1,34

Vedlegg 6: ANOVA-test over systemtillit mellom alle gruppene

	ANOVA				
	Df	Sum of squares	Mean Square	F	p-verdi
Group	3	29,2	9,731	8,148	0,000*
Residuals	968	1156,1	1,194		

Vedlegg 7: Tukey HSD-test over systemtillit mellom alle gruppene

TUKEY HSD-test				
Grupper	Forskjell	Nedre grense	Øvre grense	p-verdi
Google vs ChatGPT	0,2773	0,0224	0,5322	0,0267*
Advarsel vs ChatGPT	-0,1139	-0,3696	0,1418	0,6608
Wikipedia vs ChatGPT	-0,1730	-0,4279	0,0819	0,2999
Advarsel vs Google	-0,3912	-0,6466	-0,1358	0,0005*
Wikipedia vs Google	-0,4504	-0,7050	-0,1957	0,0000*
Wikipedia vs Advarsel	-0,0591	-0,3146	0,1963	0,9333

Vedlegg 8: Ekvivalens-test av oppfattet *hjelpsomhet* mellom ChatGPT og Google

Vedlegg 9: Forskjeller i systemtillit for ChatGPT og Wikipedia

Systemtillit	Gruppe	Gjennomsnitt	Forskjeller i gjennomsnitt	t	df	p-verdi
ChatGPT vs Wikipedia	ChatGPT	5,220	0,173	1,664	484,990	0,097
	Wikipedia	5,047				

Vedlegg 10: Forskjeller i systemtillit med brukstid som uavhengig variabel

```
Call:
lm(formula = System_Trust ~ use * Group, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.5016 -0.5016  0.1170  0.7061  2.2545

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.12347    0.32778  12.580 < 2e-16 ***
use              0.31103    0.09082   3.425 0.000642 ***
GroupGoogle     -3.35210    5.43942  -0.616 0.537869
GroupWarn       0.25989    0.45877   0.566 0.571188
GroupWikipedia  0.18603    0.76207   0.244 0.807194
use:GroupGoogle  0.63502    1.09050   0.582 0.560489
use:GroupWarn   -0.11117    0.12548  -0.886 0.375866
use:GroupWikipedia -0.16012    0.16687  -0.960 0.337536
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.084 on 964 degrees of freedom
Multiple R-squared:  0.04346, Adjusted R-squared:  0.03651
F-statistic: 6.257 on 7 and 964 DF, p-value: 3.421e-07
```

Vedlegg 11: Sammenhengen mellom systemtillit og brukstid for ChatGPT

```
Call:
lm(formula = System_Trust ~ chat_use, data = chat_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.2565 -0.4899  0.0769  0.6657  2.2545

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.12347    0.33999  12.128 < 2e-16 ***
chat_use         0.31103    0.09421   3.302 0.00111 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.125 on 241 degrees of freedom
Multiple R-squared:  0.04327, Adjusted R-squared:  0.0393
F-statistic: 10.9 on 1 and 241 DF, p-value: 0.001107
```

Vedlegg 12: Forskjeller i systemtillit med brukshyppighet som uavhengig variabel

```
Call:
lm(formula = System_Trust ~ freq * Group, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.5208 -0.4768  0.1164  0.6344  2.7570

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      6.06321    0.14733  41.155 < 2e-16 ***
freq             -0.34363    0.05377  -6.390 2.57e-10 ***
GroupGoogle     -0.36033    0.21409  -1.683  0.0927 .
GroupWarn       -0.20526    0.20361  -1.008  0.3137
GroupWikipedia   0.08519    0.21480   0.397  0.6918
freq:GroupGoogle  0.16160    0.13605   1.188  0.2352
freq:GroupWarn   0.04626    0.07283   0.635  0.5255
freq:GroupWikipedia -0.06653    0.07541  -0.882  0.3779
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.023 on 964 degrees of freedom
Multiple R-squared:  0.1482,    Adjusted R-squared:  0.142
F-statistic: 23.95 on 7 and 964 DF,  p-value: < 2.2e-16
```

Vedlegg 13: Sammenhengen mellom systemtillit og brukshyppighet for ChatGPT

```
Call:
lm(formula = System_Trust ~ freq, data = cleaned_data %>% filter(Group ==
"ChatGPT"))

Residuals:
    Min       1Q   Median       3Q      Max
-4.0426 -0.4922  0.1693  0.6498  1.9677

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      6.06321    0.15405  39.359 < 2e-16 ***
freq             -0.34363    0.05623  -6.112 3.94e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.07 on 241 degrees of freedom
Multiple R-squared:  0.1342,    Adjusted R-squared:  0.1306
F-statistic: 37.35 on 1 and 241 DF,  p-value: 3.94e-09
```

Vedlegg 14: Forskjeller i mennesketillit med brukstid som uavhengig variabel

```
Call:
lm(formula = Human_Trust ~ use_combined * Group, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-3.9413 -0.7658  0.1235  1.0587  2.4369

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.18490    0.36850  11.357 <2e-16 ***
use_combined      0.18911    0.10211   1.852  0.0643 .
GroupGoogle     -3.69107    6.11513  -0.604  0.5463
GroupWarn       -0.28096    0.51576  -0.545  0.5860
GroupWikipedia   0.05098    0.85674   0.060  0.9526
use_combined:GroupGoogle  0.68743    1.22596   0.561  0.5751
use_combined:GroupWarn    0.06565    0.14107   0.465  0.6418
use_combined:GroupWikipedia -0.08312    0.18760  -0.443  0.6578
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.219 on 964 degrees of freedom
Multiple R-squared:  0.01287, Adjusted R-squared:  0.005702
F-statistic: 1.795 on 7 and 964 DF, p-value: 0.08472
```

Vedlegg 15: Forskjeller i systemtillit med tilbøyelighet til å stole på nye teknologier som uavhengig variabel

```
Call:
lm(formula = System_Trust ~ dispo_combined * Group, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.0999 -0.4836  0.0671  0.6325  2.9011

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      3.45520    0.18416  18.762 < 2e-16 ***
dispo_combined    0.42150    0.04137  10.188 < 2e-16 ***
GroupGoogle      0.62814    0.27195   2.310  0.02111 *
GroupWarn        0.19296    0.25686   0.751  0.45271
GroupWikipedia   0.73595    0.26275   2.801  0.00520 **
dispo_combined:GroupGoogle -0.10842    0.05902  -1.837  0.06652 .
dispo_combined:GroupWarn   -0.07374    0.05755  -1.281  0.20035
dispo_combined:GroupWikipedia -0.22151    0.05845  -3.790  0.00016 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9726 on 964 degrees of freedom
Multiple R-squared:  0.2306, Adjusted R-squared:  0.2251
F-statistic: 41.29 on 7 and 964 DF, p-value: < 2.2e-16
```

Vedlegg 16: Forskjeller i mennesketillit med tilbøyelighet til å stole på nye teknologier som uavhengig variabel

```
Call:
lm(formula = Human_Trust ~ dispo_trust * Group, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.4394 -0.6021  0.0648  0.6943  2.8210

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      2.92628    0.20301  14.415 < 2e-16 ***
dispo_trust       0.45979    0.04561  10.081 < 2e-16 ***
GroupGoogle      0.13229    0.29979   0.441  0.65910
GroupWarn        0.14638    0.28316   0.517  0.60530
GroupWikipedia   0.74851    0.28965   2.584  0.00991 **
dispo_trust:GroupGoogle -0.05818    0.06507  -0.894  0.37144
dispo_trust:GroupWarn  -0.04176    0.06344  -0.658  0.51051
dispo_trust:GroupWikipedia -0.20770    0.06443  -3.223  0.00131 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.072 on 964 degrees of freedom
Multiple R-squared:  0.2366,    Adjusted R-squared:  0.2311
F-statistic: 42.68 on 7 and 964 DF,  p-value: < 2.2e-16
```

Vedlegg 17: Sammenhengen mellom systemtillit og oppfattet grad av menneskelignende egenskaper for ChatGPT

```
Call:
lm(formula = System_Trust ~ techvshuman_1, data = chat_data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.0052 -0.4590  0.0684  0.7726  1.9948

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.76416    0.15280  31.178 < 2e-16 ***
techvshuman_1  0.24101    0.07117   3.386 0.000828 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.124 on 241 degrees of freedom
Multiple R-squared:  0.04542,    Adjusted R-squared:  0.04145
F-statistic: 11.47 on 1 and 241 DF,  p-value: 0.0008276
```

Vedlegg 18: Sammenhengen mellom mennesketillit og oppfattet grad av menneskelignende egenskaper for ChatGPT

```
Call:
lm(formula = Human_Trust ~ techvshuman_1, data = chat_data)

Residuals:
    Min       1Q   Median       3Q      Max
-3.6184 -0.6184  0.1202  0.8588  2.3816

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.35700    0.15905  27.394 <2e-16 ***
techvshuman_1  0.26141    0.07408   3.529  5e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.17 on 241 degrees of freedom
Multiple R-squared:  0.04913, Adjusted R-squared:  0.04518
F-statistic: 12.45 on 1 and 241 DF, p-value: 0.0005002
```

Vedlegg 19: Sammenheng mellom mennesketillit og kontrollvariablene

```
Call:
lm(formula = Human_Trust ~ gender + age + education + techedu +
  party, data = cleaned_data)

Residuals:
    Min       1Q   Median       3Q      Max
-3.9938 -0.8088  0.1608  1.0751  2.3744

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.094087    0.299544  17.006 <2e-16 ***
gender        0.051605    0.066319   0.778  0.4367
age          -0.006953    0.003064  -2.269  0.0235 *
education     0.011199    0.044805   0.250  0.8027
techedu       0.016280    0.089077   0.183  0.8550
party        -0.062329    0.053598  -1.163  0.2452
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.221 on 966 degrees of freedom
Multiple R-squared:  0.007461, Adjusted R-squared:  0.002323
F-statistic: 1.452 on 5 and 966 DF, p-value: 0.203
```

Vedlegg 20: Forskjeller i systemtillit for kontrollvariablenes påvirkning for Google

```
Call:
lm(formula = System_Trust ~ gender + age + education + techedu +
    party, data = google_data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.1466 -0.5306  0.2109  0.6452  1.7637
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.687180   0.495348  11.481 <2e-16 ***
gender       -0.267443   0.123976  -2.157  0.032 *
age          -0.003043   0.005544  -0.549  0.584
education    0.052772   0.081290   0.649  0.517
techedu     -0.082187   0.157887  -0.521  0.603
party        0.153562   0.097949   1.568  0.118
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.096 on 238 degrees of freedom
Multiple R-squared:  0.02894, Adjusted R-squared:  0.008539
F-statistic: 1.419 on 5 and 238 DF, p-value: 0.2182
```

Vedlegg 21: Sammenhengen mellom systemtillit og klimasyn for ChatGPT

```
Call:
lm(formula = System_Trust ~ chat_human_act, data = chat_data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.3184 -0.5406  0.1193  0.6816  2.3416
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.2006     0.2630  23.576 < 2e-16 ***
chat_human_act -0.8822     0.2278  -3.873 0.000138 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.116 on 241 degrees of freedom
Multiple R-squared:  0.0586, Adjusted R-squared:  0.0547
F-statistic: 15 on 1 and 241 DF, p-value: 0.0001384
```