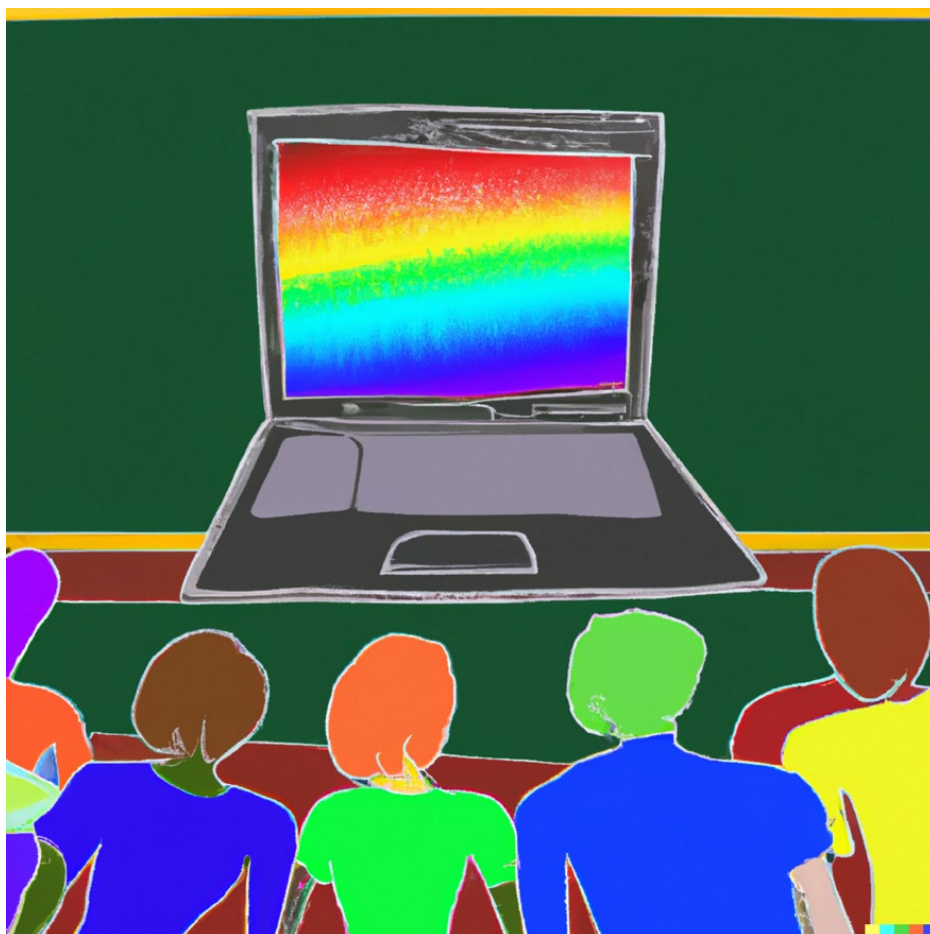


Fakultet for informasjonsteknologi og elektroteknikk

Kunstig intelligens innen utdanning

Trondheim, 18.oktober 2023

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for informasjonsteknologi og elektroteknikk



Sammendrag

Generativ kunstig intelligens (KI) og spesielt store språkmodeller, endrer måten vi lærer på, innhenter informasjon, skriver rapporter, lager modeller, skriver programkode og jobber kunstnerisk. I denne rapporten diskuterer vi hvordan generativ KI påvirker vår utdanning og hvordan fakultetet kan møte utfordringene som følger.

Kompetanse ligger til grunn for trygg og profesjonelt forsvarlig bruk av kunstig intelligens. Vi diskuterer hvilke KI-relaterte kompetanser som undervisere bør inneha. Rapporten oppfordrer fakultetet til å sørge for at faglærere og studenter får god forståelse for, og opplæring i bruk av, generative KI-verktøy. Dette vil bidra til at KI kan bli tatt i bruk i læring, undervisning og utdanning på en god måte. Vår bruk av generativ KI må være etisk, trygg, rettferdig og meningsfull, og mennesket må styre bruken av generativ KI og ikke omvendt.

Rapporten skisserer mulig bruk av generativ KI i undervisningen ved fakultetet og gir referanser til nyttige kilder om KI og bruk av KI i undervisningen. Inntoget av generativ KI utfordrer grunnprinsippet i universitet- og høyskoleloven om at studenters kunnskaper og ferdigheter skal bli prøvet på en upartisk måte dersom ikke vurderingsformene tilpasses den nye virkeligheten. Rapporten gir forslag til tolkning av NTNUs rektorvedtak fra februar 2023 om fusk og kunstig intelligens. Det er verken mulig eller hensiktsmessig å forby bruk av KI i vurderinger uten tilsyn. Rapporten skisserer hvordan bruk av KI-verktøy kan tillates og deklarerer i vurderinger uten tilsyn.

Rapporten skisserer en plan for veien videre for fakultetet og for NTNU.

INNHOLDSFORTEGNELSE

Sammendrag	2
Oppdraget	5
Mandat for arbeidsgruppen	5
Arbeidsgruppens sammensetning	5
Innledning	6
Arbeidet i gruppen	6
Rektorvedtak om bruk av kunstig intelligens	6
Avgrensning av arbeidet	6
Avgrensning av begrepet kunstig intelligens	7
Situasjonsbeskrivelse	7
Hva er nytt nå?	7
Kan undervisere og studenter trygt bruke generativ KI?	8
Kompetanse	9
Forskjell mellom kunstig og menneskelig intelligens	9
Hvordan KI kan bidra til læring.....	9
Hvordan KI kan være til hinder for læring	10
Hvordan formulere spørsmål og oppgaver for KI-systemer	10
KI-anvendelser innenfor eget fagområde.....	11
Juridiske forhold og merking av resultater fra KI-verktøy	11
Kollektive tiltak til kompetanseheving.....	12
Undervisning	12
Generativ KI som en læringsressurs.....	12
Generativ KI, læringsmål og støttende læringsaktiviteter	13
Bruk av generativ KI for å støtte underviseren	13
Vurdering	14
Vurderingsformer uten tilsyn	15
Master- og bacheloroppgaver	16
Skriftlig eksamen under tilsyn uten hjelpemidler	17
Skriftlig eksamen under tilsyn med begrensning i hjelpemidler	17
Skriftlig eksamen under tilsyn med alle hjelpemidler	17

Muntlig og praktisk eksamen	17
Arbeidsgruppens anbefalinger.....	17
Plan for veien videre ved IE og NTNU	20
Vedlegg A: Læringsressurser for å tilegne seg kunnskap innen KI og generativ KI.....	21
Elements of AI.....	21
Maskiner som tenker.....	21
Practical AI for instructors and students	21
Introduction to Generative AI & ChatGPT	21
Generative AI exists because of the transformer.....	22
Digital Norway nettkurs.....	22
Vedlegg B: Deklarasjon om bruk av KI-hjelpemidler	23

Figuren på forsiden er laget av DALL-E med promptet "digital art of a slim silver laptop as the teacher in an auditorium with a green chalkboard and students with rainbow coloured T-shirts sitting looking ahead"

Oppdraget

Fakultet for informasjonsteknologi og elektroteknikk ved NTNU (IE) etablerte våren 2023 en arbeidsgruppe for å vurdere hvilke konsekvenser den raske utviklingen innen kunstig intelligens vil ha innen fakultetets utdanningsvirksomhet. Gruppen skulle foreslå tiltak på kort og mellomlang sikt (1-2 år) som vil være hensiktsmessige for å utnytte de nye mulighetene og unngå at teknologiendringene får uheldige konsekvenser for IEs utdanningsvirksomhet.

Mandat for arbeidsgruppen

Tankekartet under viser viktige aspekter som gruppen kunne vurdere. Med utgangspunkt i dette skulle arbeidsgruppen vurdere om og eventuelt hvordan AI vil påvirke de ulike aspektene og hvilke aktiviteter eller tiltak som bør iverksettes. Gruppen stod fritt til å utforme og avgrense arbeidet.



Arbeidsgruppens sammensetning

- Rune Volden, Institutt for IKT og realfag, leder
- Rune Hjelsvold, Institutt for datateknologi og informatikk
- Mette Langaas, Institutt for matematiske fag
- Morten Dinhoff Pedersen, Institutt for teknisk kybernetikk
- Torbjørn Svendsen, Institutt for elektroniske systemer
- Gilbert Bergna-Diaz, Institutt for elektrisk energi
- Terje Brekke, Utdanningsseksjonen IE
- Vegard Rønning, Utdanningsseksjonen IE

Innledning

Oppdragsbeskrivelsen bruker AI som forkortelse for kunstig intelligens – noe som har lang tradisjon innen akademien. Arbeidsgruppen ser nå at forkortelsen KI holder på å etablere seg i Norge og har valgt å bruke denne forkortelsen i denne rapporten.

Arbeidet i gruppen

Arbeidsgruppen har i 2023 hatt ett møte før sommerferien og deretter hatt ukentlige en-times møter på Teams, hvor vi har bidratt med felles erfaringer og forståelse rundt utfordringene. Mellom møtene har gruppedeltakerne lest og vurdert hva andre har gjort på området og tatt med innspill til møtene. Ett av møtene hadde vi også med fagpersoner fra NTNUs fellesadministrasjon. Det var svært nyttig for oss. Problemstillingen er omfattende, og gruppen bestemte seg tidlig for å avgrense arbeidet for å holde tidsfristen. Det har vært en kombinasjon av diskusjoner og avklaringer i arbeidsgruppen og individuelt arbeid av den enkelte. Dette har gitt konklusjoner på kortsiktige tiltak, med forslag om videre arbeid fremover. Rapporten vår gir en rekke begrunnede anbefalinger på hvordan vi skal håndtere generativ KI innen utdanningsområdet ved IE fakultetet.

Rektorvedtak om bruk av kunstig intelligens

I februar 2023 vedtok Rektor en oppdatering av NTNUs regelverk for fusk¹. Vi kan lese at² «Det er juks å generere besvarelse ved hjelp av kunstig intelligens (Chat GPT eller lignende) og levere den helt eller delvis som egen besvarelse.» med tillegget at «Dette gjelder både obligatoriske innleveringer og eksamen». Det eksisterer en usikkerhet i tolkningen av vedtaket, som vi kommer tilbake til under avsnittet om vurdering.

Avgrensning av arbeidet

Gruppen har valgt å ikke diskutere etter- og videre utdanning og ikke infrastruktur, og vi diskuterer i mindre grad ulike KI-verktøy og serviceundervisning.

Etter- og videreutdanning (EVU) er en viktig del av NTNUs samfunnsoppdrag, men gruppen har bevisst valgt å se bort ifra dette nå. Vi vet at NAIL (NTNU Norwegian Open AI Lab) arbeider med dette på vegne av fakultetet. Verktøy/infrastruktur har gruppen i mindre grad kommentert, fordi dette bør diskuteres i NTNUs infrastrukturutvalg og driftes av NTNU IT. Serviceundervisning har vi i mindre grad kommentert fordi vi først har behov for å få en god oversikt over eksisterende emner og initiativer. I tillegg ser vi i stor grad en overlapp mot læringsressurser for opplæring av ansatte. Hvilken rolle kan NTNUs Læringsstøttesenter ha rundt opplæring i bruk av konkrete KI-verktøy?

Gruppen har valgt å konsentrere arbeidet mot tre hovedområder: kompetanse, undervisning og vurdering.

Kompetanse ligger til grunn for trygg bruk av KI. Dette gjelder både for undervisere og studenter. Undervisningen bør vise til konstruktiv bruk av generativ KI og gi studentene en grunnleggende forståelse av god KI praksis. Datasikkerhet og personvern er vesentlig i denne sammenheng. Undervisningen bør også gi studentene en forståelse for KIs begrensninger og mulige svakheter. Vurderingsformene må sikre at studentene bruker KI-verktøyene fornuftig. Besvarelsene må

1

2 <https://i.ntnu.no/wiki/-/wiki/Norsk/Juks+på+eksamen>

inneholde kritisk tenkning, gode refleksjoner, kildehenvisninger og struktur som viser ansvarlig bruk av verktøy.

Avgrensning av begrepet kunstig intelligens

Kunstig intelligens kan, som Inga Strømke sier³, ses på som en samlebetegnelse for ulike teknikker og teknologier som er utviklet for å løse helt bestemte oppgaver. Det er spesielt underområdet generativ KI som har hatt størst utvikling de siste årene, en utvikling som har gitt støtet til den pågående diskusjonen om hvordan KI vil utvikle seg, og hvilke konsekvenser denne utviklingen vil kunne ha. Generative KI-systemer lærer fra eksempler, massive mengder med data, og kjennetegnes ved at de kan generere innhold basert på brukerininput, en brukerinstruks som på engelsk kalles *prompt*. Generative KI-systemer kan generere tekst (for eksempel ChatGPT), bilder (for eksempel Dall-E), musikk (for eksempel Jukebox) og tale (for eksempel Voicemod). Den typen generativ KI som har størst påvirkning for vår utdanningsvirksomhet er uten tvil de tekst-genererende KI-systemene. Slike systemer omtales også ofte som store språkmodeller (*Large Language Models*, LLM, på engelsk). Vi vil i det følgende fokusere på tekst-genererende KI, og for enkelhets skyld benytte begrepet generativ KI om slike systemer når ikke annet er spesifisert.

Situasjonsbeskrivelse

Rektor sa i sin immatrikuleringstale i august 2023⁴: "Sett dere inn i temaet - kunstig intelligens påvirker en rekke fagfelt og endrer måter å jobbe på. Jeg vet at mange av våre undervisere nå planlegger hvordan problemstillinger knyttet til KI kan tas inn i undervisningen."

Studentene våre forventer at KI vil være en del av undervisningen vår, og samfunnet forventer at studenter som uteksamineres fra NTNU har en god forståelse for styrker og svakheter ved generativ KI og er i stand til å bruke KI-baserte verktøy i sitt fremtidige profesjonelle virke på en god måte.

Hva er nytt nå?

Generativ KI og spesielt store språkmodeller, endrer måten vi lærer på, innhenter informasjon, skriver rapporter, lager modeller, skriver programkode og jobber kunstnerisk. Verktøyene bruker den informasjonen som brukerne legger inn, til videre trening av de generative KI-modellene. Dette betyr blant annet at personlig og potensielt sensitiv informasjon kan komme på avveie dersom slik informasjon finnes i instruksjonene som brukeren legger inn. Språkmodellene reflekterer i stor grad de dataene de er trent på. Mange av dagens store språkmodeller er trent på data fra åpne kilder på Internett og vil derfor speile holdninger og kunnskap fra det åpne internettet.

De store språkmodellene er laget for å generere tekst ved å predikere hva neste ord i en setning kan være, basert på instruksjoner fra brukerne sammen med tidligere ord som er generert. Begrepet hallusinerer brukes om tilfeller hvor de store språkmodellene skriver setninger der innholdet ikke er basert på fakta. Generativ KI forstår verken instruksjonene brukerne gir eller responsen, og vil i mange tilfeller generere et svar basert på det mest sannsynlige (i trainingsdatasettet), noe som i mange tilfeller ikke vil reflektere mangfold. Det ligger ikke en kunnskapsmodell bak en stor språkmodell. Det er mange misforståelser rundt hva språkmodeller er, og hva de ikke er. Noen slike misforståelser er at man tror at modellen inneholder alle treningsdataene (i form av en database), eller at de fungerer som et avansert google-søk.

³ <https://digitalnorway.com/hva-ma-du-vite-om-kunstig-intelligens-i-2023-vi-spurte-en-ki-forsker/>

⁴ <https://www.universitetsavisa.no/anne-borg-immatrikulering-ntnu/rektors-tale-vaer-en-god-medstudent/384262>

Kan undervisere og studenter trygt bruke generativ KI?

Informasjon kan klassifiseres i fire kategorier: åpen (grønn), intern (gul), fortrolig (rød) og strengt fortrolig (sort)⁵. Informasjon som er tilgjengelig til alle uten «særskilte tilgangsrettigheter» er i grønn kategori, mens informasjon i gul kategori er begrenset med kontrollerte tilgangsrettigheter og kalles intern. I en universitetssetting vil eksamensoppgaver og forskningsartikler under arbeid være eksempler på gule data⁶.

Bruken av generativ KI bør gjøres på en «personverntrygg og profesjonelt forsvarlig måte». Det er mange selskaper som tilbyr systemer som gir tilgang til generativ KI og språkmodeller, den mest kjente er ChatGPT fra OpenAI. Hvis man skal bruke ChatGPT direkte må man lage en brukerkonto. Når man da bruker tjenesten, vil man utlevere persondata gjennom innloggingen. Videre vil dialogen man har med chatboten kunne lagres (de instruksjoner du gir, tekst og svar). Denne informasjonen kan brukes til å forbedre tjenesten, det vil si å hjelpe den å generere bedre svar, men det er vanskelig å holde oversikt over egne data og vite om disse data også sendes videre til andre tjenester. For profesjonelt forsvarlig bruk bør man være sikker på at informasjonen som legges inn i samtaleroboten ikke brukes videre. Hvis du for eksempel skriver din forskningsidé inn i samtaleroboten så må du anta at du har spredt denne idéen videre. *Denne tjeneste kan derfor ikke brukes på en personverntrygg og profesjonelt forsvarlig måte for studenter eller ansatte ved NTNU.* Men, enhver myndig person kan selvfølgelig lage bruker og/eller betale for en generativ KI-tjeneste til privat bruk.

NTNU åpnet 22.09.2023 for bruk av store språkmodeller for ansatte gjennom Bing Chat Enterprise⁷. Bruken er for informasjon klassifisert som grønn eller gul. Bing Chat Enterprise bruker GPT-4 fra OpenAI, og bruk gjennom Bing Chat Enterprise er sett på som trygg (instruksjonene du gir og svarene du får, blir slettet når du logger ut av tjenesten). Chatboten i Bing har tilgang til internett via et såkalt Bing-søk.

Når vi skriver denne rapporten, har ikke *studenter ved NTNU* fått personverntrygg og profesjonelt forsvarlig tilgang til generative KI-systemer, men vi forventer at dette skjer snart (studenter ved UiO har allerede personverntrygg tilgang, se under).

Det betyr at undervisere ved NTNU nå på en personverntrygg og profesjonelt forsvarlig måte kan bruke og vise frem bruken av generativ KI, som store språkmodeller, i sin undervisning, men *vi kan pr. oktober 2023 ikke kreve at studenter gjør egne erfaringer i bruk av store språkmodeller i et emne.*

Andre institusjoner har valgt andre løsninger enn Bing Chat Enterprise. Fra september 2023 fikk UiO-ansatte og studenter tilgang til ChatGPT 3.5 fra OpenAI gjennom tjenesten gpt.uio.no, der data behandles personverntrygt på «europisk jord»⁸. Som for løsningen ved NTNU, legges det vekt på at tjenesten kun skal brukes til såkalte grønne og gule data.

Vi er trygge på at Rektors stab følger godt med på de formelle juridiske reglene innen dette feltet, og informasjonen gitt over ser vi for oss har en kort holdbarhet.

⁵ <https://i.ntnu.no/wiki/-/wiki/Norsk/Lagringsguide>

⁶ <https://i.ntnu.no/wiki/-/wiki/Norsk/Klassifisering+av+filer+og+informasjon>

⁷ <https://i.ntnu.no/wiki/-/wiki/Norsk/Bing+Chat+Enterprise>

⁸ <https://www.uio.no/tjenester/it/aktuelt/om-it/2023/uio-gpt.html>

Kompetanse

Underviserne trenger å tilegne seg høy KI-kompetanse - spesielt knyttet til generativ KI - for både å forstå hva nye KI-systemer kan bety for valg og gjennomføring av undervisningsaktiviteter, for valg og bruk av ulike vurderingsformer og for å forstå hvordan de kan bidra til at studentene får utviklet god KI-kompetanse.

Hvilke kompetanser bør underviserne og studentene ha? På (mellom-) lang sikt bør vi *kartlegge* den eksisterende kompetansen blant faglærerne og studenter og tilby opplæring i henhold til behov i hvert enkelt tilfelle. Som en start viser vi til vårt annoterte Vedlegg A med anbefalte opplæringstilbud ved rapportens tilblivelse, og peker spesielt på micro-kurset ved universitetet i Aalborg, som vi tror vil være på et passende nivå for mange studenter og ansatte.

Forskjell mellom kunstig og menneskelig intelligens

Underviserne bør forstå de grunnleggende trekk i hvordan KI-systemer, som baserer seg på maskinlæring, fungerer og hvilke oppgaver slike systemer er gode på å løse. Det er spesielt viktig at underviserne forstå hvilken viktig rolle treningsdata spiller i slike systemer og hvilke utfordringer det kan være med tanke på skjevheter, forutinntatthet og partiskhet.

Skjevheter (bias) i generativ KI er det viktig å forstå. Hvis vi ønsker at en språkmodell skal gjenspeile vår verden, kan vi kalle det representasjonsskjevhet hvis treningsdataene ikke er representative for «vår verden» (i statistikk snakker vi om populasjon og utvalg, og da vil trolig ikke treningsdataene for en stor språkmodell være et tilfeldig utvalg fra populasjonen vi ønsker å forholde oss til). Hvis treningsdata i hovedsak er hentet fra det åpne internettet og ser det et utvalg fra en spesiell populasjon, og *det er denne populasjonen som svarene fra en språkmodell gjenspeiler*. I tillegg kan språkmodeller i mange tilfeller bekrefte informasjonene som du har skrevet i promptet, selv om denne ikke er korrekt (mange har overbevist ChatGPT om at 2+2 ikke er 4), og dette omtales som bekreftelsesskjevhet.

Slike KI-systemer kan betraktes som en «stokastisk papegøye»⁹ som bare gjentar språklige konstruksjoner i treningsdata uten å ha noen forståelse av hva disse konstruksjonene betyr. (Det er ingen underliggende kunnskapsmodell.) Undervisere bør også forstå at en utfordring med eksisterende systemer som tar i bruk maskinlæring, er at disse systemene (i hovedsak) ikke er i stand til detaljert å forklare *hvorfor* de kommer fram til sine konklusjoner og anbefalinger. Dette er fokus i fagfeltet forklarbar KI (explainable AI, XAI).

Ansatte ved biblioteket ved NTNU forteller om henvendelser fra ansatte og studenter der de ønsker tilgang til kilder de selv ikke klarer å finne frem til, der det viser seg at KI-verktøy har funnet opp kilder som ikke eksisterer.

Hvordan KI kan bidra til læring

KI-systemer basert på store språkmodeller kan ha positive bidrag til studenters læring - også når det gjelder hvordan studentene kan gjøre bruk av slike verktøy på en god, sunn og kritisk måte i sitt profesjonelle virke. Undervisere bør være kjent med hvordan KI-systemer kan bidra til studenters læring. Dette inkluderer forhold som: Slike systemer kan hjelpe studenter til å finne og sammenstille informasjon på en raskere måte enn ved bruke av klassiske søkesystemer. KI-systemene kan hjelpe studentene å finne alternative løsningsveier for en utfordring de har. Slike systemer kan også gi studentene tilbakemelding på noe de har produsert eller formulert og studentene kan raskt selv få generert spørsmål som kan teste deres faglige forståelse. Videre kan slike systemer brukes til å ha

⁹ <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

sterkere fokus på refleksjon og forståelse enn memorering, for eksempel ved at studentene må bedømme, vurdere eller sammenlikne noe KI-systemene har produsert. Systemene kan også settes opp til å delta i dialoger som fremmer læring og forståelse for fagstoffet. Dessuten kan KI-systemer bidra til å pirre studentenes faglige nysgjerrighet og engasjement. KI-verktøyene kan bidra til at studentene lettere kommer i gang med å lære av gode eksempler, og de kan få generert ulike forklaringer og analogier som kan hjelpe dem å forstå vanskelige konsepter. Dette kan være eksempelkode, startfigurer i et tegneverktøy, osv.

For å unngå at studentene bruker KI på en måte som er til hinder for læring, *må fokuset i stor grad flyttes fra å finne riktige svar på oppgaver til selve læringen*. KI kan for eksempel brukes til å oppfordre til å prøve å feile framfor å «straffe» for å få feil svar. Det kan bidra til økt studentkreativitet å vite at det å feile er «trygt» og til at studentene blir oppmuntret til å bli modigere i sine resonnementer.

Hvordan KI kan være til hinder for læring

Underviserne bør også være klar over at KI-systemer basert på store språkmodeller kan være til hinder for studentenes læring. En stor utfordring med slike KI-systemer er at det kan bli fristende for studentene å ta snarveier. Det kan bli søkelys på å finne raske svar og løsninger framfor å fremme læring og forståelse. Det er også sjanser for at KI-systemene presenterer falske – og kanskje til og med skadelige - fakta og resonnementer på en overbevisende måte - som kan føre til at studentene sitter igjen med feil faktakunnskap (hallusinasjoner) og/eller forståelse.

I en rapport fra ansatte ved Utdanningsvitenskapelig fakultet fra UiO¹⁰ ble det utført forsøk med bruk av ChatGPT blant studenter. Men fant blant annet at studentene synes svarene fra språkmodellen ChatGPT var retorisk overbevisende, og at dette var spesielt problematisk der studenten hadde manglende forkunnskap. Studentene mente videre at det er institusjonens ansvar å skape klare rammer for bruk av ulike typer KI-verktøy i studiearbeidet.

Man bør også være klar over at tilgjengelighet av KI-verktøy kan føre til store forskjeller mellom studenter som klarer å utnytte KI til bedre læring og de som ikke kan. Det er også fare for at KI-baserte dialoger resulterer i at studenter har mindre kontakt med undervisere og medstudenter – noe som kan virke negativt på studenters mentale helse og for universitetets mulighet til å bidra til utvikling av studenters dannelse. Det at slike KI-systemer er generative, kan også føre til at studentene får færre utfordringer og muligheter til å utvikle egen kreativitet og originalitet og til selvstendig å utvikle egne formuleringer. Egne formuleringer kan ofte være et viktig element i læringsprosessen, for å formidle godt til andre, er det nødvendig å forstå.

Hvordan formulere spørsmål og oppgaver for KI-systemer

Det har vist seg at kvaliteten til KI-systemer basert på store språkmodeller i stor grad påvirkes av de spørsmål som stilles. Det å utforme effektive instruksjoner – *prompt engineering* på engelsk – er en ferdighet i seg selv som undervisere bør lære seg og få god erfaring med. For å gjøre det må man forstå hvordan instruksene man gir påvirker responsen fra språkmodellen. Man må spesifisere klart hvilken bakgrunnsinformasjon som skal legges til grunn, bruke presist språk og angi hvilken målgruppe svaret skal være for. Videre er det viktig at ikke gir fra seg sensitiv informasjon. Se for eksempel nettkurset til Digital Norway om «Prompt Engineering» lenket til i Vedlegg A. Testing og validering av svaret fra en språkmodell må deretter gjøres, basert på andre kilder enn språkmodellen. Ett av rådene for prompt engineering er å skrive instruksene på engelsk og ikke på norsk, fordi

¹⁰ <https://www.uv.uio.no/om/organisasjon/idea/rapporter/lett-a-bruke-vanskelig-a-forsta-ludvigsen-morch---wagstaffe-29-aug-2023.pdf>

treningsdataene som språkmodellen er trent på ikke inneholder mange norske tekster. Vi ser frem til at arbeidet med NorGPT¹¹ sluttføres, både med hensyn på at vi da kan ha en modell trent på norske tekster og at finjustering og trygg innretning er tilpasset bruk i Norge.

KI-anvendelser innenfor eget fagområde

KI-systemer basert på store språkmodeller blir stadig oftere betraktet som fundamentale modeller – *foundation models* på engelsk – siden de er trent opp på et generelt datagrunnlag som så kan danne fundament for mer spesifikke modeller innen snevrere områder. Inntil slike mer spesifikke modeller er utviklet innen våre fagområder, bør derfor undervisere gjøre seg kjent med den enkelte språkmodellens styrker og svakheter når den brukes innenfor det spesifikke fagområdet. Dette er også noe man kan utvikle ved å skaffe seg praktisk egen erfaring i bruk av en aktuell stor språkmodell.

Juridiske forhold og merking av resultater fra KI-verktøy

Det er spesielt to juridiske forhold undervisere må forstå. Som diskutert tidligere, er det utfordringer knyttet til personvern og datasikkerhet som underviserne må være klar over. I tillegg bør undervisere være klar over de etiske utfordringene som ligger i at både tekstlige og visuelle generative KI-systemer ofte er trent opp på rettighetsbeskyttet materiale.

Det er enighet om at tekst, kode eller andre resultater som er laget ved hjelp av generative KI-verktøy skal – som andre kilder – merkes etter «gjeldende regler». Men, det er ikke enighet om hva man skal forstå med «gjeldende regler». Noen institusjoner, som University College London¹², foreslår ulike løsninger, inkludert en minimumsløsning som innebærer at man i forord eller innledning gir navn og versjon av KI-verktøyet som er brukt, navn på forlegger eller leverandør, nettsadresse og en kort beskrivelse av bruk. Eksempel: «Jeg har brukt ChatGPT 3.5 (OpenAI, <https://chat.openai.com>) til å lage et sammendrag av mine notater og til å korrekturlese det endelige utkastet». Det kan også være behov for mer informasjon, avhengig av om faglærer, institutt, eller institusjon (som NTNU) har bestemt at mer informasjon (og kanskje også direkte sitater fra bruken av KI-verktøyet) må gis. Da kan det være relevant å inkludere hvilke(n) brukerinstruks(er) (prompt) som ble brukt, datoen «samtalen» fant sted, hva som var svar fra det generative KI-verktøyet (for eksempel ved en lenke til hele samtalen, eller ved å gjengi samtalen i et vedlegg) og til slutt hvordan svaret fra KI-verktøyet er blitt bearbeidet og inkludert i arbeidet som presenteres. Et eksempel på en slik praksis kan man finne i artikkelen «What Should Data Science Education do with Large Language Models»¹³, der det er inkludert lenke til samtalen rundt analyse av et datasett¹⁴.

Det er videre en diskusjon rundt hvordan man formelt skal sitere med korrekt bruk av formelle siteringsstiler, som for eksempel APA-stil¹⁵, til bruk og resultater fra KI-verktøy. For APA-stil foreslår «the APA Style team» at man setter forlegger eller leverandør (som OpenAI) som forfatter (siden selve språkmodellen ikke kan sees som en forfatter og ta ansvar for hva som blir generert), og som dato bruker man året som versjonen KI-verktøyet man bruker er utgitt. Videre vil tittel for referansen være navnet på den generative KI-modellen, for eksempel ChatGPT, eventuelt med versjonsnummer som da kan tas med i parentes etter tittelen. Det foreslås også at en beskrivelse av det generative KI-verktøyet (som for eksempel «stor språkmodell») og nettsadresse kan tas med.

¹¹ <https://www.universitetsavisa.no/arendalsuka-chatgpt-jon-atle-gulla/norgpt-er-en-modell-som-enna-ikke-har-laert-seg-folkeskikk/384624>

¹² <https://library-guides.ucl.ac.uk/referencing-plagiarism/acknowledging-AI>

¹³ <https://arxiv.org/abs/2307.02792>

¹⁴ <https://chat.openai.com/share/daecaa84-a8d7-47cb-bc2c-93b0c36eaa1a>

¹⁵ <https://apastyle.apa.org/blog/how-to-cite-chatgpt>

Kollektive tiltak til kompetanseheving

Det virker klart at det bør skje en rask kompetanseheving blant mange undervisere for å møte de utfordringene de nye KI-systemene har gitt oss. De enkelte fagmiljø bør vurdere hvilke kollektive initiativ som kan tas for en samordnet kompetanseutvikling innad i fagmiljøet. Det er flere fordeler ved å legge opp til kollektive initiativ:

- Det kan gjøre at flere undervisere blir med på kompetansehevende aktiviteter enn dersom dette overlates helt og holdent til den enkelte.
- Kollektiv aktivitet kan gi bedre koordinering på tvers av emner (innen og på tvers av studieprogrammer). Dette vil komme studentene til gode, for eksempel ved at undervisere kan avtale en «arbeidsdeling» på tema og grunnleggende «KI-ferdigheter» og kan samkjøre innhold i emnene, bruk av KI-systemer og forventninger til studentene.
- Kollektiv aktivitet vil ofte resultere i erfaringsdeling - både om bruk av KI-systemer i undervisning generelt og om bruk av KI-systemer innen fagområdet spesielt.

Undervisning

UNESCO-rapporten fra september 2023 om «Guidance for generative AI in education and research»¹⁶ har fokus på at vår bruk av KI skal være menneskesentrert. Med det mener man at bruken av generativ KI skal være *etisk, trygg, rettferdig og meningsfull, og at det er mennesket som skal styre bruken av generativ KI og ikke omvendt.*

Det listes opp en rekke bekymringer knyttet til bruken av generativ KI innen utdanning: Vil bruk av generativ KI i undervisningen redusere interaksjonen mellom mennesker og påvirke de sosiale og emosjonelle effektene av læring? I hvilken grad vil generative KI-systemer som etteraper menneskelig interaksjon påvirke den kognitive utviklingen hos studenter? Hvilke skjevheter fra nye treningsdata (fra utdanning) vil forekomme etter hvert som generativ KI brukes innen utdanning?

Videre er det åpne og uløste problemer knyttet til rettigheter til treningsdata og til det som produseres av generativ KI (tekst, bilde, programkode, video, osv.). Det er pr nå ikke mulig å oppdage at det som er produsert (og kanskje levert inn som del av en evaluering) er generert av KI.

Generativ KI som en læringsressurs

En mulig bruk av generativ KI i undervisningen er som hjelp til å tilrettelegge for læring for studenter med utfordringer. Det er rapportert (anekdotisk) god nytte av store språkmodeller til å forbedre tekst for personer som lider av dysleksi¹⁷, men ifølge UNESCO-rapporten har kun fire land anbefalt bruk av KI for studenter med utfordringer. Generative KI-verktøy kan også være multimodale (ikke bare ta tekst som input, men også stemme og bilder). Personer med synsutfordringer vil kunne laste opp bilder til et generativt KI-verktøy og stille spørsmål om bildet. Dette kan være nyttig hvis man i undervisningen for eksempel bruker diagrammer eller grafer (som gjøres i mange IE-emner).

En annen mulig bruk av generativ KI er som en personlig læringsassistent for studenter, som en slags digital tvilling til faglærer. Dette høres i utgangspunktet ut som en veldig god ide, men dette krever at studenten har tilstrekkelige forkunnskaper innen emnet og metakognitive ferdigheter slik at hen kan oppdage mulige hallusinasjoner og feil fra det generative KI-systemet. I en slik setting bør derfor underviseren utføre utvalgte prompt (knyttet til emnet) selv og være tilgjengelig for studenten når

¹⁶ <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

¹⁷ <https://www.youtube.com/watch?v=BFusp3oJrgA&list=PLeYxHW3kH5ocx0Hk7k0WwnRgBJzSD-VRA&index=20> (Pedagogisk intelligent podkast, Episode 5: Dyslektikerens beste venn)

hen har spørsmål om svarene fra det generative KI-systemet. På den måten kan det være at en personlig læringsassistent bidrar til mer og ikke mindre arbeid for faglærer, men at det da blir de studentene som virkelig trenger hjelp fra faglærer som får denne hjelpen. Oppsummert vil det være behov for at faglærer aktivt jobber for å kvalitetssikre bruken av en slik digital læringsassistent.

Generativ KI, læringsmål og støttende læringsaktiviteter

Et viktig læringsmål knyttet til KI på programnivå bør være å oppnå en god forståelse for hva generativ KI kan og ikke kan gjøre, hvordan treningsdataene som ligger til grunn for verktøyet vil påvirke svarene man får, hvilke statistiske prinsipper som ligger til grunn for teknologien, kritisk refleksjon rundt svar fra generative KI-verktøy og kritisk bruk av kilder. *Vi ser ikke at det er naturlig å definere generelle KI-læringsmål i alle emner, men ser viktigheten av at dette blir gjort på programnivå.*

Kritisk refleksjon kan man for eksempel arbeide mot ved at undervisere legger inn sekvenser av sine prompt og språkmodellens svar i undervisningsressurser (for eksempel en øving eller forelesning), og så ber studentene kommentere og reflektere over bruken.

Som et eksempel på en læringsaktivitet som støtter opp under forståelsen for hvilke statistiske prinsipper som ligger til grunn for store språkmodeller ble det i emnet ISTx100y Statistikk (for bachelor ingeniør, fler-campusemne) laget en oppgave der studentene selv kunne bruke en enkel språkmodell (basert på betinget sannsynlighet og et datasett fra Reuters) til å generere en tekst der den underliggende modellen beregner sannsynligheten for neste ord basert på de to foregående ordene¹⁸. Denne oppgaven benytter kun svært enkel statistisk teori, men reflekterer på en god måte prinsippene bak store språkmodeller. Studentene kunne her også observere hvordan skjevheter i dataene (nyhetsartikler) knyttet til kjønn, etnisitet og religion slo ut på de predikerte setningene. Dette emnet er et såkalt serviceemne (emnet gis av fagstab i statistikk til brukere av statistikk i bacheloringeniørutdanningen).

Bruk av generativ KI for å støtte underviseren

Fra faglærers ståsted har generativ KI et potensiale til å forbedre kvaliteten på undervisningen i et emne.

I emner med tradisjonelle forelesninger kan faglærer bruke generativ KI til å hjelpe med å strukturere selve forelesningen eller forbedre strukturen på en eksisterende forelesning.

Store språkmodeller kan også utnyttes av faglærer for å produsere ulike case eller konkrete problemstillinger i forberedelse av forelesning, læringsaktiviteter eller eksamens- og vurderingsoppgaver. Språkmodellene kan da brukes både for å generere et utvalg alternative startpunkt for en problemstilling og for videre konkretisere og/eller raffinering av den.

Konsepter og tema som er vanskelig for studenter å forstå kan gjøres mer tilgjengelig ved å inkludere analogier og eksempler som er tilpasset studentene (alder, faglig bakgrunn), og her kan generativ KI bidra. Det samme gjelder kontekstualiserte eksempler mot studentenes eget studieprogram eller fra studentenes hverdag. Men, som fremhevet i UNESCO-rapporten vil et utstrakt bruk av generativ KI på denne måten potensielt kunne virke negativt på faglæreres kreative prosesser.

Studentene opplever ofte undervisning som bygger på *learning by doing*, som spesielt motiverende og engasjerende fordi de gjøre umiddelbare erfaringer med bruk av teori, noe som kan lede dem inn

¹⁸ https://github.com/imfdrift/istx100y2023/blob/main/uke3/enkel_spraakmodell.ipynb

i en positiv læringsløype. I noen tilfeller vil *learning by doing* bestå i å anvende teori i en simuleringsomgivelse mens i andre tilfeller vil det kunne bestå i å utvikle programkode for å teste det teoretiske stoffet. Slike læringsaktiviteter kan gi bedre læringsutbytte og mestringsfølelse enn ren forelesningsbasert undervisning, men de krever at studentene legger ned mer tid og dermed at færre tema kan bli dekket i emnet. Det er i denne sammenhengen at store språkmodeller kan ha en positiv effekt. De kan bidra til at studentene bruker mindre tid på selve kodingen og mer tid til å teste og videreutvikle sin forståelse av teorien. En slik effektivisering av *learning by doing*-undervisning kan bety at det da blir mer plass til andre temaer og at det blir mulig å beholde fordelene ved *learning by doing* uten å måtte redusere emnets innhold og omfang.

Arbeidsgruppen er av den oppfatning at det er den enkelte emnelærer som må finne ut hvordan KI påvirker hens emne med tanke på læringsutbyttebeskrivelser, undervisning og vurdering.

Vurdering

I følge universitets- og høyskolelovens paragraf § 3-9(1)¹⁹ skal studentens kunnskaper og ferdigheter bli prøvet på en upartisk og faglig betryggende måte. Inntoget av generativ KI utfordrer etterlevelsen av denne lovteksten på flere måter. Studentens *kunnskaper og ferdigheter* vil, i mange situasjoner, ikke være reflektert i en besvarelse generert av kunstig intelligens. Men, i en ny undervisningssituasjon med stort innslag av KI vil konstruktiv og kreativ bruk av KI være nettopp det som reflekterer de ønskede kunnskapene og ferdighetene. Videre utfordres kravet om en *faglig betryggende* vurdering av i) den senkede terskelen for fusk ved hjelp av KI og ii) uuttalte og motstridene forventninger til bruk av KI mellom emneansvarlig og sensor. Der hvor det foreligger slike motsetninger kan også lovtekstens ordlyd om *upartisk* vurdering bli underminert: en sensor som vurderer i den oppfatning at KI genererte besvarelser er uønsket, vil premiere studenter hvor slike elementer er vanskelige å oppdage eller ganske enkelt fraværende. Studenter som har anvendt generativ KI i henhold til god praksis (hvis en slik praksis en dag blir definert) og emneansvarliges forventninger vil straffes urettmessig.

Et viktig problem i den kommende tiden vil bli å harmonisere i) emneansvarliges forventninger til og rammer for bruk av KI i besvarelser, ii) studentens forståelse av redelig og konstruktiv bruk av KI og iii) sensors forståelse av vurderingskriterier som kan komme til å inkludere bruken av generativ KI.

De ovenstående utfordringene vil manifestere seg i varierende grad og på forskjellig vis avhengig av vurderingsform. Ved fakultetet benyttes det i dag en rekke ulike vurderingsformer. I løpet av studieåret 2022/2023 ble det gjennomført nærmere 1 300 ulike avsluttende vurderinger ved fakultetets 42 studieprogram (ph.d. ikke inkludert) Skriftlig skoleeksamen sammen med muntlig eksamen utgjorde ca. 40 % av alle vurderingsformene. Det betyr igjen at øvrige vurderingsformer gjennomført uten tilsyn utgjorde ca. 60 %.

Arbeidsgruppen har vurdert hvordan KI vil kunne påvirke gjennomføring av vurdering, og har definert følgende kategorier av vurderingsformer

1. Vurderingsformer uten tilsyn
2. Master- og bacheloroppgaver spesielt
3. Skriftlig eksamen under tilsyn uten hjelpemidler
4. Skriftlig eksamen under tilsyn med begrensninger i hjelpemidler

¹⁹ <https://lovdata.no/lov/2005-04-01-15/§3-9>

5. Skriftlig eksamen under tilsyn med alle hjelpemidler
6. Muntlig eksamen
7. Praktisk eksamen

Vurderingsformer uten tilsyn

Vurderingsformer uten tilsyn vil for eksempel være hjemmeeksamen, prosjektoppgaver, arbeider som inngår i mappevurdering, gruppeoppgaver mm. Dette kan både være som eneste vurderingsform i et emne eller inngå som en delvurdering. Ved NTNU, som ved de fleste andre undervisningsinstitusjoner, er det lang tradisjon for å benytte vurderingsformer uten tilsyn. Slike vurderingsformer er altså ikke nye, det nye nå i et KI-perspektiv er at det eksisterer verktøy og hjelpemidler som i mye større grad kan bistå studentene i besvarelsene, sågar utforme besvarelsene for studentene. En sikker måte å unngå at studentene fusker ved å bruke KI til å utferdige besvarelser, vil være å gå bort fra alle andre vurderingsformer enn skriftlig eksamen under tilsyn uten hjelpemidler eller muntlig eksamen. Arbeidsgruppen ser dette som lite hensiktsmessig, ikke minst da en slik beslutning vil overstyre den enkelte faglærers rett og anledning til å velge det hen mener er den/de beste vurderingsformene i et emne. Arbeidsgruppen er av den oppfatning at vi fortsatt skal operere med vurderinger som gjennomføres uten tilsyn, og at en heller bør prioritere å se på behovet for retningslinjer og reguleringer for disse vurderingsformene. Verktøy i tilknytning til generativ KI er kommet for å bli, og det både er og vil bli en større og mer sentral del av samfunns- og arbeidsliv fremover.

Rektorvedtaket av februar 2023 vedrørende fusk og bruk av kunstig intelligens (se sitat tidligere i denne rapporten) kan tolkes på flere måter, blant annet kan det velges å tolkes dit at det er ok å få hjelp til å lage besvarelse ved hjelp av kunstig intelligens, så lenge man opplyser om det. Det er med andre ord behov for å komme med klargjøringer/presiseringer når det gjelder bruk av kunstig intelligens i studentenes besvarelser.

Arbeidsgruppen anbefaler derfor at fakultetet tolker rektorvedtaket dit at det er aksept for at studentene benytter KI i arbeidet med sine besvarelser, *men at studentene som velger å benytte KI i forbindelse med utarbeidelse av besvarelser må opplyse om dette i form av hvilke verktøy som er benyttet, hvordan dette/disse er benyttet, og at tekst som direkte er produsert av KI må markeres etter gjeldende regler.* I en vurderingssammenheng må dermed type bruk og omfang av bruk av KI gis vekt i den samlede vurderingen av en besvarelse. Videre, sensorveiledninger må fremover utarbeides slik at forventninger til bruk av KI fremgår tydelig og i henhold til emnets læringsmål.

Arbeidsgruppen anbefaler at man betrakter resultater fra KI-systemer som sekundærreferanser - det vil si referanser som om nødvendig må få bekreftet ved å oppsøke primærkilder. Undervisere bør fortelle studentene at de forventes å ta i bruk tradisjonelle søkesystemer for å finne passende primærreferansene. Dette vil tilsvare situasjoner der studenter finner noe på wikipedia, og vi ber dem om å gå til en forskningsdatabase for å finne solide kilder.

Arbeidsgruppen er av den oppfatning at det ikke gir mening å sette begrensninger/forbud knyttet til vurderinger som vi ikke kan følge opp/kontrollere. Dersom det er behov for å gjennomføre vurdering uten bruk av KI må en velge en vurderingsform hvor denne bruken kan kontrolleres.

Obligatoriske aktiviteter er i en annen kategori. Her vurderer arbeidsgruppen det dit at om emneansvarlig finner det hensiktsmessig, så er det ingenting i veien for at slike aktiviteter har begrensninger i bruk av KI.

Master- og bacheloroppgaver

Skrivingen, og den påfølgende vurderingen, av bachelor og masteroppgaver (vi regner også 4-klasse prosjektoppgaver i sivilingeniørstudiet her) er kanskje studieaktiviteten som er mest eksponert for de nye mulighetene som generativ KI kan tilby.

Hovedutfordringen er åpenbar: det er svært enkelt for en student å generere relevant materiale for en rapport ved hjelp av KI. Og, det genererte materialet kan være svært vanskelig å avdekke for sensor²⁰.

Rektorvedtaket for bruk av generativ KI minner svært mye om retningslinjene for plagiat som man for eksempel kan finne i studieforskriften §6-4. Generativ KI er *likt* plagiat i den forstand at det fjerner en byrde med selvstendig tenkning fra studenten, men *ulikt* i den forstand at det genererte materialet ikke kan spores tilbake til en bestemt kilde. Det er tilsynelatende ingen gode verktøy for å avdekke bruken av generativ KI, i motsetning til plagiat hvor man har flere verktøy for plagiatkontroll. Generativ KI er dermed en lett tilgjengelig og effektiv metode for fusk som er svært vanskelig å avdekke.

Vi har fremhevet likheten mellom generativ KI og plagiat, som også er reflektert i Rektorvedtaket. Det er nå betimelig å merke seg at reglene om plagiat ikke er ment for å forhindre bruk av andres materiale og resultater, men for å forhindre i) tyveri av immaterielle rettigheter og ideer samt ii) uærlig fremstilling av egne ferdigheter. Det er, ved første øyekast, bare det siste punktet som er relevant for fusk med generativ KI.

Motsatsen til plagiat er etisk og redelig bruk av andres arbeid i utarbeidelse av en rapport, gjerne gjennom en solid fundering i tidligere arbeid forankret i en fullstendig referanseliste. Dette er en svært viktig del av akademisk arbeid som universitetet formidler til studentene gjennom kursing og veiledning. Men, hvor plagiat har en motsats i kildehenvisning, så er en slik motsats ikke veldefinert i tilfellet generativ KI.

Generativ KI er, på lik linje med litteraturstudier og kildehenvisninger, et verktøy som kan bidra positivt både til studentens rapporter og deres personlige utvikling og læring. KI-systemer basert på store språkmodeller (tekstverktøy) slik som ChatGPT kan bidra med idémyldring og kritisk refleksjon mens Grammarly (som er KI basert) kan avhjelpe utfordringer med rettskriving og grammatikk. Kodeverktøy, slik som GitHub Copilot, kan bidra med prototyping og debugging av programmer som utvikles som del av rapportskrivningen, mens bildegenereringsverktøy, slik som Midjourney, kan bistå med illustrasjon.

Det hadde vært svært uheldig hvis Rektorvedtaket om bruk av generativ KI blir tolket slik at bruk av de nye verktøyene forbyes. Med plagiat og sin motsats kildehenvisning som analogi vil vi foreslå at *master og bacheloroppgaver leveres med en KI-deklarasjon som redegjør for bruken av generativ KI.*

Arbeidsgruppen har utarbeidet et forslag til skjema som tar for seg dette, og som vi foreslår skal fylles ut av studentene for alle master- og bacheloroppgaver. Se vedlegg B med forslag til skjema. Studenten laster opp utfylt skjema i Inspira sammen med sin bachelor- eller masteroppgave. Denne løsningen er ment å bidra på følgende vis:

²⁰ Debby R. E. Cotton, Peter A. Cotton & J. Reuben Shipway (2023) Chatting and cheating: Ensuring academic integrity in the era of ChatGPT, *Innovations in Education and Teaching International*. DOI: 10.1080/14703297.2023.2190148

1: Sensor blir informert om hvilke verktøy som har blitt brukt, i hvilken utstrekning og til hvilke formål. Dette ivaretar formålet med vurderingen som skal reflektere studentens kompetanse. Videre kan studentens bruk av generative verktøy også bli utsatt for vurdering, i det denne bruken har blitt synliggjort.

2: Studenten betrygges om at deres bruk av generativ KI ikke kvalifiserer til fusk, i det den har blitt tydelig deklart. Dette vil kunne fostre redelig og konstruktiv bruk av verktøyene. Innfyllingen av skjemaet kan også ha en bevisstgjørende effekt på både student og veileder.

3: Fakultetet har dermed tilgang til skjemaet, og det kan brukes til å få en oversikt over bruken av KI-verktøy blant studentene. Dette vil kunne rette arbeidet med utarbeiding av kursing for både studenter og ansatte.

Som beskrevet over, så mener arbeidsgruppen at det ikke er riktig vei å gå og forby bruk av KI i studentenes besvarelser. For bachelor- og masteroppgaver spesielt mener arbeidsgruppen også at det skal være mulig å benytte KI for studentene og at det skal opplyses om dette. Konstruktiv bruk av KI med gode refleksjoner og kritisk tenkning bør ses på som positivt.

Ved noen studieprogrammer inngår en muntlig presentasjon i vurderingsgrunnlaget for bachelor- og/eller masteroppgave, for andre kan en muntlig presentasjon være en obligatorisk aktivitet og for noe er det ingen muntlig presentasjon. Videre så er det noen studenter som skriver slike oppgaver helt alene, mens andre skriver i grupper av ulik størrelse. Ved utstrakt bruk av KI-verktøy kan det være fornuftig å legge vekt på at studenten(e) i tillegg til den skriftlige oppgaven gir en muntlig redegjørelse for arbeidet, der sensorer (eller andre) kan stille spørsmål og på en god måte kan passe på at vurderingen gir på bakgrunn av studentenes læring og ikke kun på studentenes bruk av KI-verktøy.

Skriftlig eksamen under tilsyn uten hjelpemidler

Dette er en vurderingsform som ikke gir studentene anledning til å benytte verktøy for kunstig intelligens.

Skriftlig eksamen under tilsyn med begrensning i hjelpemidler

Her vil faglærers vurdering av tillatte hjelpemidler være avgjørende for om KI skal kunne være et hjelpemiddel for studentene. Hvis det tillattes, er det arbeidsgruppens anbefaling at det også her skal opplyses om hvilket verktøy som er benyttet, hvordan det er benyttet, og at ren «avskrift» skal merkes etter gjeldende siteringsregler.

Skriftlig eksamen under tilsyn med alle hjelpemidler

Samme vurdering som punktet over; det skal opplyses om hvilket verktøy som er benyttet, hvordan det er benyttet, og ren «avskrift» skal merkes etter gjeldende siteringsregler.

Muntlig og praktisk eksamen

Det er arbeidsgruppens anbefaling at det også for disse vurderingsformene angis i emnebeskrivelsene hvilke hjelpemidler som er tillatt. Selve bruken av KI vil kunne observeres av sensorer, og følgelig ser vi ikke noen grunn til at det skal presenteres hvordan KI er benyttet.

Arbeidsgruppens anbefalinger

Arbeidsgruppen forventer at NTNU har personverntrygge og profesjonelt forsvarlige KI språkmodellverktøy (som ivaretar GDPR og sikkerhet på en god måte) som NTNUs studenter kan bruke (senest) fra høsten 2024.

IE må tilby **faglærerne ved IE opplæring** i forståelse og bruk av generative KI-verktøy, slik at de kan bruke disse i sine emner på en god måte. Dette bør skje i god tid før høsten 2024, for at faglærerne kan være godt forberedt på arbeid knyttet til eventuelle læringsmål for KI i IEs emneportefølje. Dette bør prioriteres høyt. Det finnes allerede gode nettbaserte kurs (som for eksempel mikrokurset fra Aalborg universitet, se Vedlegg A), men vi ser store fordeler ved at et undervisningsopplegg også inneholder diskusjon blant kolleger. *Vi anbefaler i første omgang at IE skisserer et opplegg basert på mikrokurset ved Aalborg universitet, koblet med oppgaver som er relevante for faglærere på hvert institutt og at gjennomføring av opplæringen skjer på instituttnivå våren 2024 (med mulig støtte fra ressurspersoner på IE).*

Når det gjelder serviceundervisning til studenter, så bør dette komme i neste versjon av våre anbefalinger. Etter emnerevisjonene bør fakultetet kartlegge behovet for grunnleggende undervisning til studentene, vurdere hva vi allerede har og beskrive behovet for serviceundervisning. I tillegg må fakultetet se på kapasitet for å lage til emner og delemner samt mulighet for å gjennomføre disse.

NTNU og/eller IE må tilby **studentene ved IE opplæring** (IEs studieprogram og emner IE som tilbys andre) i forståelse og bruk av generative KI-verktøy, slik at de kan anvende KI på en god måte og i tråd med de regler og retningslinjer som er/blir etablert. Dette bør prioriteres og implementeres ideelt sett fra våren 2024, senest fra høsten 2024. Studentene bør inkluderes i arbeidet med å utvikle opplæringen. I arbeidet med dette bør både form og innhold vurderes. Når det gjelder form, anbefaler vi at nettbaserte kurs vurderes, og at NTNUs Læringsstøttesenter eller andre serviceaktiviteter ved NTNU bidrar med kunnskap og ressurser. Vi mener videre at innhold må vurderes både ut fra behov og ut fra hva studentene tilbys allerede i dag. Det er stor diversitet mellom studieprogrammene ved NTNU både når det gjelder innholdet i utdanningstilbudene og behovet for KI som en del av dem. Det finnes videre elementer ved KI som tangerer andre ressurser ved NTNU. Vi mener blant annet at det bør vurderes om *Universitetsbiblioteket gis i oppgave å utarbeide regler for hvordan man viser til bruk av KI-verktøy (deklarasjon og direkte sitater) deretter gi opplæring i hvordan referere til KI-verktøy i besvarelser.*

Når det gjelder **bruk av KI-verktøy i undervisningen**, så er det nå begrenset av at studenter ikke har personverntrygg og profesjonelt forsvarlig tilgang, og bruk bør derfor enten gjøres frivillig eller at man begrenser seg til oppgaver/dialog laget og kommunisert med generativ KI-verktøy av underviser. IE bør samle inn gode eksempler på bruk av KI og KI-verktøy i undervisningen, gjerne gjennom nestledere for utdanning ved hvert institutt. Dette kan gjøres tilgjengelig i opplæringsopplegget for ansatte (skissert over).

For emner der emneansvarlig ønsker å benytte **vurdering uten tilsyn** anbefaler arbeidsgruppen det spesifiseres i emnebeskrivelsen hvordan bruk av KI-verktøy kan/skal gjøres og dokumenteres i arbeidet som skal vurderes. Dette må være samstemt med læringsmålene som er satt i emnet, og trygg bruk av KI-verktøy bør da være en del av undervisningen i emnet. En deklarasjon av bruken i form av et skjema kan selvfølgelig benyttes ved behov bestemt innen hvert emne.

Videre, har vi for **bachelor- og masteroppgaver** laget et utkast til en "Deklarasjon om KI-hjelpemidler" som vi anbefaler at studentene skal fylle ut og laste opp i Inspira sammen med oppgaven. Bruken av KI i oppgaveskriving er på nåværende tidspunkt beheftet med svakt definerte retningslinjer og studentene kan ikke forventes å ha fått opplæring i redelig bruk av de nye verktøyene. Deklarasjonen synliggjør bruken av KI i en oppgave slik at i) bruken kan bli utsatt for vurdering og ii) kandidaten betrygges om at bruken ikke misforstås som fusk. Vi betrakter skjemaet som i) en midlertidig løsning før et mer komplett rammeverk for bruk og sitering av KI-verktøy

kommer på plass og ii) et førsteutkast som kan tilpasses behovene ved det enkelte institutt/fakultet. Det forventes at sensurveiledninger for bachelor- og masteroppgaver oppdateres for å reflektere bruk av deklarasjonen. Videre anbefaler vi emneansvarlige i samråd med studieprogram og tilhørende institutt(er) for denne typen oppgaver å vurdere om en **mundlig presentasjon av oppgaven** (der veileder og sensurkomité er til stede) bør inngå (hvis det ikke allerede gjør det) og om denne presentasjonen bør gjøres før eller etter at oppgaven er levert. Arbeidsgruppen har ikke gått inn i grundige vurderinger av dette, og ser at det må gjøres nøye vurderinger av om muntlig presentasjon av oppgaven skal inngå. Vi ser både fordeler og ulemper med dette. En av de opplagte fordelene er at en på en god måte kan validere at vurderingen gis på bakgrunn av studentenes læring og ikke kun på studentenes bruk av KI-verktøy. Klageprosesser knyttet til klage på karakter på vurdering som inneholder justerende muntlig er ressurskrevende og omfattende, og vurderes derfor her til en ulempe. Formen på muntlig presentasjon vil derfor være viktig å avklare, sammen med en vurdering om en slik presentasjon skal være inngå i det endelige resultatet på oppgaven.

Plan for veien videre ved IE og NTNU

Det er viktig å ha en god prosess rundt det videre arbeidet med bruk av generativ KI ved IE og med anbefalingene vi har skissert i denne rapporten. Følgende er steg som bør gjennomføres på kort sikt.

- 1) Rapporten og anbefalingene presenteres på Ledersamlingen ved IE 17.oktober
- 2) Rapporten og anbefalingene presenteres til studieprogrammer gjennom Studieutvalget ved IE, og til nestledere for utdanning gjennom IE Utdanning (hvis nødvendig, da disse deltar på ledersamlingen over).
- 3) Nestleder for utdanning ved hvert institutt gis ansvar for at alle faglærere ved instituttet blir kjent med rapporten og anbefalingene.
- 4) Basert på diskusjon av våre anbefalinger, gjør emneansvarlige endringer i emnebeskrivelser og spesielt læringsutbyttebeskrivelser i Emner på nett innenfor høstens frister. Ved behov diskuterer emneansvarlige sine forslag til endringer med nestleder for utdanning.
- 5) IE utarbeider et opplærings-/kursopplegg for faglærere ved IE, og dette gjennomføres på hvert institutt i løpet av våren 2024. Hvis det er mulig ønsker vi at IE først utarbeider et instrument (for eksempel en spørreundersøkelse) for å kartlegge kompetansenivå og kompetansebehov innen KI for faglærere ved IE.
- 6) På studieprogramnivå arbeider man med å oppdatere læringsutbyttebeskrivelser for programmet, for deretter å harmonisere med endringer som er gjort i emnebeskrivelser for emner som er obligatoriske og valgbare for studieprogrammets studenter. Dette ser vi for oss at ikke det er mulig å gjøre til fristen for Emner på nett høsten 2023, men at endringer som studieprogrammer-emneansvarlige og emneleverandører (instituttene) kommer frem til kan endres høsten 2024 (med effekt fra høsten 2025).

Arbeidsgruppen ser det er et stort behov for snarlig å ta diskusjonen med NTNU, om hvordan vi skal håndtere KI ved NTNU. Vi forventer at NTNU har på plass godkjente KI verktøy for studenter i god tid før høsten 2024. Dette er nødvendig for å muliggjøre de anbefalingene vi har gitt.

Forventningsavklaring når det gjelder beslutninger og aktivitet hos henholdsvis Rektor, fakultetene og instituttene, er essensielt i det videre arbeidet med KI ved NTNU.

Vedlegg A: Læringsressurser for å tilegne seg kunnskap innen KI og generativ KI

Elements of AI

<https://www.elementsofai.com/no>

«Elements of AI er et gratis nettkurs utviklet av Reaktor, Helsingfors Universitet og Feed. NTNU tilbyr kurset på norsk. Kurset passer for deg som vil lære om hva kunstig intelligens (KI) er, hvordan KI utvikles, og hva teknologien kan (og ikke kan) brukes til. Kurset er tekstbasert og kombinerer teori med praktiske oppgaver. Gjennomsnittlig tid for å fullføre kurset er 25 timer, og du kan ta det i ditt eget tempo.»

Her kan du lese mer om prisen kurset har fått: <https://www.ntnu.edu/ailab/elements-of-ai>

Elements of AI har i september 2023 annonsert at kurset nå også inneholder nye elementer om nye teknikker innen KI og maskinlæring, store språkmodeller og ChatGPT, samt etiske og samfunnsmessige spørsmålsstillinger.

Maskiner som tenker

<https://bok.norli.no/maskiner-som-tenker>

[Inga Strømkes](#) bok om kunstig intelligens har ligget på topplisten over de mest solgte bøkene i Norge i mange uker, og egner seg svært godt til å få en god oversikt over KI og utfordringene vi står overfor.

Practical AI for instructors and students

The Wharton School (Ethan and Lilach Mollich) har laget en videoserie består av fem videoer. Disse er laget på et slikt nivå at dette trolig vil passe godt for studenter og ansatte ved NTNU, og fokus er på generativ KI.

1. Introduction: <https://www.youtube.com/watch?v=t9gmyvf7JYo>
2. Large language models: <https://www.youtube.com/watch?v=ZRf2BfDLIIA>
3. Prompting: <https://www.youtube.com/watch?v=wbGKfAPIZVA>
4. AI for teachers: <https://www.youtube.com/watch?v=SBxb5xW7qFo>
5. AI for students: <https://www.youtube.com/watch?v=ZorvXYUZtRg>

Introduction to Generative AI & ChatGPT

Aalborg universitet lanserte i august mikrokurs, som er åpne også for studenter og ansatte utenfor Aalborg universitet, men man må lage en brukerkonto for å ta kurset.

Her finner du kurset: "Introduksjon til Generativ AI & ChatGPT": <https://www.micro.aau.dk/available-aau-micros/introduction-to-generative-ai-chatgpt>

Kurset er delt i to deler, i første del får man høre om hvordan (generativ) KI virker og hvordan man kan bruke generativ KI. I andre del er fokus på hvordan man bruke generativ KI. Kurset har en lengre

del (samtale) der ulike etiske aspekter (kritisk refleksjon) rundt KI står i fokus. Hvis man fullfører kurset får man kursbevis, dette krever at man svarer på flervalgsoppgaver underveis.

[Generative AI exists because of the transformer](#)

Financial Times har en god framstilling som kan gi innsikt i både hvordan språkmodellene virker og lærer, og berører også transformer-teknologien på en lettfattelig måte. Her finner du websiden:

<https://bit.ly/455smxb>.

[Digital Norway nettkurs](#)

Digital Norway er eid av bedrifter i norsk næringsliv, og er en non-profit organisasjon. Nettkursene listet under er for nybegynnere og dermed på introduksjonsnivå. For å få tilgang til nettkursene må du lage bruker. Her starter du: <https://digitalnorway.com/kurs/>

Hvordan bruke generativ kunstig intelligens? <https://digitalnorway.com/kurs/hvordan-bruke-generativ-kunstig-intelligens/>

Hva er og hvordan bruke ChatGPT? (Finnes også engelsk versjon.)

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

Lær prompting med språkmodeller som ChatGPT: <https://digitalnorway.com/kurs/prompt-enigneering/> Kurset er tekstbasert og består av 6 moduler.

Innføring i GDPR: <https://digitalnorway.com/kurs/innforing-i-gdpr/>

Vedlegg B: Deklarasjon om bruk av KI-hjelpemidler



Deklarasjon om KI-hjelpemidler

Har det i utarbeidingen av denne rapporten blitt anvendt KI-baserte hjelpemidler?

- Nei
 Ja

Hvis ja: spesifiser type av verktøy og bruksområde under.

Tekst

- Stavekontroll.** Er deler av teksten kontrollert av:
Grammarly, Ginger, Grammarbot, LanguageTool, ProWritingAid, Sapling, Trinko.ai eller lignende verktøy?
- Tekstgenerering.** Er deler av teksten generert av:
ChatGPT, GrammarlyGO, Copy.AI, WordAi, WriteSonic, Jasper, Simplified, Rytr eller lignende verktøy?
- Skriveassistanse:** Er en eller flere av ideene eller fremgangsmåtene i oppgaven foreslått av:
ChatGPT, Google Bard, Bing chat, YouChat eller lignende verktøy?

Hvis ja til anvendelse av et tekstverktøy - spesifiser bruken her:

Kode og algoritmer

- Programmeringsassistanse.** Er deler av koden/algoritmene som i) fremtrer direkte i rapporten eller ii) har blitt anvendt for produksjon av resultater slik som figurer, tabeller eller tallverdier blitt generert av:
GitHub Copilot, CodeGPT, Google Codey/Studio Bot, Replit Ghostwriter, Amazon CodeWhisperer, GPT Engineer, ChatGPT, Google Bard eller lignende verktøy?

Hvis ja til anvendelse av et programmeringsverktøy - spesifiser bruken her:

Bilder og figurer

- Bildegenerering.** Er ett eller flere av bildene/figurene i rapporten blitt generert av:
Midjourney, Jasper, WriteSonic, Stability AI, Dall-E eller lignende verktøy?

Hvis ja til anvendelse av et bildeverktøy - spesifiser bruken her:

- Andre KI verktøy:** har andre typer av verktøy blitt anvendt? Hvis ja spesifiser bruken her:

- Jeg er kjent med NTNUs regelverk: *Det er ikke tillatt å generere besvarelse ved hjelp av kunstig intelligens og levere den helt eller delvis som egen besvarelse.* Jeg har derfor redegjort for all anvendelse av kunstig intelligens enten i) direkte i rapporten eller ii) i dette skjemaet.

 Underskrift/Dato/Sted