KIs effekt på finansiering av forskning og innovasjon

Tre scenarier fram mot 2031

1. Bakgrunn

Som bakgrunn for seminaret er det utarbeidet tre scenarier som beskriver hvordan kunstig intelligens kan endre måten Forskningsrådet vurderer og finansierer forskning og innovasjon på fram mot 2031. Ingen av dem er det forventede utfallet. De er et verktøy for å teste veivalg. Scenariene er sortert etter hvor mye KI påvirker den skriftlige søknadens egnethet som seleksjonsmekanisme. Notatet bygger på Forskningsrådets interne gjennomgang av litteraturen om KI i fagfellevurdering. Tabeller, utvalgte funn og kilder står i vedleggene.

2. Hva scenariene varierer

Tre drivkrefter behandles i de tre scenariene:

- Volum — KI kan øke antallet søknader. Dette er et kapasitetsproblem med kjente tiltak: to-trinns søknader, karantene, strengere kvalifikasjonskrav, begrensninger på antall søknader per person. Det er krevende, men løsbart med verktøy vi allerede har.
- Signal — KI kan svekke sammenhengen mellom søknadsteksten og kvaliteten på ideen og søkeren. Dette er et validitetsproblem. Det kan ikke i samme grad møtes med velprøvde verktøy, og kan tvinge fram grunnleggende endringer i hvordan vi vurderer.
- Virkemiddelinnretning — KI kan påvirke hvorvidt prosjektfinansiering fortsatt passer med hvordan forskning og innovasjon faktisk bedrives. Hvis KI korter ned forsknings- og innovasjonssyklusene og flytter ressursbehovet fra årsverk til regnekraft, data og infrastrukturtilgang, er det en risiko for at et prosjekttilskudd med forhåndsdefinerte arbeidspakker og detaljert budsjett blir en lite velegnet måte å finansiere forskning og innovasjon på. Dette kan endre hva Forskningsrådet finansierer.

En fjerde variabel går på tvers av alle tre: asymmetri. Vi kan befinne oss i ulike scenarier i ulike deler av porteføljen samtidig. Presset er tydeligst der terskelen for å søke er lavest og vurderingen hviler tyngst på innsendt tekst — SkatteFUNN først, deretter karriereordninger — og svakest i sentre og infrastruktur (vedlegg A). Ett felles svar for hele Forskningsrådet vil derfor bomme i begge retninger, og det er nærliggende å rette eksperimenter mot de mest eksponerte ordningene først.

De tre scenariene i korte trekk

	1. Absorbert	2. Erodert	3. Omformet
Kjernepåstand	Modellen bøyer seg, men holder	Søknadsteksten sier ikke lenger hvem vi bør finansiere	Prosjekttilskudd er ikke lenger riktig enhet
Volum	Moderat vekst, konsentrert i ordninger med lav terskel	Kraftig og ujevn vekst	Mangedobles; volum ikke meningsfullt
Søknaden som signal	Svakere, men fortsatt brukbar	Liten informasjonsverdi	Uten informasjonsverdi

	1. Absorbert	2. Erodert	3. Omformet
Hva som må endres	Prosess: to-trinns søknadsprosess, karantene, antallsbegrensning	Vurdering: distribuert fagfellevurdering, kortere søknad, dokumenterbare forhold, intervjuer, loddtrekning	Virkemidler: finansiere team/individ, ikke fastlåste prosjekt; institusjonell støtte; fasedelt støtte
Hvor det avgjøres	Innenfor dagens praksis	Innenfor Forskningsrådets egen myndighet	Berører mandat og forvaltningsrett
Hovedrisiko	Selvtilfredshet basert på indikator med etterslepp	Å møte et signalproblem med volumverktøy	Å handle for sent — eller demontere for tidlig

3. Scenario 1 — Absorbert: Systemet holder

Historien. I 2031 er KI-støtte i søknadsskriving like alminnelig, og like lite verdt å opplyse om, som tekstbehandling. Men nettoeffekten på seleksjonen er håndterbar. Volumet vokser moderat, konsentrert i ordninger med lav terskel, SkatteFUNN først og fremst. Panelene lærer å se bort fra poleringen og kriteriene forskyves mot det som kan etterprøves. Velprøvde tiltak som to-trinns prosesser, karantene, søknadsbegrensing, i kombinasjon med effektiviseringsgevinstene fra TIBI viser seg tilstrekkelige. Søknadsteksten blir et svakere signal, men forblir brukbar fordi den leses sammen med merittliste og team.

Hvordan Forskningsrådet må svare. Forskningsrådet har så langt ikke sett den volumveksten som rapporteres internasjonalt: 3 332 søknader hittil i år mot 4 898 for hele 2025, uten porteføljeomfattende utslag som kan tilskrives KI. Flere finansiører internasjonalt rapporterer betydelig vekst, men uten enighet om at KI er hovedårsaken. I dette scenariet er dagens kurs i hovedsak riktig, men det er behov for å øke bruken av to-trinns søknadsprosess, karantene, og antallsbegrensning på innsending av søknader.

Risikoen. Risikoen er ikke at scenariet er galt, men at det konkluderes med at man befinner seg i det på grunnlag av en indikator med etterslep. Forskningen på fagfellevurderingens «dødsspiral» (Bergstrom og Gross, med påfølgende arbeid) beskriver en selvforsterkende sløyfe: når antall søknader øker og overvelder ekspertene, blir vurderingen mindre presis, utfallene mer tilfeldige og oddsen for svake søknader bedre — noe som gir enda flere søknader. Billig KI-produksjon av søknader øker farten i sløyfen, og når kostnaden ved å søke faller under en kritisk terskel, er ikke kollapsen gradvis, men abrupt. Stabilitet i dag er da en svak indikasjon på stabilitet om tre år. Det viktige spørsmålet er ikke hvor raskt antall søknader stiger, men hvor nær vi er at kostnaden ved å søke faller under den kritiske terskelen.

4. Scenario 2 — Erodert: Søknadsteksten er ikke lenger et pålitelig signal

Historien. I 2031 har volumet vokst kraftig og ujevnt, og karakterfordelingen er presset sammen. Det Marie Skłodowska-Curie-utlysningen for postdoktorstipend opplevde i én runde — at andelen søknader under 70-prosentsterskelen falt fra 20 til 5 prosent, samtidig som antallet søknader nesten doblet seg — blir normalen i flere av Forskningsrådets ordninger. Mekanismen er eksellensmetning (Bulla og Mikula): KI-verktøyene optimaliserer søknadene direkte mot de publiserte vurderingskriteriene, terskelen for opplevd eksellens flytter seg oppover, og ekspertene presses inn i

flisespikkeri mellom det perfekte og det enda mer perfekte. Tildelingene får da karakter av loddtrekning — men en som mangler transparensen en uttalt loddtrekning ville gitt.

KI-bruk i søknader til NIH og NSF har økt kraftig siden 2023 og henger sammen med lavere semantisk særpreg, slik at prosjektene legger seg nærmere det som nylig er finansiert (Qian m.fl., PNAS 2026). Et nettbasert frilansmarked viste et lignende mønster da generativ KI gjorde velskrevne tilbud trivielle å produsere: søknadskvaliteten sluttet å skille, og ansettelsen av de mest kompetente falt med 19 prosent, mens den steg med 14 prosent for de minst kompetente (Silbert og Galdin). Konsekvensene forsterker hverandre: kostnaden per tildeling stiger, søkerne opplever utfallene som vilkårlige, og rask gjeninnsending etter avslag — allerede synlig i SkatteFUNN — blir standardatferd.

KI-bruk risikerer å gi en konkurransefordel til velfinansierte miljøer som kan kjøpe de mest avanserte verktøyene. På den annens side kan KI-verktøy redusere konkurranseulempen for søkere som ikke har engelsk som førstespråk. I et randomisert eksperiment med over 31 000 manuskripter hadde KI-tilbakemeldinger størst effekt på forfattere fra regioner der engelsk ikke er hovedspråk og for yngre og mindre siterte team, og de økte forfatternes senere KI-bruk (Wang m.fl.). Det styrker likhetsargumentet for KI-støtte — men også argumentet om at tekstkvalitet slutter å skille.

Hvordan Forskningsrådet kan svare. Vurderingsvekten kan flyttes fra teksten til noe som er vanskeligere å generere — og man må slutte å late som om instrumentet har en presisjon det ikke lenger har:

- Kortere, standardiserte prosjektbeskrivelser og to-trinns prosesser med en kort idé og sentrale CV-er i første trinn. Et nederlandsk eksperiment som sammenlignet vurderingen foretatt kun på basis av abstrac vs full søknadstekst fant at det å holde tilbake søknadsteksten ikke ga påviselig endring i rangeringen.
- Distribuert fagfellevurdering for å motvirke volumøkningen kortere søknader potensielt utløser: søkerne vurderer andre søkeres anonymiserte skisser som prisen for å sende inn sin egen.
- Mer vekt på dokumenterbare og etterprøvbare forhold — gjennomføring av tidligere prosjekter, team, tilgang til infrastruktur og data, CV.
- Intervjuer ved større tildelinger og i individrettede ordninger, slik britiske MRC har gjeninnført for alle søkere som kommer videre til fase 2. Intervju tester om søkeren forstår sitt eget prosjekt, men er arbeidskrevende og skalerer dårlig. Et alternativ er KI-gjennomførte intervjuer. Det reduserer skaleringsutfordingen, og beholder menneskelig skjønn ved at panelet leser transkripsjonen og setter karakter. Det er imidlertid mer sårbart for søkermanipuleringen enn tradisjonelle intervju.
- En grovere karakterskala i den individuelle fasen — innvilg, muligens innvilg, ikke innvilg — med loddtrekning i den omstridte midten. Valget står ikke mellom rangering og loddtrekning, men mellom en uttalt loddtrekning med kjente regler og en uuttalt en forkledd som rangering.

Avveiningene. Å flytte vekt til merittliste forsterker Matteus-effekter og stiller yngre forskere, nykommere og framvoksende fagfelt svakere. Øremerking av midler til disse gruppene og anonymisering i første seleksjonsrunde er mulige tiltak. Delvis loddtrekning er forsvarlig ut fra kunnskapsgrunnlaget, men krever at man sier åpent at man forbi et visst punkt ikke kan rangere pålitelig — noe som er sant og upopulært. Effektivisering av søknadsarbeidet kan tære på merverdien av å søke: to av tre søkere i NIFUs undersøkelse av Forskningsrådets søknadsprosess oppga at søknadsarbeidet hadde verdi uavhengig av utfallet.

Risikoen. At man svarer på et signalproblem med volumverktøy, som distribuert fagfellevurdering. Brukt alene vil volumverktøy få systemet til å se roligere ut mens det velger dårligere. Fordi trusselen treffer søknadsarbeidet før den treffer utføringen av selve forskningen, er det feil å konsentrere forsvaret rundt forsvar av søknaden som seleksjonsmekanisme. Vurderingsinnsatsen bør heller flyttes mot etterprøvbar merittliste og prosjektresultater.

5. Scenario 3 — Omformet: Prosjektmodellen er uegnet i KI-tidsalderen

Historien. I 2031 har to ting skjedd samtidig, og det er kombinasjonen som er avgjørende. For det første er det å produsere en søknad blitt tilnærmet gratis. Agenter trent på forskerens eget forfatterskap, på panelets kriterier og på nylig innvilgede søknader genererer velrettede søknader på minutter. Søknadsvolumet mangedobles og erklæringer om KI-bruk blir meningsløse. Systemet vurderer ikke lenger kvaliteten på ideene, men hvor godt agentene har lært å etterligne det finansiørene tidligere har belønnet (Rees og Wilsdon).

For det andre har KI endret forskningen og innovasjonen som finansieres. Arbeidsprosessene går langt raskere og den bindende ressursbegrensningen flyttes fra årsverk mot regnekraft, data, tilgang og kapasitet til etterprøving: hypoteser blir billige, mens validering forblir dyr og treg. En forsker kan ikke lenger troverdig forplikte seg til en fastlagt arbeidsplan tre år fram, og rapportering mot den låste planen blir en fiksjon begge parter opprettholder. Den første tallfestingen fra et KI-laboratorium viser hvor raskt det kan gå: OpenAIs forskningsorganisasjon gikk på fire måneder fra agentbruk som lå under antall menneskelige arbeidsdager til 3,1 agent-arbeidsdager per menneskelige arbeidsdag.

Hvordan Forskningsrådet må svare. Ikke med prosessreform, men med virkemiddelreform, som berører Forskningsrådets mandat:

- Finansiere team og individer fremfor prosjekter. Bevilgningen følger menneskene og ikke et forhåndsdefinert prosjektløp. En mulig oppside er økt evne til å selekter og finansiere det særlig nyskapende. Individrettede ordninger har historisk vist seg mer effektive enn prosjektbaserte ordninger på å oppnå innovative resultater.
- Velge ut på team og idé alene, og kreve full søknad med arbeidspakker og budsjett bare fra dem som er foreløpig valgt ut — detaljert for første år, med fleksibilitet deretter.
- Endre balansen mot institusjons-, senter- og infrastrukturfinansiering
- Bruke programlederskjønn etter ARPA-, ARIA- eller SPRIND-modell: rask, skjønnsbasert tildeling foretatt av enkeltindivider uten full fagfellevurdering.
- Fasedelt, utforskende finansiering som hovedregel: små beløp i første fase, observasjon av resultatene, og deretter oppskalering — med videre finansiering betinget av verifiserte delresultater.
- Finansiere validering framfor idéer: tilgang til laboratorie-, infrastruktur- og datakapasitet, for eksempel gjennom valideringskuponger eller tildelingsmodeller fra forskningsinfrastrukturfasiliteter.

Hva som gjør dette scenariet forskjellig i art. Scenario 1 og 2 kan håndteres innenfor Forskningsrådets egen myndighet. Scenario 3 kan ikke det. Det berører mandatet, balansen mellom institusjonell og konkurransebasert finansiering, forvaltningsretten og den offentlige legitimiteten — og skal disse

endres kan det ta år. Hvis man ønsker at dette alternativet skal finnes i 2031, må forberedelsene starte mens det fortsatt ser for tidlig ut.

6. Hvis Forskningsrådet bruker KI i vurderingen

Kunnskapsgrunnlaget gjør det mulig å være konkret mht. begrensningene og mulighetene knyttet til bruk av KI i vurderingen.

Hva KI gjør bra. Strukturerte, faktabaserte kontroller — kvalifikasjon, fullstendighet, tematisk treff, intern konsistens og ekspertmatching. Forfattere anser KI-vurderinger som nyttig hjelp før innsending.

Hva KI gjør dårlig. Nyskaping. I en studie av 139 søknader var nyskaping den eneste dimensjonen som var signifikant forbundet med uenighet mellom modell og menneske: søknader innvilget av mennesker, men ikke av modellen, skåret gjennomgående høyere på nyskaping. Modellene skiller også dårlig i toppen, fordi karakterene samles i øvre mellomsjikt og systematisk er høyere enn menneskers karakterer. KI favoriserer det som dominerer treningsmaterialet — noe som rammer mindre fagfelt og lite utbredte metoder og tilnærminger hardest. KI-vurderinger er minst like sårbare for betydningen av polering som menneskelige: å la en språkmodell skrive om en tekst hever karakterene KI-vurderere gir, uten at innholdet er endret (Baumann m.fl.).

Tre ting KI ikke bør få gjøre.

1. Forankre den menneskelige vurderingen. Individuelle vurderinger må utformes og registreres før eventuelle KI-vurderinger er synlige for å unngå å påvirke den menneskelige vurderingen.
2. Bli én enkelt vurderer som alle optimaliserer mot. Modeller foretrekker tekst generert av modeller fremfor mennesker, og søkere som bruker samme modell som vurderingsmodellen, får en klar fordel. Skjevheten kan mer enn halveres med enkle grep, og det mest robuste grepet er å bruke et knippe av ulike modeller.
3. Gjøre tilfeldig feil om til systematisk feil. Flaks og uflaks i trekningen av fagfeller er ille, men et konsistent tap i hver eneste trekning er verre. Menneskelige paneler bommer i ulike retninger; én modell bommer i samme retning hver gang.

Arbeidsdelingen som følger, er å be modellen om fakta, ikke om skjønn. «Oppfyller søknaden kvalifikasjonskravene og er påstandene internt konsistente» er et faktaspørsmål; «er prosjektet verdt å finansiere» er et spørsmål som må forbeholdes mennesker. Det må også legges inn sperrer mot skjulte instruksjoner rettet mot en maskinell vurderer.

Hvor andre står. I Science Europes kartlegging av 21 medlemsorganisasjoner tillater 19 av 20 organisasjoner søkerne å bruke KI, mens 13 forbyr KI i vurderingsprosessene og 3 tillater bruk i sikre, kontrollerte miljøer. Forbud er sannsynligvis ikke en stabil posisjon: en undersøkelse fra Frontiers tyder på at over halvparten av forskerne allerede bruker KI i fagfellevurdering, ofte i strid med retningslinjene. Tyske DFG tillater fra april 2026 KI som støtteverktøy for eksperten, med fullt ansvar, konfidensialitet og opplysningsplikt hos eksperten. Retningslinjene er bindende, bruken oppgis strukturert i søknadssystemet, og brudd kan regnes som vitenskapelig uredelighet.

7. Hva som holder i alle tre scenarier

Grep som kan være verdt å ta uansett hvilket scenario som viser seg å ligge nærmest sannheten.

1. Mål, per ordning, fra nå. Forskningsrådet kan i dag ikke si hvilket scenario vi befinner oss i. Måling av KI-bruk må skje på overordnet virkemiddelnivå, fordi deteksjon av KI-tekst i enkeltsøknader er upålitelig. Egenerklæringer er ikke nok: undersøkelser viser høy bruk, lav åpenhet og utbredt frykt for å bli straffet for å opplyse om bruk. TIBIs strukturerte data gjør måling mulig. De viktigste varselindikatorene er søknadsvekst per ordning og runde, andelen søknader med toppkarakter, tiden fra avslag til gjeninnsending, andelen prosjekter der KI inngår i selve arbeidet og — senest, men mest direkte — resultater i ettertid målt mot opprinnelig karakter. Derfor må nullpunktet etableres nå (vedlegg B). RORIs KI prosjekt vil være essensielt her.
2. Mennesker må forbli ansvarlige for hver eneste tildelingsbeslutning.
3. Bygg kapasitet til å eksperimentere og evaluere. I scenario 2 og 3 må seleksjonen endres under tidspress. Om det skjer på grunnlag av kunnskap eller mageføelse, avhenger av om mulig ny verktøykasse er testet ut i forkant gjennom kontrollerte forsøk, skyggepaneler, etc.
4. Kort ned søknadene og legg mer vekt på idé og dokumenterbare forhold nå. Det er en effektiviseringsgevinst i scenario 1, en nødvendighet i scenario 2 og et skritt på veien i scenario 3 — forutsatt at volumkonsekvensene av kortere søknader håndteres samtidig.
5. Bevar merverdien av å søke. Effektivisering bør fjerne det søkerne opplever som verdiløst, ikke det de opplever som utviklende. Tidligere undersøkelser tyder på at idégenerering og utvikling av nye samarbeidskonstellasjoner er elementer søkerne setter høyt i søknadsarbeidet.
6. Samordne internasjonalt. Et globalt signalproblem løses ikke av en norsk regel. De relevante kanalene er RORI, ERA-tiltaket om vurdering og belønning av forskning, Science Europes KI-arbeidsgruppe og Global Research Council.

8. Spørsmål til diskusjon

1. Hvor langt bør Forskningsrådet gå i forberedelsene? Forberedelse er forholdsvis rimelig og improvisasjon dyrt, men forberedelse har også en signaleffekt, internt og i sektoren.
2. Hvor mye vekt kan flyttes til merittliste og dokumenterbare forhold, og til hvilken kostnad for nykommere — og er øremerking og anonymisering i første runde del av pakken?
3. Finnes det virkemidler — loddtrekning, tak på antall søknader, programlederskjønn, KI-støttet siling — som anses som prinsipielt uaktuelle?
4. Bør forbudet mot eksterne eksperters KI-bruk anses som en varig tilnærming eller en bro mot robust bruk? De realistiske alternativene er DFGs regulerte tillatelse med opplysningsplikt eller et sikkert internt miljø.

Vedlegg A. Eksponering på tvers av porteføljen

De tidlige norske signalene peker mot at presset er tydeligst der terskelen for å søke er lavest og vurderingen hviler tyngst på innsendt tekst.

Ordning	Nivå	Hvorfor
SkatteFUNN	Høyest	Lav terskel, høyt volum, i hovedsak tekstbasert vurdering; økt søknadstrykk og raske re-innsendinger rapporteres allerede.
Innovasjonsprosjekter	Høy	Rapporter om mer polerte, men mer generiske beskrivelser; stor søkermasse med kommersielt insentiv til å søke bredt.
Forskerprosjekter (FRIPRO og liknende)	Middels	Høyt volum og tekstavhengig, men vekt på merittliste gir delvis beskyttelse.
Sentre og store samarbeidsordninger	Lavere	Forpliktelser mellom flere parter og institusjonelle interesser begrenser hvor langt en generert tekst kan bære en søknad.
Infrastruktur	Lavest	Vurderingen hviler på dokumenterbar kapasitet, kostnad og forpliktelse snarere enn på fortelling.

Vedlegg B. Tidlige varselindikatorer

Scenariene lar seg skille fra hverandre på forhånd hvis de riktige tingene måles, per ordning og mot et nullpunkt.

Indikator	Hva den forteller oss	Peker mot
Vekst i søknadsvolum per ordning, per runde	Om kostnaden ved å søke har falt nok til å endre atferd	Vedvarende vekst over omtrent 15 prosent i året, konsentrert i ordninger med lav terskel: scenario 2. Mangedobling innenfor én runde: scenario 3.
Andel søknader med toppkarakter(er)	Om karakterene fortsatt skiller søknadene	En fallende andel med lave karakterer, og en stigende andel helt i toppen, er det klareste tidlige tegnet på eksellensmetning — Marie Curie-mønsteret.
Andel avslag og frafall blant fagfeller	Om ekspertgrunnlaget absorberer eller avviser belastningen	Økende avslag er et tidlig varsel om kapasitetssvikt, og ligger foran volumtallene.
Tid fra avslag til gjeninnsending	Om omskriving er blitt tilnærmet gratis	Gjeninnsending innenfor samme syklus i stort omfang viser at kostnadsstrukturen er fundamentalt endret.
Estimat på virkemiddelnivå av KI-assistert tekst	Om karakteren av innsendelsene endrer seg	Mer enn halvparten av artiklene publisert i 2025 viser språklige spor av språkmodeller, mot rundt én av ti i 2023.
Ikke-verifiserbare eller oppdiktede referanser	Om integritetsproblemer følger med volumet	Allerede rapportert av finansiører i Science Europe; automatiserte kontroller fant hundrevis av oppdiktede siteringer i antatte konferanseartikler som fagfellene hadde oversett.

Indikator	Hva den forteller oss	Peker mot
Resultater i ettertid mot opprinnelig karakter	Om vurderingen fortsatt forutsier noe	En svekket sammenheng er det mest direkte beviset på signalsvikt — og det som viser seg senest, og derfor grunnen til at nullpunktet må etableres nå.
Konsentrasjon av tildelinger blant etablerte søkere	Om svarene våre produserer Matteus-effekter	Økende konsentrasjon ville vise en rettferdighetskostnad ved våre mottiltak.
KI i selve prosjektene	Om forskningspraksis endrer seg	En økende andel prosjekter med KI i arbeidsplanen og budsjettposter til regnekraft og agenter peker mot scenario 3.

Vedlegg C. Utvalgte funn fra kunnskapsgrunnlaget

Tall og eksempler som ligger til grunn for hovedteksten.

Søknadsvolum og mottiltak hos andre finansiører

- INFLECTION-arbeidsgruppen av finansiører fra Europa, Nord-Amerika og Australia rapporterer betydelig vekst hos blant andre La Caixa, Wellcome, SSHRC, CIHR, ZonMW, ARC og NHMRC, men peker på at etterspørselsstyrende tiltak, endringer i ordningene og bredere systempress også påvirker volumet. Det er uklart om KI-assisterte søknader ennå har noen konkurransefordel. Den første fagfellevurderte studien av søknader finner at KI-bruk i NIH- og NSF-søknader har økt kraftig fra 2023, er tydelig todelt mellom minimal og omfattende bruk, og henger sammen med lavere semantisk særpreg. Hos NIH har KI bruk positiv effekt på gjennomslag, mens det hos NSF ikke har effekt, uten kjent forklaring (Qian m.fl., PNAS 2026).
- NIH innførte fra september 2025 et tak på seks søknader per prosjektleder per år, etter at enkelte forskere sendte inn mer enn førti i én runde. NIH vil heller ikke behandle søknader som i vesentlig grad er utarbeidet av KI. ERC begrenser gjeninnsending, men trakk i april 2026 tilbake en varslet innstramming etter reaksjoner fra forskersamfunnet, og flere finansiører krever at søkerne opplyser om KI-bruk.
- Villum-fondet tester ut loddtrekning for å bestemme hvem som får sende inn en skisse — et eksempel på volumet kortere skjemaer kan utløse.
- Britiske Medical Research Council gjeninnførte i mars intervjuer for alle kortlistede søkere, uttrykkelig for å motvirke KI-forsterkning av søknadene. Wellcome og La Caixa peker på intervjuer som sikkerhetsmekanisme. Intervjuer favoriserer samtidig de trygge og godt forberedte, og vil selv være KI-forberedte innen et år etter innføring.
- Etterspørselsstyring delegerer deler av Forskningsrådets portvokterfunksjon til institusjonene, som vil anvende sine egne kriterier, ujevnt og uten klageadgang.

Skjevheter, deteksjon og merverdi

- Silisiumtaket (Bulla og Mikula) svekker argumentet om at KI-assistert skriving i hovedsak virker utjevnende, som har vært en av de bedre grunnene til en tillatende holdning. Det samme mønsteret er synlig på europeisk nivå, i forskjellen i forskningsadministrativ støtte og søknadskapasitet mellom institusjoner i Nord- og Vest-Europa og i Sentral- og Øst-Europa. Et randomisert felteksperiment trekker i motsatt retning: blant over 31 000 arXiv-manuskripter

var forfattere som fikk KI-tilbakemelding, 12,55 prosent mer tilbøyelige til å revidere, og deres senere KI-bruk økte. Effekten var størst for forfattere fra regioner der engelsk ikke er hovedspråk, og for yngre og mindre siterte team (Wang m.fl., ikke fagfellevurdert).

- Rees og Wilsdon, som selv anbefaler å flytte tyngdepunktet mot prosjektlederen, teamet og deres tidligere og pågående forskningsprogrammer, navngir øremerking og anonymisering i første seleksjonsrunde som mottiltak mot Matteus-effektene.
- Identifisering av KI-generert tekst på enkeltdokumentnivå er upålitelig og kanskje prinsipielt umulig: OpenAI trakk tilbake sin egen klassifikator på grunn av dårlig treffsikkerhet, og detektorer er sårbare for omgåelse. Estimering på korpusnivå virker derimot. Mer enn halvparten av artiklene publisert i 2025 viser språklige spor av språkmodeller, mot rundt én av ti i 2023 (Siler, PNAS 2026).
- Søknadsteksten var delvis frikoblet fra forskningen allerede før KI: i 400 par av innvilgede søknader og tilhørende publikasjoner fra Kina og USA blåses policyrettet språk opp i søknaden og forsvinner i publikasjonen (Valdivia m.fl.).
- En sveitsisk studie fant at deltakelse i en finansieringskonkurranse økte publisering og utvidet samarbeidsnettverk enten søkeren vant eller ikke. Mot denne merverdien står kostnaden: søknadsprosessene koster rundt 13 prosent av tilskuddenes verdi, og 89 prosent av kostnaden bæres av søkerne (Pollitt m.fl., 12 617 respondenter hos to finansiører). Det er denne kostnaden agentene gjør billig.

Utviklingstrekk og virkemidler i scenario 3

- OpenAIs forskningsorganisasjon brukte i august 2026 3,1 agent-arbeidsdager per menneskelige arbeidsdag, mot under én før juni (egenrapportert). Hos en stor internasjonal finansiør er KI i selve prosjektene så langt forbundet med beskjedne resultatforbedringer, men med større team, mer midler til personell og bredere oppgaver — omorganisering snarere enn effektivisering (Hosseinioun m.fl.).
- Howard Hughes-modellen med å finansiere personer framfor prosjekter er den etablerte varianten av teamfinansiering.
- Fasedelt, utforskende finansiering er tilpasningen Silbert og Galdin identifiserer i markeder der forhåndssignalene har brutt sammen; ARPA-lignende mekanismer er den etablerte parallellen i forskningsfinansiering. Google DeepMind gir et beslektet argument: agenter gjør idéer billige, mens etterprøving forblir kostbar og treg.
- KI-innfødt vurdering slik Rees og Wilsdon skisserer den: å profilere søkere på tvers av publiseringsgjennomslag, samarbeidsrekkevidde, metodisk bredde og karrierebane, og bruke det rikere bildet som vurderingsgrunnlag.
- Om forbud argumenterer Rees og Wilsdon for at de sannsynligvis vil bli bredt ignorert fordi insentivet til å bruke KI er sterkt og sannsynligheten for å bli oppdaget lav, og at et forbud stiller dem som ikke skriver engelsk som førstespråk dårligere. INFLECTION-gruppen rapporterer at fagfeller bruker KI også der retningslinjene fraråder eller forbyr det. Google DeepMind anser det å gi fagfeller tilgang til agenter som en av de viktigste prioriteringene for finansiører. DFG har valgt en regulert tillatelse: deres retningslinjer krever at ekspertene bruker lokalt eller institusjonelt driftede systemer eller kommersielle systemer med avtale som utelukker trening, at bruken oppgis, at eksperten foretar kritisk kontroll og at ansvar som ikke kan overføres.

Fagfeller oppgir i søknadssystemet om KI er brukt, om søknadsdokumenter er lastet opp og om generert tekst er brukt.

KI i vurderingen

- De fleste finansiørene i INFLECTION-gruppen tester allerede kontroll av tematisk treff og kvalifikasjon, ekspertmatching, møtereferater, generering av tilbakemeldingsbrev og interne kunnskapsassistenter. Den første testen av automatisk siling mot en nasjonal utlysnings faktiske komitébeslutninger (959 søknader) viser risikoen: en transparent nøkkelordmodell fanget opp 79 prosent av komiteens valg, en språkmodellbasert klassifikator 46 prosent (Nagarajappa m.fl.).
- Ekspertmatching: én studie rapporterer 73 prosent reduksjon i tiden det tar å identifisere fagfeller.
- Samsvar: GPT-4s tilbakemeldinger overlappet med en enkelt menneskelig fagfelle i rundt 31 prosent av tilfellene for Nature-tidsskriftene, mot 29 prosent menneske-til-menneske, og 39 mot 35 prosent for en maskinlæringskonferanse (Liang m.fl.). For karakterer på søknader er samsvaret svakere: mellomstore åpne modellers karakterer korrelerte 0,22 i snitt og 0,33 i beste fall med ekspertenes snitt, om lag 56 prosent av samsvaret mellom fagfeller (Sandström m.fl.).
- Nyskaping: i studien av 139 søknader sammenfalt KI-genererte vurderinger med menneskelige tildelingsbeslutninger i 60 prosent av tilfellene. I tidsskriftvurdering kommenterte en språkmodell forskningens implikasjoner 7,3 ganger oftere enn mennesker, og nyskaping 10,7 ganger sjeldnere.
- Karaktersetting: språkmodeller forutsier enkeltkarakterer dårlig, er moderat gode til å rangere innenfor sammenlignbare sett, og gir omtrent samme karakter fra tittel og sammendrag som fra fullteksten. Anvendt på en allerede sammenpresset fordeling presser de den ytterligere sammen. Karakterene er dessuten for høye: frontmodeller ga søknader 3,2–3,8 på en firepunktsskala mot et menneskelig snitt på 2,5 (Li m.fl.), flere GPT-modeller godkjente over 98 prosent av manuskriptene i et testsett (Verma m.fl.), og i 11 000 maskinvurderinger var karakterene mindre spredt og for positive og overmodige, mens svakheter mennesker fant, ofte ble oversett (Żurawicki m.fl.).
- Selvpreferanse: i et stort kontrollert ansettelseseksperiment lå selvpreferansen mot menneskeskrevet materiale på mellom 67 og 82 prosent på tvers av store kommersielle og åpne modeller, og kandidater som brukte samme modell som vurdereren, hadde 23 til 60 prosent høyere sannsynlighet for å bli kortlistet. Skjevheten kunne mer enn halveres med enkle grep, der det mest robuste var at et ensemble av ulike modeller stemte. I en test av søknadsvurdering slo vurdering seksjon for seksjon både én samlet gjennomgang og et «råd» av personaer (Thorne m.fl.). Og karakterene KI-vurderere gir, kan heves ved å la en språkmodell skrive om teksten, samtidig som vurderingene er for samstemte («hivemind») (Baumann m.fl.).
- Arbeidsdelingen mellom maskin og menneske tar de standardiserte delene av en vurdering — prosjektledelse, gjennomførbarhet, formalia — og lar fagfellene beholde den delen de synes er verdt å gjøre. Hvordan vurderere reagerer på maskinelle innspill, bør Forskningsrådet undersøke snarere enn anta. De første feltforsøkene gir et utgangspunkt: i AAAI-26 fikk alle 22 977 innsendinger en merket KI-vurdering, og forfattere foretrakk den framfor menneskelige vurderinger på teknisk presisjon (Biswas m.fl.). Forfattere vil helst bruke KI før innsending, ikke som erstatning for fagfeller (Lemberger m.fl.).

- Den gjennomgående rådgivningen i litteraturen er nøye avgrensede piloter med eksplisitte evalueringskriterier, åpenhet og ansvarlighet, snarere enn enten ukritisk adopsjon eller refleksmessig avvisning. Åpenhet kan kombineres med vern mot manipulering: en ekstern kontrollør kan bekrefte hvilken modell, vurderingsmal og instruks som ble brukt, uten at søkerne får innsyn — noe som er viktig fordi søknader kan inneholde skjulte instruksjoner rettet mot en maskinell vurderer (Bicakci).

Vedlegg D. Kilder

Notatet bygger på Forskningsrådets interne litteraturgjennomgang av KI-bruk i fagfellevurdering, og gjennom den på det nedenstående.

- Bulla, M. og Mikula, P., When AI turns grant evaluation into a lottery — eksellensmetning, «Red Queen»-kappløpet i søknadsskriving, silisiumtaket, og argumentet for loddtrekning over en kvalitetsterskel.
- Rees, G. og Wilsdon, J., Could agentic AI topple grant-funding systems? (Nature 652, 1119–1121, 2026) — agentbasert søknadsskriving, hvorfor forbud svikter, forskyvningen mot prosjektleder, team og program, verifisering av merittliste, intervjuer og etterspørselsstyring, og KI-innfødte finansieringssystemer.
- Silbert, J. og Galdin, A., Making Talk Cheap: Generative AI and Labor Market Signaling — sammenbruddet i skriftlige signaler i et nettbasert arbeidsmarked, og forskyvningen mot utforskende kontrakter.
- Simsek, M., de Vaan, M. og van de Rijt, A., Do grant proposal texts matter for funding decisions? A field experiment (Scientometrics, 2024) — det nederlandske forsøket med å holde tilbake søknadstekster.
- Liang, W. m.fl., om språkmodellgenererte tilbakemeldinger i fagfellevurdering — overlapp med menneskelige fagfeller, forskeres oppfatninger, og forskjellen mellom maskinens og menneskets oppmerksomhet mot nyskaping.
- Thelwall, M. og Langfeldt, L. m.fl., om egenskapene ved språkmodellers kvalitetskarakterer — sammenpressing, rangering kontra karaktersetting, tilgjengelighets- og gjennomsnittsskjevhet, og råd om å bruke språkmodeller til å støtte snarere enn å erstatte ekspertskjønn.
- Bergstrom, C. og Gross, K., og påfølgende arbeid om fagfellevurderingens dødsspiral — den selvforsterkende sløyfen mellom innsendelsesvolum og vurderingspresisjon, og dens ikke-lineære kollaps under billig KI-assistert produksjon.
- Russo, G. m.fl., The AI Review Lottery (Proc. ACM Hum. Comput. Interact., 2024) — utbredelsen av KI-assisterte vurderinger ved ICLR 2024 og virkningen på karakterer og antakelse.
- Liang, W. m.fl., estimering på korpusnivå av språkmodellbearbeidet tekst i fagfellevurderinger — 6,5–16,9 prosent ved store KI-konferanser, og den tilhørende homogeniseringen av tilbakemeldinger.
- Porsdam Mann, S. m.fl., AI and the Future of Academic Peer Review (2025) — argumentet for nøye avgrensede piloter med eksplisitte evalueringskriterier, åpenhet og ansvarlighet.
- Science Europes KI-arbeidsgruppe, kartlegging av KI-praksis på tvers av 21 medlemsorganisasjoner (juni 2026), og DFGs posisjon fra april 2026 om KI i fagfellevurdering.

- Møtenotater fra INFLECTION-arbeidsgruppen — finansiørenes rapporter om søknadsvolum, KI-bruk hos søkere og fagfeller, intern utprøving, og intervjuer som sikkerhetsmekanisme.
- NIH, NOT-OD-25-132 — retningslinjer for KI-bruk i søknader og taket på seks søknader i året, med virkning fra september 2025.
- Elseviers forskerundersøkelse 2025 (3 234 forskere, 113 land) — utbredelsen av KI og bruken av den til å utarbeide søknader.
- RoRI, Agentic AI and the grant-funding system: developments and debates since April 2026 (arbeidsnotat, 10. september 2026) — andre oppdatering etter Rees og Wilsdon; grunnlag for oppdateringene i denne versjonen.
- Wallace, D. m.fl., Conjecture Machines: AI agents and the new validation bottleneck in science (Google DeepMind, 15. juli 2026) — valideringsflaskehalsen og prioriteringer for finansiører.
- OpenAI, Research acceleration: the view inside OpenAI (6. september 2026) — egenrapportert agentbruk i forskningsorganisasjonen.
- Qian, Y. m.fl., The rise of large language models and the direction and impact of US federal research funding (PNAS, 2026) — KI-bruk og semantisk særpreg i NIH- og NSF-søknader.
- Siler, K. (PNAS 123, e2605754123, 2026) — språkmodellspor i publiserte artikler 2023–2025.
- Wang, B. m.fl., Human-AI collaboration in science at scale: a global large-scale randomized field experiment (arXiv, 2026).
- Hosseinioun, M. m.fl., Artificial intelligence in science: returns, reallocation, and reorganization (2026).
- Valdivia, W. D. m.fl., Strategic decoupling between grant and publication language in AI and cancer research (Frontiers in Research Metrics and Analytics, 2026).
- Pollitt, A. m.fl., The costs and benefits of research grant funding peer review (F1000Research, 2026).
- Biswas, J. m.fl., AI-assisted peer review at scale: the AAAI-26 AI review pilot (arXiv, 2026), og Lemberger, T. m.fl., How do authors want to use AI for review? (EMBO Reports, 2026).
- Verma, N. m.fl., All accept, no reject: evaluating LLMs as «peer» reviewers (CHI 2026), og Li, W. m.fl., Large language models for peer review in biotechnology (New Biotechnology, 2026).
- Sandström, U. m.fl., Can large language models evaluate grant proposal quality? Revisiting the Wenneråas and Wold peer review data (arXiv, 2026).
- Baumann, J. m.fl., Stop automating peer review without rigorous evaluation (arXiv, 2026), og Żurawicki, K. m.fl., PRAIB: Peer Review AI Benchmark (arXiv, 2026).
- Nagarajappa, C. m.fl., Recall, risk, and governance in automated proposal screening for research funding (arXiv, 2026).
- Thorne, W. m.fl., Evaluating LLM-based grant proposal review via structured perturbations (arXiv, 2026), og Bicakci, K., Making AI-assisted grant evaluation auditable without exposing the model (arXiv, 2026).
- DFG, AI in the review process: retningslinje (skjema 4.04), FAQ (oppdatert 27. mai 2026) og white paper.
- ERCs vitenskapelige råd, ERC Scientific Council readjusts rules for reapplication (29. april 2026).

Kunstig intelligens og konsekvenser for forsknings- og innovasjonssystemet

Diskusjonsnotat

Innledning

Kunstig intelligens (KI) er i ferd med å endre hvordan forskning utføres, kvalitetssikres, vurderes og finansieres. Høstens publisering av over 700 KI-genererte matematiske preprints fra OpenAI viser hvor raskt utviklingen går. Den reiser også spørsmål som angår hele sektoren: Hvem får kreditt, hvem kvalitetssikrer, og hvem setter dagsorden for forskningen?

Notatet er et utgangspunkt for diskusjon. Det oppsummerer nyere forskning og erfaringer, peker på spenninger og avveininger, og avsluttes med spørsmål vi ønsker innspill på. Det overordnete spørsmålet er hvordan sektoren kan ta i bruk KI på måter som styrker kvalitet, mangfold og tillit.

Hva vet vi om KI i forskningen i dag?

KI-bruken i forskningen vokser raskt, men den er ujevnt fordelt og ofte usynlig. Antall naturvitenskapelige publikasjoner som omtaler KI, økte nesten 30 ganger fra 2010 til 2025 (Stanford HAI, 2026). Samme rapport viser til at en artikkel som utelukkende var generert av KI ble antatt på en fagfellevurdert workshop i 2025, men at listen over eksperimentelt bekreftede KI-funn fortsatt er kort. Ved forskningsfronten går det likevel fort. Googles «Co-Scientist» kom på to dager fram til den samme mekanismen for genoverføring mellom bakterier som forskere ved Imperial College hadde brukt år på å avdekke.

Forskerne er selv ambivalente. I en Nature-undersøkelse blant 1 907 forskere i 2026 var 48 prosent negative til KI, men nesten 60 prosent fryktet å bli hengende etter uten. I Australias første nasjonale kartlegging oppga nesten tre av fire av 1 034 forskere at de bruker KI, mens 54 prosent mente institusjonens opplæring var utilstrekkelig (Science & Technology Australia, 2026).

Vi har foreløpig ingen tilsvarende nasjonal kartlegging i Norge. Storbritannia gjennomfører nå sin første nasjonale undersøkelse av KI i forskning, i regi av UK Metascience Unit, med resultater i desember 2026 (RAND Europe, 2026). Tabellen oppsummerer muligheter og risikoer langs forskningsprosessen, slik de drøftes i resten av notatet.

Ledd i forskningsprosessen	Muligheter	Risikoer
Problemvalg og idéutvikling	Rask oversikt over litteratur; nye koblinger på tvers av fag	Mange konvergerer mot de samme, populære spørsmålene
Gjennomføring og analyse	Automatisering, nye modeller og raskere hypotesetesting	Feil som er vanskelige å oppdage; validering blir flaskehals
Skriving og publisering	Lavere språkbarrierer, særlig for forskere uten engelsk som morsmål	Tekstkvalitet mister signalverdi; flere svake artikler og artikkelfabrikker
Utdanning og rekruttering	Tilpasset veiledning; ny kompetanse i grenseflaten mellom fag og KI	Færre øvingsoppgaver for unge forskere; svekket faglig dømmekraft

Kvalitet, mangfold og forskningsintegritet

KI ser ut til å gi gevinster for den enkelte forsker, men kan svekke bredden i forskningen som helhet. En studie av 41,3 millioner naturvitenskapelige artikler finner at forskere som bruker KI, publiserer tre ganger så mye, får nesten fem ganger så mange siteringer og blir prosjektledere nesten halvannet år tidligere (Hao mfl., 2026). Samtidig dekker den KI-støttede forskningen et snevrere tematisk felt, og forskerne samhandler 22 prosent mindre med hverandre.

Mønsteret finnes også i eksperimenter: Tilgang til KI-genererte ideer gjør enkelttekster mer kreative, men tekstene blir likere hverandre (Doshi og Hauser, 2024). Det som er rasjonelt for den enkelte, kan bli uheldig for fellesskapet. Nature har kalt risikoen en «diffus monokultur»: mange temaer, men samme mal for å produsere resultater.

Virkningene varierer likevel mye mellom fag. En studie i Research Policy av over 80 millioner publikasjoner i mer enn 170 fagfelt finner at KI i snitt henger sammen med mer nyskapende og mer sitert forskning. Gevinsten er størst der KI er godt integrert i faget, og der kunnskapsfeltet er fragmentert og kombinatorisk komplekst. Konsekvensene av KI må altså vurderes fag for fag.

En tredje utfordring gjelder tekst som kvalitetssignal. Når velskrevne tekster blir billige, skiller god språkføring ikke lenger sterk fra svak forskning, viser analyser av over 260 000 manus og preprints. Det gjør det også lettere for artikkelfabrikker å masseprodusere formelt ryddige, men faglig tynne bidrag. Skriving er dessuten en måte å tenke på, og overføres den til maskinen, kan sentrale ferdigheter svekkes.

KI-bruk kan også svekke allmennhetens tillit til forskningen. I en forhåndsregistrert, representativ britisk undersøkelse støttet 77 prosent bruk av skattepenger til forskning generelt, men bare 42 prosent når forskerne brukte KI, og 34–40 prosent ville stole mindre på både forskningen og forskerne (Bretter mfl., 2026). Samtidig støttet over 80 prosent KI-bruk når forskerne sto for en vesentlig egen innsats. Hvordan KI-bruken beskrives og begrunnes, betyr altså mye.

Til sist reiser KI nye spørsmål om kreditering og ansvar. Da OpenAI i oktober la ut over 700 KI-genererte matematiske preprints, ble tre trukket dagen etter på grunn av en fortegnsfeil, ifølge Nature. Hvem skal kvalitetssikre maskinproduserte resultater?

Utdanning, kompetanse og karriere

KI kan svekke nettopp de ferdighetene som trengs for å bruke KI forsvarlig. Forskere trenger dypere fagkunnskap enn før for å vurdere KI-genererte ideer, mens studenter og stipendiater får færre anledninger til å bygge den kunnskapen gjennom eget arbeid. I et felteksperiment med nesten 1 000 elever fant man at fri tilgang til en vanlig chatbot ga bedre resultater under øving, men 17 prosent dårligere resultater når hjelpemiddelet ble tatt bort (Bastani mfl., 2025).

En versjon med pedagogiske sperrer fjernet det meste av den negative effekten. Studien gjelder videregående skole, men lærdommen er relevant for høyere utdanning og forskerutdanning: Hvordan KI tas i bruk, betyr mer enn om den tas i bruk. Verdien av menneskelig arbeid flyttes mot å stille gode spørsmål, velge problemer, vurdere metoder og oppdage feil, og det stiller nye krav til veiledning, emnedesign og ph.d.-utdanning.

KI utfordrer også hvordan forskere vurderes. Når KI-brukere publiserer og siteres mer, kan rekruttering og opprykk som vektlegger kvantitative mål, belønne volum fremfor faglig bidrag. Det styrker argumentene for ansvarlig forskningsvurdering i tråd med DORA-erklæringen, som Forskningsrådet støtter, og arbeidet i CoARA.

Gevinstene fordeles heller ikke nøytralt. KI senker språkbarrierer for forskere som skriver på engelsk som andrespråk, men karrieregevinstene ser ut til å tilfalle forskere med engelsk som morsmål i større grad (Nature-kommentar, 2026). Tverrfaglige team med forskningsingeniører, dataspesialister og fagfolk blir viktigere, mens karriereveiene for slike roller ofte er svake.

Institusjonene har begynt å svare. Kravene fra institusjoner, tidsskrifter og finansiører er likevel ikke samordnet.

Innovasjon, infrastruktur og internasjonalt samarbeid

Tilgang til modeller, regnekraft og data blir en forutsetning for konkurransedyktig forskning, og tilgangen er skjevt fordelt. En rådgivende gruppe av matematikere advarte i høst mot et todelt system der KI-selskapenes laboratorier, med egne lukkede modeller, løper fra resten av fagfeltet. Privat sektor trekker samtidig til seg en voksende andel av KI-forskerne (Jurowetzki mfl., 2025).

Norge har satset betydelig. KI-milliarden finansierer seks nasjonale forskningssentre for KI fra høsten 2025, og Forskningsrådet hadde over 600 aktive KI-prosjekter samme år. EUs satsing på KI åpner nye muligheter for norske miljøer.

KI kan også styrke innovasjon. Billigere og raskere forskning kan øke etterspørselen etter forskning og forskere, slik nye selskaper innen proteindesign, materialer og værvarsling allerede viser. Norske virksomheter tar likevel KI i bruk gradvis: Enkel bruk øker raskt, mens integrert bruk i offentlig sektor fortsatt er på et tidlig stadium (Krogstie, 2025). Kandidater med både fagkunnskap og KI-kompetanse blir dermed en viktig kanal for kunnskapsoverføring.

Geopolitikken påvirker også samarbeidet. EU, USA og Kina regulerer data, modeller og teknologi ulikt, og eksportkontroll og forskningssikkerhet legger nye føringer på internasjonalt samarbeid. Beslutningene de neste fem årene om hvem som får tilgang til KI-basert forskningsinfrastruktur, kan forme kunnskapsgeografien i en generasjon (Turekian, 2026).

Spørsmål til diskusjon

Forskningsrådet ønsker innspill på **hvordan KI endrer forsknings og innovasjonsprosessen på ulike fagfelt og hvordan KI endrer institusjonene.** Eksempelvis:

1. **Kvalitet og mangfold:** Hvordan kan vi sammen sikre at KI brukes til å utforske nye spørsmål, og ikke bare til å produsere mer av det samme? Hva bør institusjoner og finansiører belønne?
2. **Forskjeller mellom fag:** Hvilke fag har mest å vinne og mest å tape, og hvordan bør det påvirke retningslinjer og virkemidler?
3. **Åpenhet og tillit:** Hvilket nivå av åpenhet om KI-bruk er realistisk og nyttig i publikasjoner, søknader og vurderinger? Hvordan unngår vi at ærlighet straffes, og at åpenhet svekker allmennhetens tillit?
4. **Forskerutdanning:** Hvilke ferdigheter må kandidater og stipendiater utvikle, og hvordan bør veiledning og ph.d.-utdanning endres?
5. **Karriere og meritter:** Hvordan bør rekruttering og opprykk ta høyde for at KI kan blåse opp publikasjons- og siteringstall?
6. **Infrastruktur og samordning:** Hvilke felles løsninger for modeller, regnekraft og data trenger sektoren, og hvor bør retningslinjer samordnes nasjonalt?

7. **Kunnskapsgrunnlag:** Er det behov for en nasjonal kartlegging av KI-bruk i norsk forskning, etter modell av de britiske og australske undersøkelsene? Bør den også omfatte befolkningens syn på KI-støttet forskning, og hvem bør gjennomføre og finansiere den?

Kilder

Forskningslitteratur

- Bastani, H. mfl. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *PNAS*, 122(26), e2422633122.
- Bretter, Todorova, Curtis og Hornsey (2026). Public views on the use of AI in scientific research. *Research Policy*, 55(10), 105632.
- Doshi, A. R. og Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290.
- Hao, Q., Xu, F., Li, Y. og Evans, J. (2026). Artificial intelligence tools expand scientists' impact but contract science's focus. *Nature*, 649, 1237–1243.
- Jurowetzki, R. mfl. (2025). The private sector is hoarding AI researchers: What implications for science? *AI & Society*.
- Kobak, D. mfl. (2025). Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11(27), eadt3813.
- Krogstie, J. (2025). Bruk av kunstig intelligens (KI) i norsk offentlig sektor: Status 2025. *NIKT*, 37(2).
- Liang, W. mfl. (2024). Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. *Proceedings of ICML 2024*, PMLR 235, 29575–29620.
- AI in science: When and where it makes a difference (2026). *Research Policy*.

Rapporter, retningslinjer, kommentarer og nyhetsartikler

- Europakommisjonen (2025). European Strategy for AI in Science og RAISE.
- Europakommisjonen og ERA-landene. Living guidelines on the responsible use of generative AI in research.
- Griffin, C. mfl. (2024). A new golden age of discovery. Google DeepMind.
- ICML (2026). On violations of LLM review policies.
- Nature (2025). AI tool detects LLM-generated text in research papers and peer reviews (om AACR-studien).
- Nature (2026). Human scientists trounce the best AI agents on complex tasks (om Stanford AI Index 2026).
- Nature (2026). Scientists have a bad case of AI FOMO, Nature poll reveals.
- Nature (2026). The uncritical adoption of AI in science is alarming (kommentar).
- Nature (2026). Will AI spark a scientific renaissance — or a diffuse monoculture?
- Nature (2026). AI co-scientists are revolutionizing how research is done.
- Nature (2026). Nyhetsartikkel om OpenAIs publisering av over 700 matematiske preprints, oktober (gjengitt i vedlegget).
- NIH (2025). Supporting Fairness and Originality in NIH Research Applications (NOT-OD-25-132).
- RAND Europe (2026). AI in research: Mapping the terrain of adoption, attitudes and impact in the UK. Prosjektside; resultater ventes i desember 2026.
- Regjeringen (2025). Dette er Norges seks nye forskningssentre for kunstig intelligens.
- Science & Technology Australia (2026). AI use widespread among researchers but support urgently needed. Første resultater fra den nasjonale undersøkelsen, 20. oktober.
- Stanford HAI (2026). Artificial Intelligence Index Report 2026.
- Turekian, V. C. (2026). Platform science and the fragmentation of global scientific cooperation. *PNAS Nexus* (lederartikkel).