

Bakpå og frampå

Programnotat til fakultetets KI-strategi

**Erik Børve Rasmussen, Anders Graver Knudsen,
Rebecca Schmidt og Kim Tallerås**

OsloMet Skriftserien 2024 nr. 7



Bakpå og frampå

Programnotat til fakultetets KI-strategi

Av Erik Børve Rasmussen, Anders Graver Knudsen, Rebecca Schmidt og Kim Tallerås

CC-BY-SA versjon 4.0

OsloMet Skriftserien 2024 nr 7

ISSN 2535-6984 (online)

ISBN 978-82-8364-622-1 (online)

OsloMet – storbyuniversitetet

Universitetsbiblioteket

Skriftserien

St. Olavs plass 4,

0130 Oslo,

Telefon (47) 64 84 90 00

Postadresse:

Postboks 4, St. Olavs plass

0130 Oslo

Trykket hos Byråservice

Trykket på Scandia 2000 white, 80 gram på materiesider/200 gram på coveret

Godkjent og kvalitetssikret av studieleder Alejandro Miranda Nieto, Institutt for sosialfag, Fakultet for samfunnsvitenskap

Innhold

Sammendrag	3
Mandat	4
Utredningsutvalget.....	4
Innledning	5
Noen innledende avklaringer	6
Fremgangsmåte	7
Rundspørringen	7
Arbeidsmøtet.....	8
Den mindre systematiske informasjonsinnhenting	9
Resultater: svarandeler	10
Angående KI som verktøy	10
Angående KI som tema.....	10
Resultater: innholdsanalyse	11
KI som verktøy	11
KI som tema.....	14
Utfordringer knyttet til bruk av KI i UH-sektoren	16
Diskusjon av utfordringer og muligheter	21
Gråsoner	21
Systemutfordringer	23
Muligheter for SAM	25
Avrundende kommentar	26
Konklusjon: råd og anbefalinger	28
Litteratur	29

Sammen drag

Hvordan skal vi forholde oss til kunstig intelligens (KI) i høyere utdanning? I dette notatet analyserer vi en rundspørring blant UH-ansatte på SAM, om deres tanker om og erfaring med KI i høyere utdanning og forskning. Rundspørringen ga både kvantitative og kvalitative data, og analysen kombinerer enkel statistikk med en enkel innholdsanalyse. Analysen trekker også på opplysninger fra et arbeidsmøte hvor KI i høyere utdanning ble diskutert, og på en ustrukturert informasjonsinnhenting fra instituttene ved SAM. Basert på analysen argumenterer vi for at fakultet både er langt frampå og langt bakpå vedrørende KI-teknologi. UH-ansatte ved SAM er langt frampå med bruk av KI-teknologi og bruker KI mye og til mye forskjellig, inkludert sentrale akademiske oppgaver. Derfor er også fakultetet langt bakpå med strategiarbeidet, og det haster å komme i gang med fellestenkningen rundt hva vi regner som god og grei bruk. Vi argumenterer også for at SAM har mange strategiske muligheter hva gjelder bruk og tematisering av KI i forskning og undervisning. Dessuten argumenterer vi for at SAM har et særlig ansvar for å bidra med systemiske perspektiver til universitetets retningslinjer og fagmiljøenes tenkning om KI. Analysen av rundspørringen gir flere nyttige innspill i så måte, og i notatet diskuterer vi flere av disse. Til sist gir vi 7 gode råd til fakultetets strategiarbeid.

Mandat

Den 5. mars 2024 diskuterte fakultetets ledermøte saken «Utredning om KI-satsing ved SAM». Dekan Oddgeir Osland besluttet følgende:

«Det skal lages et kort programnotat som

- 1) gjør rede for pågående og planlagte initiativ for forskning på KI og bruk av KI i utdanningene på fakultet, og
- 2) diskuterer sentrale utfordringer og muligheter som KI representerer i utdanning og forskning og hvilke konsekvenser dette kan bør få på institutt og fakultet.

Utredningen foregår parallelt med utvikling av strategisk handlingsplan på fakultetet våren 2024, med rom for å foreslå tilhørende satsnings- og tiltaksområder, men forventes å kunne gi grunnlag for diskusjoner i FoU-, utdanningsutvalg og andre fora også etter at strategiplanen er vedtatt.

Utredningen ledes av førsteamanuensis Erik Børve Rasmussen ved Institutt for sosialfag, det oppnevnes en representant fra hvert av de andre tre instituttene til utvalget. Utvalget legger selv opp prosessen med arbeidet, slik de finner ønskelig. Selv om utredningsfristen er kort, er det ønskelig at det gis rom for innspill og diskusjoner med fagmiljøene som ikke er representert i utvalget underveis i prosessen.»

Programnotatet skulle ferdigstilles innen 24. mai 2024.

Utredningsutvalget

I tillegg til Rasmussen (SF), ble følgende utvalgsmedlemmer oppnevnt:

- Førstelektor Anders Graver Knudsen (JM)
- Førsteamanuensis Rebecca Schmidt (HH)
- Førsteamanuensis Kim Tallerås (ABI)

Innledning

Etter at den generative språkmodellen «ChatGPT» kom på banen i november 2022 stilles det spørsmål ved hvordan vi skal forholde oss til denne typen teknologi som på kort tid har fått en allmenn utbredelse og som allerede brukes til å løse vesentlige oppgaver i mange sektorer. Dette gjelder ikke minst i vår egen sektor hvor studenter og forskere allerede oppgir å bruke generative KI-verktøy til en rekke ulike formål. I likhet med UH-sektoren ellers jobber derfor Fakultet for samfunnsvitenskap (SAM) og resten av OsloMet – storbyuniversitetet med å posisjonere seg.

Dette notatet er ment som et av flere underlag i utviklingen av fakultetets KI-strategi, og gir hovedsakelig tre bidrag til strategiarbeidet:

- 1) **Vi presenterer en kartlegging av UH-ansattes pågående og planlagte KI-initiativ, og hva de ser som viktige KI-relaterte utfordringer.** Kartleggingen er basert på en nettbasert rundspørring, og består av innspill fra ABI (Institutt for arkiv-, bibliotek- og informasjonsfag), JM (Institutt for journalistikk og mediefag), HH (Handelshøyskolen) og SF (Institutt for sosialfag). I tillegg baserer den seg på et arbeidsmøte i regi av fakultetsledelsen, og noe supplerende informasjonsinnhenting (mer om datakildene nedenfor).
- 2) **Basert på kartleggingen diskuterer vi sentrale utfordringer og muligheter relatert til KI, og hvilke konsekvenser dette kan/bør få på institutt, fakultet og universitetsnivå.** I diskusjonen begrenser vi oss ikke bare til innsikter fra kartleggingen, men tar oss friheten til å trekke inn innsikter og viktige problemstillinger vi mener må med i en ansvarlig strategiprosess – og som vi ikke minst mener et samfunnsvitenskapelig fakultet er i en særstilling til å kunne håndtere på en klok måte. Basert på analysen og diskusjonene kommer vi med 7 konkrete råd og anbefalinger for det videre strategiarbeidet.¹
- 3) **I både kartleggingen og diskusjonen gjør vi bruk av et analytisk skille som vi tror er et klargjørende bidrag i det videre strategiarbeidet, mellom KI som verktøy og KI som tema.** På den ene siden brukes *KI som verktøy* til å løse sentrale oppgaver i UH-sektoren knyttet til administrasjon, FoU og undervisning. I tillegg brukes det av studentene som studieverktøy. På den andre siden er KI også *tema i undervisning og for forsknings- og utviklingsprosjekter* (altså at man underviser om og utforsker KI-bruk som samfunnsfenomen).²

Samlet sett utgjør notatet et argument med følgende struktur:

- **KI som verktøy: Vi er langt frampå og langt bakpå**
 - Vi er langt frampå med bruk, men derfor også langt bakpå med strategiarbeidet. KI brukes mye og til mye forskjellig, også sentrale akademiske oppgaver, og det haster å komme i gang med fellestenkningen rundt hva vi regner som god og grei bruk.
 - SAM har et særlig ansvar for å bidra med systemiske perspektiver til universitetets retningslinjer. Her er flere av de rapporterte bekymringene fra de ansatte nyttige innspill.
- **KI som tema: Vi er godt i gang, og har strategiske muligheter**
 - Trass i at «kvantesprangene» er relativt nye, så har vi allerede flere pågående og påtenkte, og ikke minst relevante, FoU-prosjekter gående ved hvert institutt.
 - KI er også inne/er planlagt inn i undervisningen på alle institutt, både tematisert som verktøy i studiet og i profesjonell virksomhet, og som samfunnsfenomen.

¹ Takk til Emily Weitzenboeck for innspill på deler av analysen (jussen).

² I enkelte tilfeller kan det være vanskeligere å skille KI som verktøy og tema, f.eks. i prosjekter hvor man jobber med teknologiutvikling. Her regner vi slike tilnærminger, altså hvor teknologien i seg selv er et forsknings- eller utviklingsobjekt, som en form for tematisering.

- SAM kan velge å la utviklingen fortsette (nedenfra og opp), men kan også prioritere noen retninger (ovenfra og ned).

Noen innledende avklaringer

Et par avklaringer før vi går videre med kartleggingen og diskusjonen. Først, angående vår bruk av begrepet «KI». I konklusjonen argumenterer vi for å unngå det generelle begrepet «KI» og heller snakke om konkrete KI-teknologier, altså konkrete programvarer som *ChatGPT*, *Keenious* eller *Midjourney*. Likevel brukte vi bare det generelle «KI» i rundspørringen. Grunnen er at vi antok at flere ville ha noe å komme med om vi spurte generelt (evt. med noen eksempler) enn om vi spesifiserte bestemte verktøy. Det tror vi har fungert, men det betyr likevel at også kartleggingen i hovedsak er på et generelt nivå.

Derneft, angående formen på diskusjonen vår vedrørende bruk av KI: Utgangspunktet vårt er at teknologien og praksisene er nye og derfor utgjør praktisk-etiske *gråsoner*. Vi har enda ikke (videre)utviklet adekvate forståelser av god eller grei bruk av KI, hverken i utdanning eller forskning, og vi er kommet omtrent like kort mtp studentenes bruk som de UH-ansattes bruk. Vi la til grunn at kartleggingen ville vise at vi tenker ulikt om normene for KI-bruk, noe den også viser. Det betyr nødvendigvis at flere av respondentene vil beskrive bruk som andre respondenter og SAM-ansatte vil finne problematisk. Målet vårt er å fremme en åpen og konstruktiv dialog. Vi mener derfor at det er både viktig og klokt å bruke innspillene fra kartleggingen åpent, uten dom, og med formål om å identifisere problemer og muligheter i den pågående og planlagte bruken av KI. Når vi løfter frem og problematiserer enkelte bruksformer, så er det derfor ikke med en dømmende hensikt, men med formål om å reise spørsmål som sektoren trenger å ta stilling til, og å problematisere svar vi kanskje ikke er kommet langt nok med å tenke på. Vi håper med andre ord at notatet bidrar til konstruktive samtaler om hvordan vi går frem, og til å identifisere i alle fall en del av de viktige problemstillingene og mulighetsrommene som det haster å utvikle en foreløpig mening om. Vi oppfordrer til å holde på en slik åpen og tolerant «tone» for «meningsmangfold i gråsonene» i det videre strategiarbeidet.

Fremgangsmåte

Kartleggingen støtter seg som sagt på en rundspørring, innspill fra et arbeidsmøte, og mindre systematisk informasjonsinnhenting. Her sier vi litt mer om hver av disse fremgangsmåtene.

Rundspørringen

Første utvalgsmøte ble avholdt 18. mars, og i møtet besluttet vi å lage en rundspørring, og lagde den i Nettskjema i samme møte. Invitasjon til rundspørringen ble sendt ut neste dag. Hvert utvalgsmedlem sendte til sitt institutt. På ABI og SF ble det også sendt en purring. Brorparten av svarene kom 19. og 20. mars (N=72), og 2. april (N=16). Skjemaet stengte 3. april.

i fikk til sammen 96 svar, som utgjør 39% av SAMs 248 UH-ansatte (pr. 8. april 2024).³ Den samlede svarprosenten er hverken spesielt bra eller dårlig. Vi registrerer stor variasjon i antall svar pr. institutt (se tabell 1), men mangler informasjon om hva variasjonen skyldes utover at instituttene har ulik størrelse. Fordi vi ikke ba om opplysninger om respondenten utover institutt-tilknytning får vi heller ikke gjort en såkalt frafallsanalyse. I de tilfellene hvor vi har opplevd å mangle viktig informasjon fra rundspørringen fra et institutt har vi gjort noe ekstra informasjonsinnhenting for å kompensere (mer om det nedenfor). Vi vet heller ikke om de som har svart skiller seg på viktige måter fra de som ikke har svart. Ei heller bruker vi svarene til å fastslå omfanget av ulike bruksformer og tematisering. I stedet er målet å kunne si noe kvalifisert om den faktiske bredden i ulike typer pågående og planlagt bruk og tematisering, samt ulike typer bekymring.

Tabell 1, Antall svar og svarandeler

Institutt	Antall	Andel av svar i undersøkelsen	Andel av ansatte ved instituttet
ABI	15	16%	40%
HH	24	25%	31%
JM	5	5%	11%
SF	52	54%	60%

I spørreundersøkelsen operasjonaliserte vi «initiativ» (fra mandatet) ved å etablere det overnevnte analytiske skillet mellom *KI som verktøy*, og *KI som tema*. Vi ga også eksempler på relevante former for bruk og tematisering, for å hjelpe respondentene med å tenke gjennom om sine erfaringer og planer. Undersøkelsen stilte følgende spørsmål:

1. Benytter du og/eller studentene KI-verktøy i forbindelse med undervisning eller vurdering?
 - a. (Hvis «Ja» eller «Ikke foreløpig, men har planer om det»): Kan du helt kort beskrive bruken?
2. Tematiserer du KI i undervisningen?
 - a. (Hvis «Ja» eller «Ikke foreløpig, men har planer om det»): Kan du helt kort beskrive hvordan dette tematiseres/skal tematiseres?
3. Benytter du KI-verktøy i forbindelse med FoU-arbeid?
 - a. (Hvis «Ja» eller «Ikke foreløpig, men har planer om det»): Kan du helt kort beskrive bruken?
4. Er KI tema for FoU-prosjektene dine?
 - a. (Hvis «Ja» eller «Ikke foreløpig, men har planer om det»): Kan du helt kort beskrive hvordan KI tematiseres i FoU-prosjektene dine?

³ Tallet inkluderer 2 ansatte i stilling som professor II, og 5 vitenskapelige assistenter. Merk at tallene i Tabell 1, for UH-ansatte pr. institutt, ble etterspurt senere (5. september 2024) og oppgitt å være 245 for april 2024 (38 ved ABI, 77 ved HH, 44 ved JM og 86 ved SF).

5. Benytter du KI-verktøy i forbindelse med dine administrative oppgaver?
 - a. (Hvis «Ja» eller «Ikke foreløpig, men har planer om det»): Kan du helt kort beskrive bruken?
6. Til slutt lurer vi på hva du ser som utfordringer knyttet til bruk av KI i UH-sektoren. Hva ser du som utfordringer knyttet til bruk av KI, enten for deg selv, ditt institutt/fagfelt, for SAM, eller for UH- sektoren i stort?

I analysen oppgir vi *andelene* som har svart «Ja», «Nei», og «Ikke foreløpig, men har planer om det» på de ulike svarene. Når det gjelder andelen vi oppgir for det første spørsmålet, «Benytter du og/eller studentene KI-verktøy i forbindelse med undervisning eller vurdering?», er det en svakhet ved undersøkelsen av vi ikke kan skille mellom egen bruk og studentenes bruk blant dem som har svart «Ja» og «Nei».

Vi har også gjort en enkel *innholdsanalyse* av fritekstsvarene fra dem som har gitt korte beskrivelser av bruk, tematisering og bekymringer. Vi fikk til sammen 241 fritekstsvaer av varierende lengde (se tabell 2). De fleste fritekstsvarene var ganske korte, og fra ett ord til én setning og tre var vanligst. EBR gjennomførte en første utgave av analysen, som ble diskutert og revidert av gruppa. EBR så etter svar som lignet hverandre tematisk (typisk i form av lignende bruksformer og praksiser, som idemyldring eller skrivning av førsteutkast), og klynget disse sammen. Svar som hadde lite til felles med de fleste andre svar ble klynget sammen i restkategorien «Ymse». Alle respondentene ble informert om at graden av detaljer i beskrivelser kunne gjøre det mulig å gjenkjenne dem (f.eks. dersom de beskrev aktivitet i et bestemt emne, eller det unike innholdet i et bestemt forskningsprosjekt). Noen respondenter er derfor mulig å gjenkjenne fordi de driver med forskning eller undervisning som gjelder ganske få, og fordi de har svart ganske detaljert. Der vi omtaler eller gjengir svar i slike tilfeller legger vi til grunn at de har godtatt å være gjenkjennelige. Men i tilfeller der respondenten har svart på dialekt eller annet språk enn norsk, har vi oversatt vi til norsk bokmål for å ivareta anonymiteten. Vi har også rettet enkle skrivefeil i direkte sitater.

Tabell 2, Fritekstsvarene

Spørsmål	Antall svar	Antall ord / snitt pr. svar
1a, Bruk av KI i undervisning eller vurdering	38	884/ 23,3
2a, Tematisering av KI i undervisningen	43	753/ 17,5
3a, Bruk av KI-verktøy i FoU-arbeid	53	819/ 15,5
4a, Tematisering av KI i FoU-prosjekter	20	484/ 24,2
5a, Bruk av KI-verktøy til administrative oppgaver	21	265/ 12,6
6a, Utfordringer knyttet til bruk av KI	66	3316/ 50,2

Arbeidsmøtet

Tirsdag 16. april inviterte dekanen til «Strategitorg» for å innhente innspill på flere satsingsområder som skal tematiseres i fakultetets strategiarbeid. Blant disse var KI, og en gruppe på 10-12 personer deltok i en diskusjon om KI og ga innspill til programnotatet. Vi førte ikke opp hvem som deltok, men det var en sammensatt gruppe som blant annet bestod av en studentrepresentant, en utdanningsleder, UH-ansatte fra ulike institutt, samt to fra fakultetsadministrasjonen. Representanter fra arbeidsgruppa (EBR og KT) ledet diskusjonen og tok notater. Basert på notatene har vi skrevet inn innspillene inn i analysen nedenfor.

Den mindre systematiske informasjonsinnhenting

For å supplere analysen der vi enten visste at rundspørringen ga knapp informasjon, eller der vi basert på kjennskap til eget institutt visste om tilleggsopplysninger vi ville ha med, har vi gjennomført en uformalisert og usystematisk informasjonsinnhenting. Enten ved å lete frem i dokument eller på nettsider, eller ved å spørre utvalgte kolleger om ting vi ville vite mer om. Supplementene er bakt inn i analysen der det passet.

Resultater: svarandeler

Her viser vi andelene som svarte «Ja», «Nei» og «Ikke foreløpig, men har planer om det», på de ulike spørsmålene om KI som verktøy og KI som tema. Vi skriver bare «Har planer» i stedet for hele «Ikke foreløpig, men har planer om det». Vi setter «Ja» og «Har planer» sammen siden begge antyder en vilje til bruk, men vi synliggjør også andelene som har svart «Ja» og «Har planer».

Angående KI som verktøy

- Benytter du og/eller studentene KI-verktøy i forbindelse med undervisning eller vurdering?
 - Ja/Har planer: (33+19=) **52%**. Nei: 48%
- Benytter du KI-verktøy i forbindelse med FoU-arbeid?
 - Ja/Har planer: (55+10=) **65%**. Nei: 35%
- Benytter du KI-verktøy i forbindelse med dine administrative oppgaver?
 - Ja/Har planer: (15+19=) **34%**. Nei: 66%

Angående KI som tema

- Tematiserer du KI i undervisningen?
 - Ja/Har planer: (41+21=) **62%**. Nei: 38%
- Er KI tema for FoU-prosjektene dine?
 - Ja/Har planer: (18+10=) **28%**. Nei: 72%

Merk at når det gjelder spørsmålet «Benytter du og/eller studentene KI-verktøy i forbindelse med undervisning eller vurdering?» så vet vi ikke i hvilken grad andelen som har svart «Ja» eller «Nei» har tenkt på *egen* bruk, *studentenes* bruk, eller *begge*. Fra dem som har svart «Nei» vet vi ingenting, mens for dem som har svart «Ja» vet vi kun noe om den bruken de har valgt å beskrive kort som fritekstsvaer.

Likevel kan vi si at **KI er i utstrakt bruk både i forbindelse med utdannings- og FoU-relaterte oppgaver**. Eksempelvis er det sånn at dersom samtlige ansatte som ikke deltok i undersøkelsen, hadde deltatt og svart «Nei» angående bruk av KI i forbindelse med FoU-arbeid, så hadde det likevel nærmere 1 av 4 ansatte oppgitt at de bruker KI aktivt i ulike deler av forskningsarbeidet. Det er med andre ord snakk om en betydelig bruk blant våre kolleger selv i det usannsynlige tilfellet hvor undersøkelsen har fanget opp alle de aktive og planlagt aktive brukerne. Som vi kommer til i neste del, så er også bruken svært variert – KI brukes allerede til nær sagt alle de sentrale akademiske oppgavene man har som fast UH-ansatt ved SAM, OsloMet, og ved landets universiteter for øvrig.⁴

Vi kan også si at KI allerede er, eller er i ferd med å bli, tema for både undervisning og forskning, og i et ikke ubetydelig omfang. Tallene forteller så klart lite om KI er et perifert eller sentralt tema i disse sammenhengene, eller om det vies mye eller lite tid og oppmerksomhet. Men vi vet, som vi skal se i neste del, at tematiseringen er variert, og på ulike måter også ofte knyttet direkte til instituttene og utdanningenes profesjonsfelt og faglige kjerneområder.

⁴ I etterkant av ferdigstilling av dette notatet for fakultetsledelsen publiserte Oxford University Press en undersøkelse med 2000 akademikere fra humaniora og samfunnsvitenskap (<https://fdslive.oup.com/www.oup.com/academic/pdf/Researchers-and-AI-survey-findings.pdf> , lest 10.09.2024). Undersøkelsen har en lignende utforming, og funnene minner om dem vi oppgir nedenfor.

Resultater: innholdsanalyse

Nedenfor ser vi på resultatene fra innholdsanalysen. Vi ser først på hva kollegene våre sa om KI som verktøy i utdanning, FoU, og administrasjon. Deretter ser vi på hva de sa om tematisering av KI i undervisning og FoU-prosjekter. Til sist ser vi på de ulike bekymringene de ansatte oppga knyttet til KI-teknologiens frammarsj. I hovedsak ser vi på svarene samlet, uavhengig av hvor respondenten er ansatt. Men der respondentene knyttet svarene til konkrete emner og utdanningsløp, eller pågående prosjekter ved instituttene, tydeliggjør vi dette.

KI som verktøy

I denne delen forteller vi om hva de ansatte har svart angående bruk av KI som verktøy i sitt daglige arbeid. Vi drøfter i hovedsak ikke funnene her, men sparer det til diskusjonskapitlet.

KI som verktøy i utdanningene

Vi spurte ansatte om hvordan de og/eller studentene benytter KI-verktøy i forbindelse med undervisning eller vurdering. De ansatte beskrev en omfattende og variert bruk, både som verktøy i eget arbeid, og som studieverktøy blant studenter. Overordnet bruker de ansatte KI til å utvikle ideer til undervisningsopplegg, og til å produsere tekst. I tillegg viser det seg at bruken er svært variert – ansatte forteller at de bruker KI til:

- **Språkforbedring**, i form av redigering av tekst i skriftlig kommunikasjon til studentene, og redigering av «slides» til PowerPoint presentasjoner.
- **Översetting** av tekst, som f.eks. å översette tekstutdrag til en «slide» i en forelesning.
- **Studentoppgaver**, herunder eksamensoppgaver, arbeidskrav, caser og andre øvings- eller seminaroppgaver (f.eks. quiz og flervalgsoppgaver). I tillegg til å få KI til å lage forslag eller gi ideer til oppgaver, brukes teknologien også som kvalitetssikring – man sjekker om KI klarer eksamensoppgaven eller arbeidskravet.
- **Forelesning**, herunder utforming av manus (førsteutkast eller mer omfattende), pedagogiske metaforer, PowerPoint-presentasjonen (førsteutkast eller mer omfattende), og bilder eller illustrasjoner til enkelte lysbilder.
- **Litteraturoversikt**, enten ved å be KI finne forskningslitteratur, eller ved å be KI gi en oversikt eller et sammendrag av kunnskapen på et felt eller tema.
- **Lek og testing**, ved at KI brukes til mye forskjellig, men i form av nysgjerrig utpröving og uten at resultatet trekkes inn i andre prosesser. En ansatt har f.eks. prøvd å bruke KI til å vurdere/ sensurere en eksamensbesvarelse (den ansatte syntes ikke vurderingen var god).
- **Utvikling av profesjonelle ferdigheter**, knyttet til hhv. undersökende journalistikk ved JM og simulert samtaletrening med barn ved SF (hvor sistnevnte er nært knyttet til pågående forskning ved SF, hvor KI brukes til å utvikle «avatarer» som etterligner barns språkferd).
- **Tilbakemeldinger til studenter**, enten generelt eller begrenset til korte tekster.
- **Avslöring av studenters juks**, basert på en tenkning der studenten regnes som avslört når KI produserer en lignende besvarelse som studentens.

Mange av de ansatte som svarte på undersökelsen fortalte også om SAM-studentenes bruk av KI som studieverktøy. Ifölge ansatte bruker studenter KI til:

- **Oppgaveskriving**, til alt fra arbeidskrav og hjemmeeksamener til MA-oppgaven, samt til å sjekke konsistensen i resonnementer.

- **Studieassistanse**, som vi ikke vet hva innebærer, men som trolig betyr at man bruker KI til å forklare fagbegreper, lage sammendrag av lange og/eller tunge tekster, og til idemyldring (trolig, fordi det minner om de ansattes bruk av KI til forskningsrelaterte oppgaver).

De to vanligst rapporterte undervisnings-relaterte bruksformene blant ansatte var språkforbedring og oversettelse, mens utvikling av profesjonelle ferdigheter, tilbakemelding til studenter og avsløring av juks, var sjeldne (ett og to tilfeller). I mange av tilfellene hvor ansatte beskriver bruk av KI til å utvikle tekster eller oppgaver eller lignende, så trekker de opp et skille mellom det å bruke KI som *støtte* i arbeidet, som omtales som greit, og det å levere noe KI har laget som et *ferdig produkt*, som omtales som ugreit. Det er imidlertid ikke klart om det kun er «rene» KI-produkter som er problematisk, eller om det er en slags terskel for hvor mye de KI-produserte tekstene må endres før de kan regnes som «eget arbeid». Dataene om studentenes bruk er knappere enn dataene om de ansattes bruk, både fordi vi har færre svar som omhandler studentenes pågående bruk, og fordi dette er basert på de ansattes observasjoner. Men det passer godt med hva man ellers leser og hører om studenters KI-bruk.

Samlet sett forteller undersøkelsen oss at KI allerede er ganske tungt inne i mange sider ved ansattes undervisningsarbeid og studentenes studiehverdag. I tillegg forteller undersøkelsen oss at KI allerede brukes til ting som befinner seg i etter hvert kjente juridisk-etiske gråsoner, knyttet til grensene for juks og for at et arbeid skal regnes som ens eget. Selv om dataene ikke lar oss beskrive de forskjellige bruksformene i detalj, så gir de anledning til å identifisere og diskutere noen mindre kjente gråsoner. Det er f.eks. i mange tilfeller uklart om de som bruker KI til å *kvalitetssikre* utkast til eksamensoppgaver ser på det som bra eller dårlig dersom KI svarer godt på oppgaven. Blant dem som spesifiserer, ser det ut til å være et mål at KI skal streve – men man kunne også tenkt motsatt, altså at det er et godt tegn dersom KI klarer å lage fornuftige svar på oppgaven. Det handler også om hvordan vi vil vekke hensynet til å forhindre juks opp mot hensynet til å lage best mulige tester på studentenes læring. Vi kommer tilbake til «gråsonene» i diskusjonen.

KI som verktøy i FoU

Vi spurte også ansatte om hvordan KI-verktøy brukes i forbindelse med FoU-arbeid. Også her beskrev de en omfattende og variert bruk. Ansatte forteller at de bruker KI til:

- **Språkforbedring**, i form av korrektur/språkvask, men også omskriving og omstrukturering av større og mindre deler av egenprodusert tekst.
- **Översetting**, av ulikt slag, som å översette egne tekster til annet språk, eller översette intervjuutdrag eller tekstutdrag fra fremmedspråklige tekster, samt for å sjekke at man har forstått en komplisert tekst.
- **Transkribering**, av kvalitative forskningsintervjuer, som en form for første grovtranskribering.
- **Tekstproduksjon**, som vi skiller fra «språkforbedring» ved at det eksplisitt handler om å få KI til å skape et mer eller – i det fleste tilfeller – mindre ferdig utkast til tekst, herunder forskningsartikler og søknader om forskningsfinansiering. Noen har også brukt KI til å lage sammendrag av/utvikle abstrakt til egne forskningsartikler (med dårlige erfaringer).
- **Litteraturoversikt**, i form av å få korte forklaringer på faglige begrep, automatiske litteratursøk, gi sammendrag av lengre tekster, inkl. monografier (så man slipper lese alt selv), få översikt/tips om teori og teoretikere, samt «screening», systematisering og sammendrag av forskningslitteratur – men hvor flere inkluderer et ledd i prosessen hvor man leser og vurderer bidrag eller deler av bidrag «manuelt».
- **Analyse**, enten helt overordnet til «å analysere data», inkl. «stordata» og «sammenfatting av intervjuer», eller til mer avgrensede operasjoner som å identifisere overordnede tema i et

datasett, få ideer til tematiske koding, bruke KI til å sammenligne/sjekke tematiske koder, utvikle kode til statistikkprogrammer og andre databehandlingsverktøy, utvikle konsistent argumentasjon, eller finne egnede ord til å beskrive noe i data.

- **Idemyldring**, beskrevet som «brainstorming» eller sparring for å komme i gang med skriving eller kreativitet, med KI beskrevet som «diskusjonspartner» i sammenhengen.
- **Utvikling av «avatarer» for ferdighetstrening**, som etterligner barns språkatferd og som kan brukes som verktøy i ferdighetstreningen.⁵
- **Datainnsamling**, i form av ekstraksjon av informasjon, f.eks. fra sosiale medier.

Som for bruk av KI i undervisningssammenheng er det språkforbedring og oversetting som trekkes frem av flest. Nederste kategori, «Datainnsamling», er nevnt av kun én. Bredden i bruken er slående – sett bort fra til ROS-analyser, DPIA, informasjonsskriv/samtykkeskjema, og det å gi og å svare på fagfellevurderinger, er det ikke mange deler av forskningsprosessen hvor KI holdes utenfor av ansatte ved SAM. Og som for bruk til undervisningsoppgaver er det også åpenbart at dagens bruk allerede er godt inne i gråsoner hvor fakultetet (og akademia for øvrig) pr nå mangler klare og gjennomtenkte etiske føringer for god og grei bruk (mer om dette nedenfor).

KI som administrativt verktøy

Til sist spurte vi om ansattes bruk av KI til administrative oppgaver. Dette er kategorien hvor færrest hadde noe å melde, og hvor bredden i bruk var klart minst. Det var også to ansatte som ikke enda hadde begynt å bruke KI til administrative oppgaver, men som takket oss for inspirerende forslag til bruk i spørreundersøkelsen. Den ene skrev «Med stikkordene under dette spørsmålet så skjønner jeg at jeg må lære mer/sette meg mer inn i mulighetene som finnes for administrativ hjelp av KI.». Den andre: «Takk for tipset om oppsummering av innspill fra studenter og revidering av pensumlister. Det skal jeg teste nå». Utover svarene fra disse to, fortalte ansatte at de bruker KI til

- **Språkforbedring og oversettelse**
- **Sensorveiledning**
- **Emnerapporter**, i form av sammendrag av studentenes emneevaluering.
- **Referater**
- **E-poster**

Det var også ansatte som foreslo bruk av KI til utvikling av kursopplegg, læringsutbyttebeskrivelser, og pensumlister, samt til teksting av opptak fra forelesning.

Selv om dataene ikke lar oss beskrive de forskjellige bruksformene i detalj, så gir de anledning til å identifisere og diskutere noen mindre kjente gråsoner. Det er f.eks. i mange tilfeller uklart om de som bruker KI til å *kvalitetssikre* utkast til eksamensoppgaver ser på det som bra eller dårlig dersom KI svarer godt på oppgaven. Blant dem som spesifiserer, ser det ut til å være et mål at KI skal streve – men man kunne også tenkt motsatt, altså at det er et godt tegn dersom KI klarer å lage fornuftige svar på oppgaven. Det handler også om hvordan vi vil vekte hensynet til å forhindre juks opp mot hensynet til å lage best mulige tester på studentenes læring. Vi kommer tilbake til «gråsonene» i diskusjonen.

⁵ Denne forskningen er vanskelig å plassere innenfor rammen av vår undersøkelse: KI brukes som verktøy for å utvikle forskningsbasert programvare som kan brukes i ferdighetstrening, og man forsker også på bruken av programvaren i ferdighetstrening i en rekke profesjoner. Derfor kommer denne med her *og* i kapitlet om KI som tema for forskning.

KI som tema

I denne delen ser vi på hva de ansatte har svart angående tematisering av KI, overfor studentene i undervisning og deretter i FoU.

KI som tema i undervisningen

Overordnet sett kan vi si at KI er tematisert i undervisningen på én av to måter – enten som studieverktøy for studentene (som overlapper en del med funnene for «KI som verktøy i utdanningene»), eller som samfunnsfenomen, knyttet til arbeidslivet generelt og profesjonene SAM utdanner til spesielt.

KI tematiseres som studieverktøy av ansatte ved alle instituttene, og fokus i fritekstsvarene vi har fått er på både muligheter og utfordringer, og de juridisk-etiske grensene for god og grei bruk. Den hyppigst tematiserte bruken er studentenes bruk av KI til utvikling og skriving av tekster. Eksamensbesvarelser (hjemmeeksamen) er mest omtalt, men det informeres, diskuteres og problematiseres også rundt bruk av KI til skriving av andre tekster, som BA-oppgaven, MA-oppgaven, og arbeidskrav. I tillegg tematiseres studentenes bruk av KI til: å sikre at man forstår oppgaveteksten man er gitt; utvikling av forskningsspørsmål til BA og MA-oppgaven; lage litteraturoversikt; koding i statistikkprogrammer; og transkribering av lydopptak. Selv om undersøkelsen ikke forteller hvor ansatte trekker grensen for god og grei KI-bruk i disse tilfellene, er det klart at den trekkes ulikt – det kommer ikke minst fram ved at noen *advarer mot* bruk av KI eller regelrett sier at det *ikke skal brukes* til slik og sånn, mens andre *oppfordrer til* (god) bruk.

Juridisk-etiske dilemmaer som fusk og korrekt referansebruk blir også diskutert (både blant entusiaster og skeptikere), inkludert hvorvidt man må oppgi at KI er brukt, og når KI har laget så mye at en ikke lenger kan si at arbeidet er ens eget. Vi vet ikke hvordan studentene konkret anbefales å tenke om disse problemstillingene, men vi legger til grunn at de får ulike råd og føringer hos ulike lærere, all den tid fakultetet pr nå ikke har utviklet felles føringer. På samme måte er det flere som formidler at studentene må være bevisste på teknologiens svakheter – men, igjen, hva som regnes som *svakheter*, antas å variere. Det som derfor virker klinkende klart, er derfor behovet for å klargjøre og utvikle felles svar på disse spørsmålene – hva kan, og evt. bør, studentene bruke KI til, og hvordan, og hva kan eller bør de ikke bruke det til? Dette kommer vi tilbake til i diskusjonen.

Undersøkelsen viser også at KI blir tematisert som samfunnsfenomen generelt og profesjonsrelevant fenomen spesielt, og at det skjer, i ulikt omfang, ved samtlige institutt på SAM.

På **ABI** tematiseres KI i forbindelse med saksbehandling og offentlig forvaltning, og bruk av KI i Nasjonalbiblioteket og Arkivverket. Juridiske problemstillinger og reguleringer knyttet til KI tematiseres også for både arkiv- og bibliotekstudenter. KI-baserte verktøy og kildekritiske utfordringer knyttet til slike drøftes i emner som handler om informasjonssøk. Foreløpige planer for et revidert masterprogram inkluderer nye emner som blant annet tar opp KI og tillitt til informasjon og offentlige institusjoner og konkrete KI-verktøy brukt til avansert gjenfinning, databehandling og utvikling av systematiske kunnskapsoversikter.

På **HH** er fokus på KIs rolle i samfunnet og fremtidens arbeidsmarked, som en del av en bredere diskusjon om automatiseringens effekter. Videre blir KI's innvirkning på spesifikke fagfelt analysert gjennom fagartikler og diskusjoner. I tillegg blir betydningen av KI for organisering og politikktutforming i offentlig sektor belyst.

På **JM** er det fokus på etikk i bruken av KI som med tanke på innholdsproduksjon og som arbeidsmetoder for journalister. Dette gjelder særlig automatisert tekstproduksjon og hvordan dette kan påvirke profesjoner som lever av tekst, men er også knyttet til manipulasjon av bilder og visuelt

materiale – det foreleses blant annet om hvordan studentene kritisk kan vurdere og verifisere bilder. KI diskuteres også kritisk i lys av teknologiens innvirkning på samfunnet og medienes rolle. I tillegg undervises det i hvordan man kan bruke KI i undersøkende journalistikk. Blant annet lærer studentene ulike metoder for å systematisere store datamengder, og noe koding (Python). Instituttet har også opprettet en videreutdanning i KI-journalistikk, der målet er at studentene skal utvikle ferdigheter innen analyse- og visualiseringsmetoder av store mengder digitale data, samt lære statistikk og koding. EVUen har også et fokus på KI-bruk i undersøkende journalistikk.

På **SF** er KI tematisert i forbindelse med generelle endringer i arbeidslivet, og mer spesifikt knyttet til hvordan KI kan simulere enkle beslutninger i profesjonsutøvelse – noe som i betydelig grad kan utfordre profesjonene. Den nasjonale forskerskolen DIGIT ledes av ansatte på SF som er tilknyttet det fremragende forskningsmiljøet CEDIC. Flere av inn lederne har KI i offentlig sektor som fokus. Det er også planlagt undervisning kommende studieår på SF, hvor KI tematiseres som verktøy og faktor i helse- og sosialfaglig praksis, og mer generelt tilknyttet digitalisering og spørsmål om strukturell makt.

Et generelt inntrykk er at KI allerede er ganske godt inne i undervisningen på SAM, og med et fokus som passer med profesjonene som utdannes ved de ulike instituttene. Et viktig unntak gjelder opplæring i god bruk av KI til akademiske basisoppgaver som en alvorlig blindflekk. Sett i sammenheng med den store og varierte bruken av KI som studie-, undervisnings-, og forskningsverktøy som vi så i forrige kapittel, fremstår dette siste som særlig problematisk.

KI som tema for FoU

Vi spurte også de ansatte om hvorvidt KI er tema for FoU-arbeidet deres, og i så fall på hvilke måter. Fordi svarene vi fikk generelt var ganske knappe, og fordi vi i en del tilfeller vet hva slags prosjekter som er omtalt, har vi supplert med mer informasjon enkelte steder.

På **ABI** forskes det på KI langs både tekniske og mer samfunnsmessige akser. Som omtalt i delen om bruk av KI som FoU-verktøy, er det flere som bruker KI til indeksering og analyse av store datamengder. Én av forskergruppene (MetalInfo) har anskaffet og installert dataprosessorer (GPUer) på en lokal server og er i gang med eksperimentering med ulike KI-modeller til analyser av bilder og tekst. Instituttet er også sammen med JM og TKD involvert i en infrastruktursøknad til NFR knyttet til KI-basert indeksering og analyser av landets største bildesamling ved Nasjonalbiblioteket. Andre prosjekter undersøker KI-baserte gjenfinningssystemer (som Keenious) fra et brukerperspektiv, og det planlegges prosjekter på hvordan automatisering og KI påvirker arbeids- og informasjonspraksiser i jobbsammenhenger. I tillegg er ansatte involvert i NFR-søknader som omhandler hvordan algoritmer og KI påvirker bruk og tilgang til kulturprodukter gjennom strømmetjenestenes grensesnitt, og det foreligger ambisjoner om å utvikle prosjekter som diskuterer konsekvensene av KI-generert informasjon og tilliten til offentlige informasjonskanaler.

På **HH** er forskere i gang med å bruke NLP for å utforske forbindelser mellom språk og psykologi, nærmere bestemt hvordan språket reflekterer psykologiske prosesser og tilstander. Andre planlegger forskning på rettslige aspekter ved bruk av KI, inkludert personvern, cybersikkerhet, og tverrfaglige perspektiv på hvordan KI påvirker høyere utdanning. Det er også sendt en søknad til NFR sammen med Arbeidsforskningsinstituttet (AFI), hvor man foreslår å undersøke kompleksiteten og utfordringene ved å regulere KI innenfor høyere utdanning.

Institutt **JM** har et større prosjekt med ulike finansieringskilder knyttet til undersøkende journalistikk. Det er i hovedsak et praktisk prosjekt med ulike samarbeidspartnere (Stiftelsen for kritisk undersøkende journalistikk, ulike redaksjoner og AI-Lab) der formålet er å forstå KI-basert journalistikk og utvikle nye metoder for å gjøre bruken tryggere, mer transparent og mer effektiv.

Målet er å forske på og utvikle ferdighetssiden og nye forskningsmetoder. Instituttet har også fått sentrale strategimidler fra OsloMet til dette arbeidet. I tillegg er det planer om prosjekter innenfor det visuelle, blant annet om publikumsbevissthet om KI-generert visuelt materiale og profesjonsforskning knyttet til bruken av KI i produksjon av journalistisk innhold og om dette driver frem nye profesjonstyper. Som nevnt over er JM involvert i en NFR-søknad sammen med ABI og TKD.

På **SF** drives forskning på og utvikling av KI-baserte avatarer som brukes for å fremme evidensbasert og effektiv ferdighetstrening (samtaletrening), enten tekstbasert eller i VR. Dette er forskning og utvikling av KI som er i den internasjonale forskningsfronten, og teknologien vil kunne brukes til samtaletrening i alle typer av menneskebehandlende profesjoner. Det er også flere innsendte og pågående søknader om forskningsmidler for å utvikle KI, blant annet et prosjekt som vil bruke KI i utviklingen av tjenester for eldre mennesker - omsorgs- og arbeidsrelaterede tjenester - på en slik måte at de minimerer aldersdiskriminering/ «ageism». I tillegg ledes det overnevnte forskningsmiljøet CEDIC (og forskerskolen DIGIT) fra SF, som inkluderer og fostrer forskning der KI både er metodeverktøy og studieobjekt.

Når det gjelder FoU er det altså samlet sett mye på vei, særlig tatt i betraktning at vi snakker om ganske «ferske» samfunnsendringer og den kjente (og kjære) tregheten som ofte kjennetegner akademia. Med mindre SAM av strategiske årsaker mener det er behov for bestemte typer av forskning, så ville vi vært mindre bekymret for progresjonen her. Men selv om det er en del egen forskning på vei, så kan SAM med fordel oppfordre til mer bruk av eksisterende og kommende forskning av relevans for våre utdanninger – det kan med andre ord vurderes grep for å få relevant forskningsbasert kunnskap om KI inn på pensum og i forelesninger og seminarundervisning i flere av våre utdanninger.

Utfordringer knyttet til bruk av KI i UH-sektoren

Vi avsluttet undersøkelsen med å spørre om hva de ansatte så som særlige utfordringer med bruk av KI i UH-sektoren, for den individuelle ansatte, institutt og fagfelt, fakultetet, eller for UH-sektoren generelt. Vi unnlot å spørre om «muligheter» fordi vi anså det som dekket (av spørsmålene der ansatte både kunne fortelle hva KI per i dag ble brukt til, og hva de tenkte de ville bruke det til fremover). Så hva er de store utfordringene, slik de ansatte ser det?

Vi oppgir først hva vi lærte fra rundspørringen, og så ser vi på hva vi lærte fra arbeidsmøtet og hvordan det passer med eller skiller seg fra innspillene fra førstnevnte.

Fusk og plagiat

Flere av utfordringene som de ansatte tar opp henger naturligvis tett sammen med hva de har fortalt oss om bruken. Mange beskriver utfordringer knyttet til spørsmål om **fusk og plagiat**: for studentene, i tilknytning hjemmeeksamen, masteroppgaven og arbeidskrav; og for de ansatte, i forbindelse med forskning (og da tilknyttet alt fra bruk av KI til delprosesser i skriving og analyse, til ren fabrikkering av forskningsdata). Enkelte løftet også frem hvordan utfordringene henger sammen med spørsmål om hvor man trekker grensen, mellom inspirasjon og fusk, og mellom eget og andres arbeid (terskelen for når man ikke lenger bør regnes som opphavsperson eller si at arbeidet er ens eget). Et relatert poeng som tas opp av et lite antall ansatte, er at det også blir **vanskeligere å vurdere studentenes læring og kompetanse i studieløpet**, som betyr også betyr at arbeidsgivere kanskje ikke kan stole på om jobbsøker er kvalifisert.

Manglende retningslinjer og opplæring

Mange etterspør også **klare retningslinjer vedrørende studentenes bruk av KI** (men ikke de ansattes bruk!). Vi innbiller oss at dette ønsket er mer eller mindre allment (alle vil ha klare kjørerregler/regler),

selv om det nok er uenighet om hvor retningslinjene skal komme fra og hva slags form de skal ha (fra forbud til anbefalinger for god og grei bruk). Merk at vi skriver «retningslinjer», selv om flere også påpeker behov for god opplæring/ veiledning, da det likevel er rimelig å tenke at man trenger noen felles rettesnorer for veiledningsarbeidet. Vi har også sortert kommentarer om bekymringer knyttet til personvern under dette punktet: det er behov for retningslinjer for hvordan man overholder GDPR ved bruk av KI i ulike sammenhenger (som vi kommer tilbake til i diskusjonen). En relatert bekymring, som riktignok bare reises av et fåtall, er at uten god opplæring og klare retningslinjer så vil KI kunne forsterke forskjellen mellom sterke og svake studenter, f.eks. ved at bare de sterkeste studentene vil klare å være selvstendige nok i møte med KI sine "forslag", også de mindre gode.

Vi legger til at mens arbeidet med dette notatet har pågått så har OsloMet sentralt lansert generelle retningslinjer for bruk av KI.⁶ Men disse retningslinjene er så generelle at selv folk som er av diametralt motsatt oppfatning hva gjelder god og grei KI-bruk vil kunne enes om de samme reglene. Man kan f.eks. slutte opp om at KI skal brukes «på en etisk, trygg og ansvarlig måte», men samtidig være helt uenige om hva man regner som etisk, trygg eller ansvarlig bruk.

KI undergraver læring og utvikling av grunnleggende ferdigheter

Det er også mange som beskriver farer ved at bruk av KI til alt det ansatte og studenter ved SAM allerede bruker KI til, vil føre til at **brukerne blir dårligere til å tenke, lese og skrive**. KI forstås med andre ord som **en trussel mot** en av de tradisjonelt sentrale og generelle læringsmålene med høyere utdanning, som vi med Immanuel Kant kan beskrive som **oppøvelse av evnen til å gjøre bruk av vår egen fornuft, og dermed også evnen til kritisk tenkning**.⁷ (Man kunne kanskje i samme vending også vært bekymret for *utøvelsen* av egen fornuft). Bekymringen knyttes sterkest til studentene, men ikke utelukkende. Det er også bekymring for at det **går ut over evnen til å tenke lange og grundige tanker**, eller til å tenke «godt og grundig», og evnen og viljen til å *bli værende* ved en vanskelig problemstilling. Noen peker også på at KI hemmer refleksjon og utvikling av språkferdigheter, mens andre peker på lekenhet og kreativitet, som de tror man blir dårligere på om mer og mer av ideutviklingen overlates til programvare. Andre er bekymret for at vi blir sløve av å støtte oss på maskiner, og at vi skal miste vår motivasjon for å lære. Og noen beskriver samme bekymring, men fra motsatt ende – at **det kan være mange ting KI gjør raskere og/eller bedre enn oss, men som vi likevel, som individer og samfunn, har godt av å gjøre selv**. Som en slags parallell til disse bekymringene er det også noen som peker på at KI gjerne kan brukes som avlastning til «ensformige og tunge oppgaver», som kan leses som en presisering om at det nettopp ikke skal brukes som avlastning for tenkning og kreativt arbeid.

Utover det generelle sees KI også som en trussel mot mer bestemte læringsmål, dersom studentene ikke eller i liten grad selv trenger å lese eller skrive for å komme seg gjennom studiet. Det er fare for at de ikke lærer – eller for at de lærer å bruk KI, men ikke det de ifølge programplanen skulle lære – når KI «gjør jobben for dem». I sum oppfattes KI som lærings- og arbeidsverktøy som lettvinnt, og **bekymringen er at man velger det lettvinnte fremfor det viktige**. En påpeker også at det blir dårligere profesjonsutøvere av sånt.

KI skaper mye og dårlig forskning, og truer tilliten til vitenskap

Dessuten er det mange som beskriver **bekymring for kvaliteten på og tilliten til forskning**. Bruk av KI kan, eller antas å være i ferd med å, resultere i *dårlig* (pregløs, kjedelig, metodisk svak og/eller intetsigende) og/eller *uredelig* forskning, i *stort* volum. Vi har hatt saker i den brede offentligheten som handler om falske forskningsartikler og røvertidsskrift, og nå er det også kommet saker der

⁶ Se <https://student.oslomet.no/policy-kunstig-intelligens>, lest 14.05.2024.

⁷ Her parafraiserer vi fra essayet *An Answer to the Question: "What is Enlightenment?"* fra 1784.

forskningsartikler er publisert med åpenbare tegn på KI-bruk i skrivningen (f.eks. ved at ChatGPT sitt svar på tekstbestillingen – «Certainly, here is a possible introduction for your topic» osv. – fortsatt står i teksten.⁸ I ytterste konsekvens, skriver flere, utgjør KI **en trussel mot faglig integritet, og vitenskapenes tillit og rolle i storsamfunnet.**

I tett forbindelse med bekymring for at det produseres mer og dårligere, er det også noen som uttrykker **bekymring for at bruk av KI skal føre til forhøyede produktivitetskrav** til UH-ansatte. Dette beskrives også som en situasjon der handlinger som er individuelt rasjonell blir kollektivt irrasjonelle – hvor den enkelte prøver å bedre en presset arbeidssituasjon ved å bruke KI til å få opp tempoet på forskning og skrivning, men hvor summen av at alle gjør det samme er at ingen er blitt, relativt sett, raskere enn røkla; i stedet jobber alle forttere for å levere normert. Dette er en type systemperspektiv vi kommer tilbake i diskusjonen.

Falske nyheter og kildekritikk

Et tema som er nært forbundet med juks og plagiering er **«falske nyheter» og nye krav til kildekritikk** både i og utenfor academia. KI gjør det vanskeligere å skille ekte fra falsk, god informasjon fra dårlig informasjon, og gjør det generelt vanskeligere å utøve god kildekritikk – særlig når falske artikler og nyhetssaker publiseres i kanaler som man eller har hatt tillit til. Fra ABI ble det særlig pekt på utfordringer med å sikre at borgerne har gode kanaler til offentligheten – som er betingelse for et opplyst ordskifte og et velfungerende demokratisk maktkretsløp. Fra JM ble det løftet frem utfordringer knyttet til verifisering av tekst og bilder, og at dette kan gi både tekst og bilde (og trolig film) en ny rolle i samfunnet.

Det ble også løftet frem utfordringer knyttet til å være kritisk overfor KI når man selv er bruker – **at vi stoler på KI når vi ikke bør.** Her fyller vi noe ut, men en fundamental utfordring ved bruk av KI til oppgaver hvor KI skal gi deg opplysninger du ikke hadde fra før, er at du ikke vet om KI gir deg rette eller gale opplysninger. Dette kommer vi tilbake til i diskusjonen, men når KI brukes for å få litteraturoversikt eller til å gi sammendrag eller forklaringer av komplekse begrep, så må man vite svaret for å kunne avdekke om Kis resultat er sakssvarende.

Hindringer for god bruk av KI

Det er også flere som skriver mer fra en posisjon som ikke bekymrer seg for KI som sådan, men for **strukturelle hindringer for god bruk av KI.** Som vi kommer tilbake til i diskusjonen, er det imidlertid uklart hva som ligger i at folk skal «lære å bruke» KI – det kan variere mellom ren teknisk operasjon til former for klokskap og visdom. KI omtales som **et uutnyttet potensial**, og noen skriver bekymret om «fare for at potensialet er underutnyttet» (f.eks. til automatisert vurdering av studenters arbeid, som kan veksles inn i mer tid til veiledning av studenter, «til selvstendig tenkning»). Det pekes også på **behov for ressurser til omstilling og KI-relatert kompetanseheving** hos de ansatte – ansatte som skal lære noe nytt må få tid til det. Dette forstås blant annet som **et ledelsesproblem** (uansett om vi tolker det som «ledelsen sitt problem» eller «noe det er utfordrende for ledere å løse»). Det er også mange som peker på bestemte former for kunnskapsmangel det må bøtes på, for å realisere god og grei bruk av KI. Ansatte trenger **kompetanse om ulike KI-teknologier og hva de kan brukes til** – som vi kommer tilbake til i diskusjonen er det viktig å komme forbi generell snakk om KI når vi skal komme videre i tenkningen om god og grei bruk. Dette omtales også som **«KI-literacy»**, og det skilles mellom generell og fagspesifikk «KI-literacy», som vi tolker som uttrykk for at kunnskap om hvordan ulik KI-teknologi kan brukes fornuftig vil ha både generelle og fagspesifikke komponenter. En del av kunnskapen må

⁸ Se <https://www.khrono.no/forskere-reagerer-pa-at-denne-artikkelen-slapp-gjennom-er-det-mulig/857948>, lest 22.05.2024.

også bidra til bevissthet om hva man vet og ikke vet om grunnlaget for det KI-teknologien lager for oss.

Det pekes også på det heseblesende utviklingstempoet vi står oppe i nå, som utgjør en ekstra belastning både for den enkelte (i form av stress), og en stor utfordring for fakultetet dersom man skal legge strukturelt bedre til rette for kompetanseheving – **det er fare for at utviklingstempoet betyr «at bruken alltid er umoden».**

Et fåtall avviser utfordringer, og peker kun på mulighetene KI byr på i positiv forstand, generelt og til konkrete oppgaver som analyser av store datasett. Som én skriver, «Ingen utfordringer, bare muligheter». Motsatt er det noen få som uttrykker sterk skeptisk til om KI tilfører noe nytt eller kan brukes til noe fornuftig overhodet.

Ymse

En del av utfordringene som beskrives er mer variert, og nevnes av én til tre ansatte, ble sortert under kategorien «ymse» (som ikke betyr at de er uviktige for fakultetets strategiarbeid!). Her løfter vi frem to: 1) KI byr på utfordringer knyttet til **forvaltning av opphavsretten**; og 2) Det er **utfordrende å forberede studentene godt på yrkeslivet** når de antatt store konsekvensene av KI foreløpig er ukjente. Som vi kommer tilbake til i diskusjonen, er det slående at så få reiste denne sistnevnte bekymringen.

Vi biter oss merke i at det jevnt over er mest bekymring knyttet til bruken av KI som arbeids- og studieverktøy, altså i den akademiske konteksten. Det er med andre ord relativt sett få som uttrykker bekymring for profesjonene vi utdanner til på SAM. Som vi kommer tilbake til i diskusjonen er også det et mulig innsatsområde for fakultetsstrategien – er det viktige sider ved profesjonenes arbeid eller vilkår som på interessante, lovende eller problematiske måter kan, eller ligger an til, å påvirkes av KI? Hvis ja, er dette noe vi må ha oppmerksomhet på både mtp undervisning og forskning.

Fra arbeidsmøtet 16. april

De overordnede bekymringene fra rundspørringen speiles i innspillene fra arbeidsmøtet, men møtet tilførte også noen nye innspill.

I likhet med i rundspørringen ble det diskutert **behov for retningslinjer og kompetanseheving** for ansatte. Deriblant diskuterte gruppa selve utformingen av retningslinjer for KI-bruk. Utkastet til nye retningslinjer for universitetet ble diskutert, og omtalt som svært generelle. Samtidig ble det reist tvil om hvorvidt presisering som sådan gir nyttigere retningslinjer. Men kanskje kan man lage mer konkrete retningslinjer knyttet til bestemte arbeidsoppgaver, som utvikling av kunnskapsoppsummeringer, og knyttet til konkrete KI-teknologier, som bruk av språktjenester som ChatGPT? Og kanskje man kan jobbe med å utvikle felles forståelser av mer generelle prinsipper ved å diskutere konkrete eksempler opp mot prinsippene, og gjerne i form av vignetter som kan justeres på jakt etter en rimelig omforent grense, mellom god og dårlig og grei og ugrei KI-bruk.

Ulik behandling av studentene ble løftet frem som et problem, knyttet opp mot studentenes rettigheter. F.eks. dersom man i ett emne får lov å bruke KI-verktøy, men ikke i et annet. Dette ser vi i sammenheng med overnevnte ønsker om retningslinjer og kompetanseheving for de ansatte. Samtidig ble det diskutert at det også kan være problematisk å likebehandle studenter som går på ulike utdanninger der det kan være faglige og pedagogiske grunner for å gjøre ting på bestemte måter, også mtp bruk av KI. Derfor er bekymringen å forstå som knyttet til problematisk forskjellsbehandling, herunder forstått som vilkårlig eller ubegrunnet forskjellsbehandling, innenfor de enkelte utdanningsløpene.

Dette ansporet igjen en diskusjon, om **hvorvidt studentene lærer av å bruke KI, og etter hvert om hva studentene lærer, og om det er kunnskap eller ferdigheter som svarer til læringsutbyttebeskrivelsene** i program- og emneplanene. Det ble sagt omtrent slik: «De går på studiet for å lære noe, og ikke bare å bruke KI» – og vi kan derfor spørre om de faktisk «lærer det de skal» når de bruker KI. Men vi kan også tenke, ble det innvendt, at fremtidens studenter og profesjonsutøvere må og skal kunne løse oppgaver ved hjelp av ulike KI-verktøy, og dermed kan man også spørre seg om dagens læringsutbyttebeskrivelser er egnet for den nære fremtids profesjonsutdanninger. Diskusjonen reiste altså både empiriske spørsmål om hva studentene lærer, og utdanningspolitiske spørsmål om hva de bør lære. Og dersom KI bør få konsekvenser for læringsmålene, kommer også spørsmålet om hva slags pedagogikk som er best egnet for å lære studenter å *lære med KI*.

En annen diskusjon gjaldt **plagiat og spørsmålet om hvorvidt, når og hvordan man skal oppgi at KI er brukt**. Noen mente at man alltid skal oppgi dette fordi det er viktig informasjon om hvordan et arbeid er blitt til, mens andre mente at KI heller bør sees på som en integrert del av arbeidet som vi ikke trenger å oppgi – litt på samme måte som at man ikke oppgir at man har brukt MS Word med stavekontroll.

En annen side ved diskusjonen om plagiat gjaldt det praktiske spørsmålet om **hvor mye tid UH-ansatte burde bruke på å avdekke plagiat og fusk ved bruk av KI**. Uten at det besvarte spørsmålet, ble det bemerket at all tid man bruker på slike ting jo er tid man kunne brukt på å heve kvaliteten på utdanning, forskning og utvikling, som viser at det hele er et prioriteringsspørsmål – og et SAMs KI-strategi bør bidra til å veilede.

Utredningsutvalget hadde på forhånd valgt ut bestemte ting vi ønsket å få innspill på, og et hovedtema gjaldt hva SAM, som *samfunnsvitenskapelig* fakultet, burde være spesielt gode på, og som kan bidra til å gi retning også når det kommer til satsing på KI som verktøy og tema. Med andre ord, er det noen **særskilte satsinger som bør inngå i strategien**, eller bør strategien være å støtte oppunder hva enn som vokser frem «nedenfra» av initiativ og retninger? Svaret som ble gitt, var «**Tillit og demokrati**». Hva kan vi stole på, hvem, og under hvilke forutsetninger? Dette ble beskrevet som sentrale spørsmål for alle SAMs institutt, kanskje til og med som et eget «exphil» for studenter på SAM. Tanken er da at spørsmål om KI-teknologier passer fint innunder dette overordnede tema, på måter som samfunnsvitere kan bidra til å belyse.

Vi legger til at det også ble påpekt at dette er et ansvar som må bæres av emneansvarlige hver og én: alle må tenke gjennom hva KI kan ha å si for tematikken i sitt emne – altså ikke bare som mer eller mindre egnet arbeidsverktøy for student eller lærer, men som samfunnsfenomen, og for det aktuelle samfunnsområdet som emnet omhandler.

Det ble også kort diskutert behovet for å fortsette med **utvikling av stillinger for å knytte teknologer til SAM**. I tillegg ble det etterspurt **en mer systematisk måte å dele på kunnskap og initiativer internt på SAM** – en form for fellesarena eller portal for dem som vil lære av hverandre mtp KI-initiativ på fakultetet. Til sist kan vi nevne at **ansatte i administrasjonen**, som vi har utelukket fra vår rundspørring, gjerne **ønsket å tas med eller å få gjennomført en lignende undersøkelse hos seg**.

Diskusjon av utfordringer og muligheter

Her diskuterer vi utfordringer og muligheter som bruk av KI byr på, og hvordan dette bør hensyntas i SAMs KI-strategi. Vi starter med utfordringene og skiller mellom *utfordrende «gråsoner» for brukerne* og *utfordringer av mer systemisk art*. Deretter ser vi kort på *mulighetene KI utgjør* for fakultetet.

Gråsoner

Vi har snakket om «gråsoner» i gjennomgangen av resultatene. Med «gråsoner» sikter vi til brukstilfeller der det er uklart eller uavklart hva som utgjør god og grei bruk av KI. Utredningen har gitt hint og pek om flere slike brukstilfeller, og her deler vi dem inn i to hovedtyper, knyttet til normer for 1) god bruk av KI (praktiske gråsoner) og 2) grei bruk av KI (juridisk-etiske gråsoner). Med 'god bruk' sikter vi altså til hva som er effektiv og rasjonell bruk av teknologien. Med 'grei bruk' sikter vi til hva som er ansett som i tråd med lov, retningslinjer og sedvane.

God bruk av KI: praktiske gråsoner

Det er særlig to typer av praktiske gråsoner som brukerne ofte står overfor, og som en KI-strategi bør ta sikte på å bidra til å tydeliggjøre.

«Er KI til å stole på?» Den første handler om at man må vite svaret for å vite om KI svarer feil, og derfor at bruk av KI i mange tilfeller handler om tillit. Når man bruker KI-teknologier som ChatGPT til å produsere sammendrag av vanskelige artikler, lange bøker eller store kunnskapsfelt, eller å forklare et komplisert begrep på en enkel måte – altså *oppgaver hvor gyldighet er en faktor* – er det forbundet usikkerhet med resultatene man får. Gir KI gyldige svar? Den praktiske gråsonen består i at man ikke vet om svaret KI gir er riktig eller feil, med mindre man allerede vet hva som er rett (i alle fall omtrentlig). Samtidig kan vi legge til grunn at de som spør KI som oftest ikke besitter slik kunnskap. De som spør og ikke vet, står derfor overfor et valg – å ettergå eller stole på svaret. En del av dem som svarte på rundspørringen understreket riktignok at svarene de får behandles som et utkast som man jobber videre med, men med mindre man vil gjøre det samme arbeidet som KI har gjort (og dermed risikere at «vinninga går opp i spinninga»), består det grunnleggende problemet også i disse tilfellene.⁹ Når man bruker KI for å få litteraturoversikt eller til å gi sammendrag eller forklaringer av komplekse begrep, så setter man derfor nødvendigvis sin lit til at algoritmen (og data!) er til å stole på.

Kan man stole på KI? *Bør* man stole på KI? Tilliten må i så fall enten bero på kunnskap om teknologienes virkemåter og virkninger – som per i dag må vurderes som mangelfull, på SAM og i sektoren og samfunnet ellers – eller på tillit til organisasjonene som lager løsningene (som Open AI og Google) eller som gir sin tilslutning til at de brukes (som OsloMet gjør når KI-verktøyet Keenious ligger på universitetets hjemmesider).¹⁰ I vår sektor, hvor gyldighet og redelighet er sentrale verdier for virksomheten, er det verdt å reise spørsmålet om KI er tilliten verdig – og kanskje også spørsmålet om å velge avholdenhet frem til mer solid dokumentasjon er bragt på banen. Fakultetets strategi bør baseres på en grundig diskusjon av disse spørsmålene, og fakultetet bør på en eller annen måte sikre at ansatte er bevisst disse problemstillingene.

«Forstår vi hva vi bruker KI til?» Den andre praktiske gråsonen handler om bruk av KI til «kvalitetssikring», og nærmere bestemt forholdet mellom det KI produserer på den ene siden og vår vurdering av relevans og kvalitet på den andre. Et viktig skille mellom denne typen utfordring og den

⁹ Så kan man innvende at man som menneske også gjør feil, men den diskusjonen forfølger vi ikke her.

¹⁰ <https://www.oslomet.no/ub/ki-verktoy>, lest 07.05.2024. Det er interessant å spørre seg hva universitetets tilsynelatende støtte til bruk av Keenious har å si for brukernes forståelse av god og grei bruk.

forrige, er at denne ikke ble løftet frem av respondentene i rundspørringen – det er vi som problematiserer respondentenes (beskrevne) bruk, fordi vi mener den ofte hviler på tatt-for-gitte ideer om hva vi lærer av «kvalitetssikringen». Eksemplet vi allerede har sett på gjaldt bruk av KI til å kvalitetssikre utkast til hjemmeeksamensoppgaver, hvor læreren sjekker hvor godt KI løser oppgaven. Den praktiske tolkningsutfordringen består i at det i prinsippet er uavklart om det lover godt eller dårlig at KI svarte godt (eller dårlig) på oppgaven. Der vi fikk mer fylldige svar virket det å bli tatt som et dårlig tegn at KI svarte bra – men vurderingen avhenger av hva slags «kvalitet» man vil sikre. Vår antakelse er at man i disse tilfellene vil sikre oppgaven mot juks – å bruke KI skal ikke være en farbar vei. Kvalitet betyr da «vanskelig å løse ved juks». Men man kunne tenkt motsatt, og vært mer opptatt av at oppgaven er lesbar og forståelig, og er en god test på den læringen man ønsker å måle. Da er kvalitet «et godt mål på studentenes læring i emnet». Svaret er ikke gitt, og motstridende svar kan forsvares av rasjonelle grunner. Poenget er altså at man må ha et reflektert forhold til hva slags kvalitet man legger vekt når man bruker KI til å kvalitetssikre – kvalitetssikring er alltid et verdivalg i forkledning.

Et annet eksempel er bruk av KI til å «sjekke» kodingen av data i kvalitative analyser. I slike tilfeller er det uklart om det lover dårlig at KI produserer noe som *avviker* fra din analyse, eller om det lover dårlig dersom KI produserer noe som *overlapper sterkt*. Inntrykket fra rundspørringen er at man har tenkt «overlapp lover godt», men motsatt tolkning er mulig og begge posisjoner kan begrunnes rasjonelt. Trolig avhenger tolkningen også av hva slags vitenskapsteoretisk posisjon man tar (for en empirist er nok overlappet mer betryggende enn for en sosialkonstruksjonist eller en hermeneutiker).¹¹

Begge eksemplene har samme struktur: Brukerne virker å ha et ureflektert forhold til det prinsipielt åpne og uavklarte forholdet mellom hva KI genererer og vår vurdering av kvalitet. Selv om det er langt fra gitt at man skal fremme enhetlig tenkning rundt disse spørsmålene, så mener vi det er en opplagt oppgave for fakultetet å inkludere dem i et kompetansefremmende opplegg for UH-ansatte på SAM. Vi trenger ikke være enige om tolkningen, men vi må være bevisst det store rommet for tolkning.

Grei bruk av KI: juridisk-etiske gråsoner.

Det er også (minst) to juridisk-etiske gråsoner som brukerne ofte står overfor, og som en KI-strategi bør ta sikte på å bidra til å tydeliggjøre.

«Er det ditt arbeid?» Den første juridisk-etiske gråsonen er gammel, men med KI kommer den i ny innpakning – det dreier seg om vurderinger og forvaltning av opphavsretten, og mer generelt om spørsmålet om et arbeid er ens eget. **Bruk av KI skaper nye utfordringer tilknyttet opphavsretten.** Det er uklart hvor grensen går for å bestemme hvor mye av arbeidet som kan være gjort av KI før det ikke lenger kan regnes som selvstendig og ditt (om du nå er student eller UH-ansatt). Og når opphavsretten er avklart er den samtidig truet dersom KI bruker ditt arbeid i genereringen av tekst, og den som brukte KI dermed bruker dine ideer uten henvisning til deg (fordi KI-verktøy ofte ikke oppgir kilde). Selv om problemstillingen ikke grunnleggende sett er ny, så skaper nye KI-praksiser behov for ny og klar tenkning.

Vi tror ikke SAM skal ta sikte på å utvikle egne retningslinjer for opphavsrett, men holde de ansatte oppdatert på relevante endringer i tenkning og praksis. Likevel, når det kommer til forskyvninger i grensene for hva som regnes som ens eget arbeid (om en er student eller ansatt), er det grunn til å tro at grensene også til dels vil være fagspesifikke. SAM bør derfor bidra til at hvert institutt og

¹¹ Skillene er analytiske, og det finnes grensetilfeller, som når man bruker KI som *kvalitetssikring* for å se om man har *rett* forståelse av et komplisert begrep eller et langt resonnement.

fagmiljø har egne diskusjoner knyttet til grensen for når noe bør og ikke lenger bør regnes som et selvstendig arbeid pga omfanget og typen av KI bruk. Det kan også være nyttig hvis universitetet tilbyr veiledning eller assistanse angående de juridiske konsekvensene av å utvikle arbeid i samarbeid med KI. Dette er spesielt relevant hvis bruk av KI-verktøy bidrar til å skape produkter eller programmer som kan kommersialiseres, og reiser juridiske spørsmål knyttet til patenter, opphavsrett og andre immaterielle rettigheter. SAM kan oppfordre til at dette sikres ordninger for veiledning og assistanse på universitetsnivå.

«Er personvernet ivaretatt?» Den andre juridisk-etiske gråsonen handler om at bruk av KI og personvernet (se også f.eks. Khowaja et al., 2024). For eksempel etter personvernerklaring til Sikt KI-chat, er det "sluttbruker som kontrollerer hvilke data som skrives inn i Tjenesten. Sluttbruker må likevel følge instruks fra sin egen institusjon/virksomhet om hva han eller hun har tillatelse til å bruke Tjenesten til." Det er derfor behov for retningslinjer for hvordan man overholder personvern generelt og GDPR spesielt ved bruk av KI, f.eks. ved å ha et forbud mot å gi personopplysninger som del av en ledetekst (prompt). SAM bør nok ikke kjøre sitt eget løp her, men frem til generelle retningslinjer foreligger bør OsloMet også drøfte hvorvidt ansatte og studenter som vil bruke KI i undervisnings- eller forskningsaktivitet, skal/må bruke KI-løsninger som til enhver tid har godkjent. Det bør også gis opplæring i hva studenter og ansatte må være obs på når det gjelder personvern og personvernregelverket, som å være bevisst hva som deles av data når man «instruerer» en KI som Sikt KI-chat eller Chat GTP.

Systemutfordringer

Som samfunnsvitenskapelig fakultet bør SAM levere på systemperspektiver på KI-bruk til hele universitetet/ sektoren. Her løfter vi frem to systemutfordringer som ble antydnet i innspillene vi har fått, og som vi har tenkt videre på i arbeidet med programnotatet. Dessuten løfter vi selv frem en utfordring – som ikke ble nevnt, men bør med – nemlig hensynet til energibehovet som utbredt bruk av KI medfører. Vi tenker på utfordringene som systemutfordringer fordi de handler om hvordan individuelle valg og endringer i deler av et system forplanter seg som skadelige ringvirkninger i systemet som helhet.

«Kappløp mot avgrunnen?»

Samfunnsvitenskapene er fulle av artikler og bøker som diskuterer typer av kollektive handlingsproblem eller situasjoner hvor det som er rasjonelt for den enkelte er irrasjonelt for kollektivet. Blant dem finner vi «race to the bottom», eller «kappløpet mot avgrunnen», som i utgangspunktet beskriver en situasjon der bedrifter lempet på alt av kvalitetsstandarder i produksjonen, arbeidsmiljøhensyn og ansattlønninger for å hevde seg i konkurransen mot andre bedrifter som gjør det samme, og hvor til slutt ingen vinner konkurransen, men mye går tapt i forsøket.¹² Det er verdt å diskutere om bruk av KI kan lede til lignende kappløp mot avgrunnen i vår sektor.

En måte denne systemutfordringen ble antydnet i innspillene var i forbindelse med bekymringen for at vi bruker KI som en «lettvin» løsning for å produsere *mer og dårligere* forskning – noen var også bekymret for at man med det svekker tilliten til sektoren, og andre var bekymret for at muligheten til å produsere raskere med KI ville føre til forhøyede produktivitetskrav til UH-ansatte. Bekymringen er altså at når den enkelte velger «lettvin» løsninger på tunge arbeidsoppgaver, så vil det kunne få uheldige systemvirkninger i form av dårligere forskning, redusert tillit til academia og forverrede arbeidsvilkår for de ansatte. KI ble også omtalt som en «effektiv» løsning på utfordringer i

¹² <https://dictionary.cambridge.org/dictionary/english/race-to-the-bottom> , lest 08.05.2024.

undervisning. Eksempelvis var det én som fortalte at hen ofte henviste studentene til KI som en form for hjelpelærer, da vedkommende ikke hadde tid til å veilede alle studentene som ba om hjelp. *Spør KI, så får du hjelp raskere*. Slik praksis kan problematiseres med henvisning til «gråsonene» vi diskuterte over, da det ikke gitt at veiledningen KI gir holder mål. Samtidig er det viktig å understreke det rasjonelle i lærerens tenkning – det er en måte å håndtere arbeidsmengden på. Problemet vi ønsker å belyse her er likevel i hvert tilfelle de *systemiske* konsekvensene av summen av den enkeltes individuelt rasjonelle forsøk på å bedre en presset arbeidssituasjon.

Vi mener det er grunn til å følge med på tegn til utvikling av kappløp-tendenser på SAM. En av respondentene sa det sånn: «Vi må bruke KI for å rekke alle oppgavene vi skal nå over i samfunnet». Når KI fremstilles, slik flere av respondentene gjorde det, som en «effektiv» løsning på krevende arbeidsoppgaver og stort arbeidspress, så står vi i fare for å overse det underliggende problemet – nemlig at vi i utgangspunktet, før KI, hadde for mange oppgaver til å kunne løse dem skikkelig. Å løse problemet individuelt ved bruk av KI er å tilsløre problemets strukturelle karakter og årsak. I et slikt systemisk perspektiv fremstår ikke KI nødvendigvis som en etterlengtet løsning på faglige utfordringer, men også som et verktøy for å skru opp tempo i en organisasjon.

Merk at vi ved å reise denne systemutfordringen overhodet ikke fordømmer KI-bruk som letter arbeidsbyrder eller avviser KI-teknologier som sådan. KI er åpenbart bra for UH-sektoren dersom læringsutbyttet av utdanning og kvaliteten på FoU-arbeid heves. Men det er ikke gitt at bruken først og fremst har lærings- og kvalitetsheving som mål eller konsekvens. Poenget er derfor å mane til systemtenkning i strategiarbeid som angår en teknologi som garantert vil få konsekvenser på systemnivå. Strategien bør bygge på, og fostre, kritisk systemisk tenkning som får oss til å tenke gjennom hvorvidt det vi trenger er å jobbe raskere. Alternativt kan vi tenke at det er arbeidstrykket som må ned, for å sikre kvaliteten på det vi driver med og arbeidsvilkårene til de ansatte.

Samfunnsoppdraget: utdanning og dannelse

Et relatert systemperspektiv handler om å se bruk av KI opp mot samfunnsoppdraget til et profesjonsuniversitet som OsloMet, og et samfunnsvitenskapelig fakultet som SAM. Utgangspunktet for systemutfordringen er også den samme – at studenter og UH-ansatte tar i bruk KI som en «lettvin» eller «effektiv» løsning på sine kjerneoppgaver. Flere av innspillene vi har fått kan leses som bekymringer knyttet til hvordan vi ivaretar vårt samfunnsoppdrag som utdanningsinstitusjon, eller som varianter av følgende spørsmål: **Lærer studentene det de skal når de bruker KI som læringsverktøy?** Da tenker vi både på *læringsutbyttebeskrivelsene* i de ulike utdanningsprogrammene, som angir konkrete mål for utdanningsløpet, og på de mer *generelle akademiske ferdighetene* som inngår i en «akademisk dannelse», som lesing, skriving, analyse, argumentasjon, og kompetent deltakelse i saklig diskusjon.

Så langt utredningsutvalget erfarer mangler vi foreløpig solid forskningsbasert kunnskap som besvarer spørsmålet. Noen hint og pek finnes, men ikke tydelige svar. Oversiktsstudien til Wu og Yu (2024) viser at studenter som bruker KI ala ChatGPT har bedre læringsutbytte («learning outcomes») enn kontrollgruppene – men det fremgår ikke av studien om kontrollgruppene også mottok en form for intervensjon (noe som er relevant mtp «Hawthorne-effekten»), og effekten avtar når intervensjonen strekker seg over tid (noe som passer med en evt. «Hawthorne-effekt»¹³). Samtidig er det ikke gitt at

¹³ «Hawthorne-effekten» gjelder eksperimentstudier der (deler av) effekten på intervensjonsgruppa skyldes deltakelse i studien heller enn den konkrete intervensjonen. I vårt tilfelle, at studentene blir ekstra motivert for, eller fokusert på, læring fordi de deltar i en studie om læring, og at det økte læringsutbyttet skyldes dette heller enn den konkrete bruken av ChatGPT. Selv om tvil er reist tvil om den opprinnelige «Hawthorne-effekten», er effekter av samme type påvist i senere studier (se f.eks. McCarney et al., 2007).

målene på læringsutbytte får tak på faktisk økt forståelse hos studenten. Et mulig utfall – som også er hva Darvishi og kolleger (2024) fant i sin eksperimentstudie – er at man lærer å løse oppgaver ved hjelp av KI, men ofte uten å forstå oppgaven man har bedt KI løse, eller svaret KI lager. Vi kan med andre ord skille mellom å bruke KI som en *katalysator* – som gjør at studentene lærer mer og bedre – eller som en *krykke* – som gjør at man får jobben gjort, men uten å lære stort av det.

Gitt at det er dagens læringsmål utdanningene skal innfri, er det opplagt at det er «katalysator-bruk» vi vil fremme og «krykke-bruk» vi vil hindre – men det er uklart hva som fostrer den ene og den andre bruken. KI-strategien bør i så fall må ha som mål å fremme «katalysator-bruk», og fram til man vet hvordan, bør det vurderes om foten bør være på bremsene snarere enn på gasspedalen. I strategiarbeidet bør man også trekke et skille mellom typer av bruksområder, som f.eks. bruk av KI til simulert samtaletrening med barn – noe man ikke kan trene på i samme omfang uten å bruke KI, og som det er vanskelig å se at vil bli en krykke for studentens læringsutbytte – og f.eks. bruk av KI til å lage sammendrag av krevende pensumtekster eller til å forbedre egne tekster – som åpenbart *kan* hindre utviklingen av grunnleggende akademiske lese- og skrive-ferdigheter.

Mot dette kan man innvende at dagens læringsmål ikke er tilstrekkelig framtidrettede, og at det vi kaller «krykke-bruk» snarere bør forstås som en sentral ferdighet for fremtidens profesjonsutøvere, så lenge KI-teknologiene man lærer å mestre i studiene også blir en integrert del av profesjonenes praksis. I et slikt perspektiv vil enkelte av dagens læringsmål – inkludert de generelle akademiske ferdighetene – kanskje også fremstå som utdaterte. Man trenger ikke kunne alt selv, kunne man sagt, når det finnes verktøy som lar oss løse oppgaver raskere og bedre enn før. Det er også mye KI allerede er i stand til å gjøre mye raskere og bedre enn mennesker, og dette bør, kan man innvende, få konsekvenser for læringsmålene i høyere utdanning. Følger man denne tankegangen bør også KI-strategien være mindre bekymret for «krykke-bruk».

Vi mener begge ståstedene bør diskuteres opp mot fakultetets KI-strategi, men legger til at særlig sistnevnte ståsted innebærer en fare for målforskyvning. Det underliggende argumentet har følgende form: nye metoders potensiale for økt effektivitet bør føre til endring av læringsmålene. Men effektivitet er ikke nødvendigvis det fremste målet for et profesjonsuniversitet, og selv om KI alt nå kan løse enkelte oppgaver raskere og bedre enn mennesker, kan det likevel være bedre for oss – som mennesker, som borgere, og som samfunn – at profesjonelle lærer å gjøre dem selv. For å motvirke målforskyvning i diskusjoner om KI bør vi derfor ha følgende spørsmål i mente: 1) Hvilke av oppgavene KI løser raskere og bedre enn oss har vi likevel godt av å (lære å) gjøre selv? Og 2) Hvilke av våre målsettinger som profesjonsuniversitet er det vi strever med å nå, og er KI den beste løsningen på problemet?

KI som energisluk

En siste systemutfordring handler om klima og miljø. **Bruk av KI er forbundet med stort energibehov** (Khowaja et al., 2024), og omfattende integrering av KI i studenters og ansattes arbeidsmåter kan sånn sett komme på kant med både universitetets, og nasjonale og internasjonale klima- og bærekraftsmål. I tillegg til alle usikkerhetsmomentene, som alene tilsier noen grad av forsiktighet i fakultetets KI-strategi, mener vi disse sikre skademomentene må inkluderes (og selvsagt på en måte som sikrer en adekvat forståelse – utredningsutvalget har ikke kompetanse til å anslå det konkrete strømforbruket som vil følge av utbredt KI-bruk blant ansatte og studenter på SAM).

Muligheter for SAM

KI byr så klart også på mange og positive muligheter. Flere av dem har allerede vært antydnet i gjennomgangen av bruksformer, bekymringer og utfordringer, nemlig at KI *kan* by på muligheter for:

- effektivisering av oppgaver knyttet til administrasjon, FoU og utdanning.
- økt kvalitet på utdanning og på forskning- og utviklingsprosjekter.
- økt læringsutbytte for studenter som bruker KI som læringsverktøy.

Utover disse *generelle* mulighetene byr også KI på muligheter til å utvikle kunnskap og løsninger som er spesielt relevante for et *samfunnsvitenskapelig* fakultet, og for de konkrete *profesjonene* vi utdanner til, i utdanningen, praksisfeltet, eller begge. Noen eksempler har vi også vært innom i gjennomgangen av pågående og planlagte initiativ, både i form av utviklings- og forskningsprosjekter. KI-strategien på SAM bør si noe om hvordan slike muligheter skal gripes an framover, og man kan overordnet skille mellom tre tilnæringer: 1) *nedenfra og opp* (hvor man støtter og prioriterer blant de initiativene som vokser frem i fagmiljøene); 2) *ovenfra og ned* (hvor man bestemmer seg for hva konkret man ønsker at fagmiljøene skal utvikle); eller 3) en kombinasjon av begge.

Dersom man vil tenke *nedenfra og opp*, enten helt eller delvis, så trenger man i strategiarbeidet ikke gjøre mer enn å definere noen rutiner for å plukke opp satsinger man vil prioritere, på enten institutt- eller fakultetsnivå. Alternativt kan man velge som strategi at man skal la initiativer vokse frem før man evt. bestemmer seg for å gjøre én eller flere av initiativene til søyler i en *ovenfra og ned*-satsing på fakultetet.

Dersom man derimot vil tenke *ovenfra og ned*, enten helt eller delvis, så kan man i strategiarbeidet ta sikte på å velge ut én felles tematisk satsing for hele SAM – som innspillet vi fikk på arbeidsmøtet, «tillit og demokrati» – og invitere til eller definere satsinger innenfor dette. Alternativt kan man tenke i form av en to-trinns-rakett, der man først kartlegger hvilke sider ved profesjonenes arbeid eller vilkår som på interessante, lovende eller problematiske måter ligger an til å påvirkes av KI, og deretter bruker kartleggingen som utgangspunkt for å definere én eller flere tematiske satsinger for fakultetet.

Denne siste alternative tilnærmingen er verdt å nevne ekstra, da det er slående hvor få av respondentene i rundspørringen som reiste spørsmål og bekymringer knyttet til hvordan KI kan lede til endringer for profesjonene vi utdanner til, og hvordan dette i sin tur kan skape behov for endringer i både profesjonsutdanningene og profesjonsrettet forskning. Det er derfor utvalgets vurdering at en eller annen form for kartleggingsprosess – ikke av oss selv, men av profesjonsfeltet vi utdanner til – vil være et avgjørende grep for å sikre en responsiv KI-strategi.

Avrundende kommentar

Det er noe hastverk-preget over tilnærmingen til KI, i sektoren og samfunnet ellers – vi tror vi må velge mellom å henge med eller bli stående igjen på perrongen. Vi henger *i* for å henge *med*, og derfor stresser vi med å komme i gang. Og når vi stresser med å komme i gang, må vi også stresse med å utvikle retningslinjer og klar tenkning knyttet til hva vi *allerede* bruker KI til og hvordan.

I et kritisk perspektiv er vi på denne måten med på å gjøre de enorme teknologiselskapenes markedsførings-strategier til *selvoppfyllende profetier* (Merton, 1948), hvor brukertallene øker raskt i tråd med prognosene, og bransjen svømmer i friske penger fra investorer som forventer stor avkastning grunnet veksten i brukertall. Vi innretter oss bransjens fortellinger om at KI *er kommet for å bli*, og at KI *kommer til å forandre alt*.¹⁴ Slike fortellinger har en egen produktiv kraft – forventninger

¹⁴ Som å overta 95% av konsulentbransjens nåværende oppgaver og løse dem nærmest kostnadsfritt – en av mange prediksjoner sjefen for Open AI, Sam Altman, skal ha levert (uten videre begrunnelse), se f.eks. <https://www.cmswire.com/digital-marketing/sam-altman-ai-will-replace-95-of-creative-marketing-work/?ref=wheresyoured.at> (lest 15.05.2024).

til fremtiden gir retning her og nå, og bidrar til fortsatt vekst for teknologibedriftene (se f.eks. Borup et al., 2006).

Det kan derfor være grunner til å tenke at vi både kan og bør puste med magen – ta oss tid. Selv om vi har kjøpt fremtidsvisjonen, er det ikke gitt at det blir sånn. Flere stemmer i akademia og journalistikken stiller også kritiske spørsmål til det juridiske og økonomiske grunnlaget for særlig generative KI-teknologier. Det er reist spørsmål om treningen av KI-modellene er gjort i overensstemmelse med opphavsrettigheter, og flere rettsaker er på vei for å avklare om publikumsfavoritter som ChatGPT i det hele tatt er lovlige.¹⁵ Det er også reist spørsmål om når, og hvorvidt, teknologien blir lønnsom for bedriftene og bærekraftig mtp material- og energibehov.¹⁶ Derfor er også KI omtalt som en ny «tech-boble» som kan være i ferd med å sprekke (Naughton, 2024). Det ville ikke vært første gangen markedet kjøper bransjens luftige vyer for sure penger.¹⁷

Selv om vi naturligvis ikke skal tro så mye mer på kritikerne enn på proponentene, så kan vi ta usikkerheten som oppstår som en gave, en anledning til å utvikle en strategi der vi gir oss noe mer tid enn vi tror vi har. Vi må så klart regulere den bruken som allerede er i gang – men kanskje skal vi samtidig prøve å lage en strategi som er litt mindre bekymret for å miste toget, og litt mer bekymret for å bli sittende fast på et tog som enten står eller sporer av.

¹⁵ Se f.eks. <https://bipartisanpolicy.org/blog/legal-challenges-against-generative-ai-key-takeaways/> (lest 15.05.2024).

¹⁶ Se f.eks. Naughton (2024), <https://www.washingtonpost.com/technology/2023/06/05/chatgpt-hidden-cost-gpu-compute/?ref=wheresyoured.at> (lest 15.05.2024), eller <https://www.wheresyoured.at/peakai/> (lest 15.05.2024).

¹⁷ Et relevant eksempel er den såkalte «dot-com boblen», se <https://www.britannica.com/event/dot-com-bubble> (lest 15.05.2024).

Konklusjon: råd og anbefalinger

I tillegg til å nærlese analysen over, anbefaler utredningsutvalget følgende til arbeidet med fakultetets KI-strategi:

- 1) **Å snakke om konkrete KI-verktøy** (som Sikt KI-chat) **og teknologityper** (som generativ KI) **fremfor å snakke generelt om «KI»**. Det vil tilføre presisjon til både diskusjoner og tenkning.
- 2) **Å skille mellom KI som verktøy** (som vi og studentene bruker i vårt daglige virke), **og KI som tema** (som teknologi vi *omtaler* i undervisning, *studerer* i forskning, og *utvikler* i utviklingsprosjekter) i strategiarbeidet.
- 3) **Å utvikle og kommunisere nyttige retningslinjer for KI-bruk, av ulike art.** (Vi bør i den sammenheng skille mellom faglige og juridiske hensyn i strategiarbeidet).
 - a. Noen retningslinjer kan avklares i form av rimelig klare forbud og påbud, basert på en tolkning av eksisterende lover og regelverk.
 - b. Andre retningslinjer om grei bruk, av den sorten OsloMet sentralt har utviklet, må konkretiseres og gjøres relevant og nyttig. Dette kan ikke alltid gjøres ved å spesifisere retningslinjene. For å samkjøre forståelsen av grei bruk anbefaler utvalget derfor at det gjennomføres seminarer og kurs hvor man diskuterer konkrete eksempler og hvor grensen skal gå. Slike retningslinjer bør utvikles nedenfra, og kommuniseres lokalt.
- 4) **Å utvikle kompetanse og sikre adgang til informasjonsmateriell** (illustrasjonsvideoer, oppslagsverk, manualer, osv.) **som fremmer god KI-bruk**. Å utvikle kompetanse innebærer også strategisk rekruttering av eksperter og ressurser til opplæring av ansatte.
- 5) **Å vurdere satsing på prioriterte områder for utdanning og forskning**. Er det særskilte behov mtp endring i profesjonene/feltene vi jobber? Er det særskilte tema som forener instituttene og fagmiljøene? Bør vi posisjonere oss for å kunne agere på hastige og store utlysninger? Osv.
- 6) **Å samkjøre KI-strategien med andre relevante foreliggende strategier og retningslinjer**.
- 7) **Å huske å tenke kritisk, som kompetente samfunnsvitere, i strategiarbeidet**. Vi må vurdere effektene av bruk over tid, og på samfunnsnivå.

Litteratur

- Borup, M., Brown, N., Konrad, K., & Van Lente, H. (2006). The sociology of expectations in science and technology. *Technology Analysis & Strategic Management*, 18(3–4), 285–298.
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967.
<https://doi.org/10.1016/j.compedu.2023.104967>
- Khowaja, S. A., Khuwaja, P., Dev, K., Wang, W., & Nkenyereye, L. (2024). ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. *Cognitive Computation*. <https://doi.org/10.1007/s12559-024-10285-1>
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: A randomised, controlled trial. *BMC Medical Research Methodology*, 7(1), 30.
<https://doi.org/10.1186/1471-2288-7-30>
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8(2), 193–210.
- Naughton, J. (2024, April 13). From boom to burst, the AI bubble is only heading in one direction. *The Observer*. <https://www.theguardian.com/commentisfree/2024/apr/13/from-boom-to-burst-the-ai-bubble-is-only-heading-in-one-direction>
- Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, 55(1), 10–33.
<https://doi.org/10.1111/bjet.13334>