

Kortlægning af bedømmelsespraksis i gymnasiet i lyset af generativ AI



INDHOLD

Kortlægning af bedømmelsespraksis i gymnasiet i lyset af generativ AI

1	Resumé	5
<hr/>		
2	Indledning	10
2.1	Baggrund	10
2.2	Formål	11
2.3	Datagrundlag	12
2.4	Afgrænsning	14
2.5	Begrebsafklaring	14
2.6	Læsevejledning	15
<hr/>		
3	Litteraturopsamling – Bedømmelsespraksis i lyset af AI	16
3.1	Udfordringsbilledet	17
3.2	Greb til at håndtere generativ AI i bedømmelsessituationer	19
3.3	Persondata og etiske hensyn	21
<hr/>		
4	Brug og bedømmelse af AI til prøver på universitetet	22
4.1	Universiteternes retningslinjer for brug af generativ AI	22
4.2	Prøveformer på universiteterne i lyset af generativ AI	27

5	Bedømmelsespraksis og prøver i gymnasiet i lyset af AI	29
5.1	Konkrete initiativer	29
5.2	Udfordringer med bedømmelse	31
5.3	Bedømmeskriterier	33
5.4	Prøveformer	41
5.5	Rammevilkår	43

	Appendiks A – Litteraturliste	45
--	--------------------------------------	-----------

	Appendiks B – Figur og tabelbilag	47
--	--	-----------

1 Resumé

I denne rapport præsenterer Danmarks Evalueringsinstitut (EVA) resultaterne af en kortlægning af bedømmelsespraksisser i gymnasiet i lyset af generativ AI. Kortlægningen er gennemført på opdrag fra Styrelsen for Undervisning og Kvalitet (STUK).

Generativ kunstig intelligens (AI) er blevet en del af hverdagen på gymnasierne, siden udbredelsen af teknologien for alvor tog fart i slutningen af 2022 med lanceringen af ChatGPT. Det udfordrer særligt de hidtidige bedømmelsespraksisser i gymnasiet, der knytter sig til skriftlig fremstilling.

Fremkomsten af generativ AI stiller således gymnasierne i en udfordrende position ift. bedømmelsespraksisser, der dels bidrager til en valid bedømmelse af elevernes selvstændige faglige niveau og dels gør det muligt at bevare blikket for og understøtte elevernes læring, så de klædes på til eksamenssituationer, hvor de indtil videre ikke må benytte sig af teknologien. Grundet det generelle forbud mod anvendelse af generativ AI til prøver i gymnasiet er der på nuværende tidspunkt begrænset viden omkring, hvordan fremkomsten af generativ AI kan håndteres i denne forbindelse.

Denne rapport kortlægger på den baggrund erfaringer med bedømmelsespraksisser på de gymnasiale uddannelser i lyset af generativ AI. Kortlægningen bygger først og fremmest på gymnasielevere og -læreres erfaringer med bedømmelsessituationer ifm. undervisningen, da der kun findes sparsomme erfaringer med bedømmelse af generativ AI til prøver i gymnasiet. Erfaringerne med bedømmelsessituationer i undervisningen skal bidrage med viden om, hvad det er vigtigt at være opmærksom på i tilrettelæggelsen af prøverne i gymnasiet, hvis eksaminatorer og censorer skal have indblik i og mulighed for at bedømme elevernes selvstændige faglige niveau i lyset af generativ AI. Rapporten præsenterer derudover en litteraturopsamling af danske og internationale studier om bedømmelse af generativ AI i en uddannelseskontekst og fremlægger erfaringer med at tillade brug af generativ AI til prøver på danske universiteter.

Med afsæt i de forskellige datakilder tegner rapporten de udfordringer op, som generativ AI har medført ift. eksisterende bedømmelsespraksisser, efterfulgt af en udlægning af de forskellige bedømmelseskriterier, som undervisere i tiltagende grad lægger vægt på for at bevare blikket for elever og studerendes selvstændige faglige niveau. Endelig fremlægger rapporten de overvejelser, som de forskellige erfaringer giver anledning til ift. prøver i gymnasiet.

Rapporten henvender sig først og fremmest til Styrelsen for Undervisning og Kvalitet (STUK), som kan bruge kortlægningens resultater i sit videre arbejde med at udvikle prøver i gymnasiet i lyset af generativ AI. Rapporten kan derudover være relevant for ledere, lærere, eksaminatorer og censorer i gymnasiet som inspiration til deres møder med generativ AI i bedømmelsessituationer.

Datagrundlag

Kortlægningen, der præsenteres i rapporten, er baseret på de følgende datakilder:

- **Kortlægning af erfaringer på gymnasierne.** EVA har gennemført interview med ledere og lærere fra 12 udvalgte gymnasier på tværs af htx, hhx, stx og hf. Gymnasierne er udvalgt, fordi de på forskellig vis har særlig fokus på arbejdet med at integrere generativ AI ifm. opgaver og afleveringer. Interviewene har givet indsigt i ledere og læreres erfaringer med at stille og bedømme opgaver og forløb, hvor brug af generativ AI er tilladt.
- **Afdækning af erfaringer på universiteterne.** EVA har gennemført interview med repræsentanter fra syv universiteter om deres erfaringer med at udvikle og benytte retningslinjer for brug af generativ AI til udvalgte prøver.
- **Litteraturopsamling.** EVA har gennemført en litteraturopsamling af danske og internationale studier om brug og bedømmelse af generativ AI i uddannelsessektoren.
- **Sparringsgruppe.** En sparringsgruppe bestående af tre eksperter har bidraget med at ridse kortlægningens genstandsfelt op, anbefalet referencer til relevant litteratur inden for området og har desuden kvalificeret kortlægningens hovedpointer opsummeret i resuméet.

Resultater

Bedømmeskriterier i gymnasiet i lyset af generativ AI

Det fremgår af kortlægningen, at fremkomsten af generativ AI især udfordrer lærere, eksaminatorer og censorers grundlag for at bedømme elevernes selvstændige faglighed ifm. skriftlige opgaver, hvor eleverne kan generere hele eller delvise opgavebesvarelser ved brug af generativ AI. Kortlægningen viser i forlængelse heraf, at lærerne i gymnasiet benytter sig af en række bedømmelsesmæssige greb som reaktion på fremkomsten af generativ AI, som de oplever dels understøtter deres føling med elevernes læring og dels muliggør, at lærerne kan bedømme eleverne ud fra deres selvstændige faglige niveau, uafhængigt af om eleverne benytter sig af generativ AI eller ej.

De følgende afsnit udfolder kortlægningens resultater vedrørende, hvordan fremkomsten af generativ AI hhv. påvirker gymnasielærernes bedømmelsespraksis og får dem til at anvende bedømmeskriterier, som både knytter sig til elevernes grundfaglighed og til metakompetencer, der er relevante for elevernes almene dannelse og brug af teknologien.

Generativ AI udfordrer bedømmelse af elevernes selvstændige faglighed

Kortlægningen viser på tværs af litteraturopsamlingen samt erfaringerne fra universiteterne og de 12 udvalgte gymnasier, at udbredelsen af generativ AI i uddannelsessystemet udfordrer eksisterende praksisser for at bedømme elevernes selvstændige faglige niveau. Det gør sig især gældende ifm. bedømmelse af skriftlige produkter, hvor bedømmerne kan have svært ved at vurdere, hvilke hjælpemidler elever eller studerende har brugt for at nå frem til deres afleverede resultat, og dermed også om de rent faktisk bedømmer elevens eller den studerendes selvstændige præstation. Litteraturopsamlingen viser, at det både kan resultere i, at elever og studerendes ureglementerede

brug af generativ AI ikke bliver opdaget, men også at de risikerer at blive mistænkt for at have snydt gennem brug af generativ AI, selvom det ikke er tilfældet. Litteraturopsamlingen viser desuden, at usikkerheden generelt kan føre til, at undervisere, eksaminatorer og censorer bedømmer opgaver hårdere, når de har en mistanke om, at elever eller studerende har snydt gennem brug af generativ AI i deres besvarelse – også selvom det ikke kan bevises og ikke nødvendigvis er tilfældet.

Disse pointer går igen i interviewene med gymnasielærere, som forklarer, at de bruger meget tid og energi på at kontrollere, om elever har snydt ifm. skriftlige opgaver. Lærerne peger på, at det kan gå ud over fagligt dygtige og pligttopfyldende elever, der bevidst kan finde på at integrere fejl i deres skriftlige opgaver, fordi de er bange for at blive uretmæssigt anklaget for at have snydt.

Lærerne skruer derfor op for bestemte greb i deres bedømmelsespraksis for at styrke deres blik for elevernes selvstændige faglige niveau i lyset af generativ AI og forebygge den beskrevne mistillid, som de finder uhensigtsmæssig ift. at skabe et trygt læringsrum og en retfærdig bedømmelse af eleverne.

Lærere bedømmer elevernes faglighed ved at øge fokuset på grundfaglighed, arbejdsproces, mundtlighed og konstruktiv anvendelse af generativ AI

Det fremgår af kortlægningen, at de elementer, som lærerne benytter sig af for at styrke deres føling med og mulighed for at bedømme elevernes selvstændige faglige niveau, knytter sig til at skruer op for kendte greb i form af 1) forskellige øvelser uden hjælpemidler, 2) fokus på elevernes arbejdsproces, herunder deres brug af læringsunderstøttende prompts, og 3) mundtlige præsentationer pba. skriftlige opgaver.

Kortlægningen viser, at sammensætningen af greb varierer, alt efter hvilke kompetencer lærerne vil træne og bedømme tilegnelsen af hos eleverne. De forskellige øvelser uden hjælpemidler bliver brugt til at træne og bedømme elevernes grundfaglighed, mens elevernes ageren i opgaveprocesser bliver fremhævet som en måde at udvikle og bedømme mere avancerede faglige kompetencer, herunder metakompetencer såsom kritisk tænkning, refleksion over egen læring, nysgerrighed og grundighed. Endelig bliver de mundtlige præsentationer dels brugt til at træne og bedømme elevernes grundfaglighed, fx ift. sproglig korrekthed eller applicering af matematisk teori, og dels som opfølgning på skriftlige opgaver for at vurdere, om elevens mundtlige formidling af opgavens pointer indikerer, at der er blevet arbejdet selvstændigt med opgaven.

Det fremgår også af kortlægningens litteraturopsamling, at et øget fokus på mundtlighed, proces og prøver uden hjælpemidler kan være relevante veje at gå, hvis man vil opretholde valide bedømmelsespraksisser i lyset af generativ AI. Netop fordi det kan være svært at foretage en retvisende bedømmelse af det selvstændige faglige niveau, hvis bedømmerne kun har skriftlige opgavebesvarelser at forholde sig til uden indblik i, hvilke hjælpemidler eleverne har benyttet sig af. Det understreger ligeledes betydningen af tilsyn ved skriftlige stedprøver, så det er muligt at have blik for, om elever benytter sig af hjælpemidler, der ikke er tilladte ved den pågældende prøve.

På baggrund af kortlægningens litteraturopsamling og interview med gymnasielærere kan der opstilles følgende fire fokuspunkter ift. bedømmeskriterier i lyset af generativ AI, der kan understøtte, at lærere, eksaminatorer og censorer kan bedømme elevernes selvstændige faglige niveau.

Bedømmelseskriterier i lyset af generativ AI

1. Fokus på elevernes grundfaglighed

- *Bedømmelsesgrundlag:* Skriveøvelser med pen og papir eller test uden hjælpemidler og opgavebeskrivelser, der bevidst er udformet, så opgaven er vanskelig at besvare gennem brug af generativ AI.
- *Bedømmelsesfokus:* Basale faglige kompetencer.

2. Fokus på elevernes arbejdsproces

- *Bedømmelsesgrundlag:* Projektarbejde og åbne opgaver.
- *Bedømmelsesfokus:* Elevernes ageren i en opgaveproces med fokus på mere avancerede faglige kompetencer, herunder metakompetencer såsom kritisk tænkning, nysgerrighed og refleksion over egen læring.

3. Fokus på mundtlig formidling

- *Bedømmelsesgrundlag:* Mundtlige præsentationer og gruppedrøftelser eller videoer og podcasts som supplement til skriftlige produkter.
- *Bedømmelsesfokus:* Basale faglige kompetencer ift. mundtlig formidling, og om elevernes mundtlige præsentation indikerer, at de har arbejdet selvstændigt med deres skriftlige produkt.

4. Fokus på elevernes interaktion med generativ AI

- *Bedømmelsesgrundlag:* Bilag ifm. skriftlige opgaver, hvor eleverne vedlægger deres prompts og evt. deres refleksion over, hvordan de har brugt dem til at forbedre deres opgave.
- *Bedømmelsesfokus:* Elevernes brug af læringsunderstøttende prompts, herunder om elevernes prompts indikerer, at de har benyttet sig af generativ AI som en konstruktiv sparringspartner i udarbejdelsen af deres opgave eller som et redskab til at udlicitere udarbejdelsen af dele af eller hele opgaven.

Prøver i gymnasiet i lyset af generativ AI

På tværs af kortlægningens litteraturopsamling og interview med repræsentanter for universiteterne samt ledere og lærere på de 12 udvalgte gymnasier kan der opsummeres to bud på potentielle veje for skriftlige prøver i gymnasiet i lyset af generativ AI, der understøtter, at eksaminator og censor kan få indblik i og bedømme elevernes selvstændige faglige niveau. Det drejer sig hhv. om et øget fokus på mundtlige forsvar pba. skriftlige produkter med tilladt brug af digitale hjælpemidler og skriftlige stedprøver, der inkluderer elementer uden hjælpemidler.

- **Skriftlige stedprøver, der inkluderer elementer uden hjælpemidler.** Det fremgår af kortlægningen, at skriftlige stedprøver, der inkluderer elementer uden hjælpemidler, kan være en relevant reaktion på udbredelsen af generativ AI i nogle fag. Skriftlige stedprøver uden hjælpemidler giver eksaminator og censor et retvisende grundlag for at bedømme elevernes grundlæggende faglige niveau, samtidig med at denne prøveform kan give eleverne incitament til at lægge sig i selen for at tilegne sig basale kompetencer i faget, selvom de i princippet kunne udlicitere denne læring gennem brug af generativ AI.

- **Skriftlige prøveformer med mundtligt forsvar.** På baggrund af kortlægningen fremstår mundtlig opfølgning på skriftlige produkter som en meningsfuld måde at styrke validiteten i bedømmelsen af elevernes skriftlige arbejde med brug af digitale hjælpemidler, da det giver eksaminator og censor mulighed for at udprøve og bedømme, om eleverne har arbejdet selvstændigt med indholdet i deres skriftlige produkt.

Prøveformerne knytter sig til justeringer af de skriftlige eksaminer, da mundtlige prøver med både lang og kort forberedelsestid generelt beskrives som velegnede til at imødekomme udviklingen i generativ AI's udbredelse og kapacitet. Mundtlige prøvesituationer giver således eksaminator og censor mulighed for at spørge ind til og bedømme elevernes selvstændige forståelse af relevante faglige mål.

De to beskrevne prøveformer afspejler i vid udstrækning tendenserne, der ses i udviklingen af bedømmeskriterier efter fremkomsten af generativ AI, ligesom de afspejler ekspertgruppen for ChatGPT og andre digitale hjælpemidlers anbefaling om prøver, der skal gå på to ben og udprøve både med og uden digitale hjælpemidler.

Kompetencer i brug og bedømmelse af generativ AI

Kortlægningen viser på tværs af litteraturopsamlingen samt erfaringerne fra universiteterne og de udvalgte gymnasier, at det er et vigtigt opmærksomhedspunkt for skolerne, at lærere i fag, hvor generativ AI evt. bliver tilladt som del af den afsluttende prøve, har de rette kompetencer. Det indebærer, at lærerne skal have kompetencer til dels at kunne undervise eleverne i læringsunderstøttende brug af generativ AI, som de vil blive bedømt ud fra til eksamen, og dels at kunne tage højde for elevernes brug af teknologien som del af bedømmelsesgrundlaget til prøven.

De pågældende læreres kompetencer i brug og bedømmelse af generativ AI er på den måde centrale for at understøtte, at eleverne så vidt muligt kan gå til prøver på lige vilkår.

2 Indledning

Styrelsen for Undervisning og Kvalitet (STUK) har bedt Danmarks Evalueringsinstitut (EVA) om at foretage en kortlægning af bedømmelsespraksisser på tværs af de gymnasiale uddannelser (htx, hhx, stx og hf) i lyset af generativ kunstig intelligens (AI).

Kortlægningen henvender sig først og fremmest til STUK, som kan bruge den i sit videre arbejde med at udvikle prøverne i gymnasiet. Kortlægningen er derudover relevant for ledere, lærere, eksaminatorer og censorer i gymnasiet, der kan bruge kortlægningen som inspiration i deres møde med generativ AI i bedømmelsessituationer og i deres arbejde i øvrigt.

2.1 Baggrund

Siden fremkomsten og udviklingen af gratis og let tilgængelig generativ AI, såsom ChatGPT, begyndte at tage fart i slutningen af 2022, har teknologien været blandt de mest debatterede emner inden for uddannelsesområdet generelt og det gymnasiale område specifikt. Der er i gymnasiesektoren en udbredt oplevelse af, at generativ AI udgør et paradigmeskifte, bl.a. ift. lærernes føling med og bedømmelse af elevernes læring og selvstændige faglige niveau. Ifølge både ledere, lærere og elever i gymnasiet aktualiserer fremkomsten af generativ AI derfor udvikling af prøver, der kan understøtte en valid bedømmelse af elevernes indsats (se fx Danske Gymnasier, 2024; Gymnasieskolen, 2024a, 2024b; Ritzau, 2024). I dag er det ikke tilladt at anvende generativ AI ved langt de fleste prøver i gymnasiet.¹

Børne- og undervisningsministeren nedsatte i 2023 en ekspertgruppe, der havde til opgave at komme med anbefalinger til, hvordan ChatGPT og andre digitale hjælpemidler kan håndteres ved prøverne i folkeskolen såvel som på ungdoms- og voksenuddannelser. Ekspertgruppen var optaget af at flytte fokus på generativ AI fra snyd til læring, så den tillidsfulde lærer-elev-relation styrkes gennem en fælles nysgerrighed på teknologiens potentialer.

Ekspertgruppen har i deres anbefalinger taget udgangspunkt i, at generativ AI er til stede i elevernes hverdag, og at de under deres uddannelse skal lære at bruge teknologien reflekteret og beherske den på linje med øvrige tilgængelige værktøjer. Ekspertgruppen anbefaler derfor, at der fremover både skal være prøver uden digitale hjælpemidler og prøver, hvor de kan indgå. Det skal sikre,

¹ Fagene informatik B og C, IT A, programmering B og C, teknikfag A samt digitalt design og udvikling er dog undtaget, da der i disse fag vurderes at være særlige faglige grunde, der taler for brug af AI i prøvesituationer.

at det både er muligt at udprøve elevernes basale færdigheder uden hjælpemidler og mere avancerede kompetencer, hvor digitale hjælpemidler kan inddrages (BUVM, 2024).

På nuværende tidspunkt findes der ikke meget viden om, hvordan man kan integrere generativ AI i prøver i gymnasiet på en meningsfuld måde, og hvordan bedømmere arbejder med at bevare blikket for og vurdere elevens selvstændige faglige niveau mhp. at sikre en valid bedømmelse i lyset af generativ AI.

2.2 Formål

Formålet med kortlægningen er at styrke vidensgrundlaget for STUK's videre arbejde med at udvikle prøverne i gymnasiet i en tid, hvor generativ AI er en faktor, man må tage højde for. Ifm. dette udviklingsarbejde er det relevant at kortlægge, hvad lærere i gymnasiet på nuværende tidspunkt har af konkrete erfaringer med at bedømme elevernes selvstændige faglige indsats i lyset af generativ AI. Disse erfaringer kan bidrage med viden om, hvad man kan opstille af relevante bedømmelseskriterier efter fremkomsten af generativ AI, og hvordan prøver kan tilrettelægges for at understøtte bedømmernes brug af disse kriterier.

Ud over ledere og læreres konkrete erfaringer med bedømmelse i lyset af generativ AI på de gymnasiale uddannelser bygger kortlægningen på en litteraturopsamling af bedømmelsespraksisser efter fremkomsten af generativ AI i uddannelsessektoren bredt set og en afdækning af danske universiteters erfaringer med at udvikle og benytte retningslinjer for brug af generativ AI til udvalgte prøver.

Kortlægningen besvarer de følgende fire overordnede undersøgelsesspørgsmål:

1. Hvilke erfaringer findes der med brug og bedømmelse af generativ AI til prøver og ifm. opgaver og afleveringer som del af undervisningen på de gymnasiale uddannelser, herunder udarbejdelse af bedømmelseskriterier, der understøtter udprøvning af elevernes selvstændige faglighed og lige prøvevilkår for alle?
2. Hvad er erfaringerne, og hvilke forskelle kan der identificeres, ift. brug og bedømmelse af generativ AI til prøver og ifm. opgaver og afleveringer som del af undervisningen på tværs af fagområder og prøveformer?
3. Hvad findes der af erfaringer fra dansk og international forskning samt de lange videregående uddannelser ift. brug af generativ AI til prøver og bedømmelse heraf, som man kan lade sig inspirere af inden for det gymnasiale område?
4. Hvilke udfordringer og potentialer peger erfaringerne samlet set på, herunder erfaringer fra forskning, universiteterne og lokale tiltag og arbejdsformer på gymnasieområdet, hvor man har tilladt brug af generativ AI til udvalgte projekter eller skriftlige afleveringer og opgaver i undervisningen?

Det er en grundlæggende præmis for kortlægningen, at bedømmelseskriterier og prøver i gymnasiet efter fremkomsten af generativ AI er et genstandsfelt, der er opstået pludseligt og med en rivende udvikling inden for de seneste to år. Det afspejler sig i den tilgængelige litteratur om emnet og i de erfaringer, som ledere og lærere på gymnasierne har gjort sig indtil nu, der varierer, alt efter hvor langt de enkelte gymnasier er nået med at igangsætte initiativer, der skal imødekomme udbredelsen af teknologien. Det betyder også, at kortlægningens resultater om bedømmelsespraksis i lyset af generativ AI er foreløbige og kan udvikle sig, i takt med at teknologiens kapacitet og lærernes og elevernes erfaringer med teknologien udvikler sig.

2.3 Datagrundlag

Kortlægningen bygger på kvalitative interview med ledere og lærere på 12 gymnasiale uddannelser på tværs af htx, hhx, stx og hf, interview med repræsentanter fra universiteterne samt en opsamling af relevant dansk og international litteratur om brug og bedømmelse af generativ AI. Der har derudover været tilknyttet en sparringsgruppe til kortlægningen.

2.3.1 Sparringsgruppe

EVA har ifm. kortlægningen nedsat en sparringsgruppe bestående af tre eksperter med indsigt i prøveformer og/eller generativ AI i relation til uddannelses- og gymnasieområdet. Sparringsgruppen har dels bidraget til kortlægningen ved at gøre os opmærksomme på aktuelle problemstillinger og relevant litteratur inden for området og dels ved at kvalificere kortlægningens hovedpointer.

Konkret har vi afholdt bilaterale telefonmøder med de tre deltagere i sparringsgruppen, hvor de har delt deres perspektiver på potentialer, udfordringer og opmærksomhedspunkter ved tilladt brug af generativ AI til prøver i gymnasiet og i uddannelsessektoren generelt. Derudover har deltagerne i sparringsgruppen hver især peget på op til tre forskningsartikler eller andre undersøgelser om emnet, de finder særligt interessante. Endelig har de læst og givet feedback på kortlægningens hovedpointer opsummeret i resuméet.

Sparringsgruppe

Kortlægningens sparringsgruppe består af de følgende tre eksperter inden for prøveformer, uddannelse og generativ AI:

Jens Tofteskov, underviser og kursusansvarlig ved Copenhagen Business School

- Har forsket i og udgivet bøger om eksamen og eksamensformer.

Kristine Bundgaard, lektor i it, organisation og sprogteknologi ved Aalborg Universitet

- Har ekspertise i sprogteknologi med fokus på integration og implikationer af generativ AI i en uddannelseskontekst, herunder sprogteknologiens betydning for fremmedsprogundervisning.

Tine Wirenfeldt Jensen, ph.d. i læring og uddannelse og ejer af METoDO

- Har ekspertise i AI og læring med særligt fokus på akademisk skrivning i lyset af den digitale udvikling og var medlem af ekspertgruppen for ChatGPT og andre digitale hjælpemidler.

Sparringsgruppen har haft en rådgivende funktion. Kortlægningens indhold og resultater er EVA's ansvar.

2.3.2 Litteraturopsamling

Kortlægningens litteraturopsamling bidrager til et indledende overblik over, hvad der allerede findes af viden om udfordringer og potentialer ifm. bedømmelse og prøver i lyset af generativ AI.

Denne viden bidrager som en overordnet forståelsesramme for kortlægningen og har derudover ligget til grund for udarbejdelsen af spørgeguides til kortlægningens kvalitative interview med ledere og lærere på de gymnasiale uddannelser og repræsentanter for universiteterne.

Litteraturopsamlingen er dels baseret på de forskningsartikler og undersøgelser, som sparringsgruppens tre medlemmer har udpeget, og dels på åbne søgninger via Google og Google Scholar samt forskellige forskningsdatabaser. Litteraturopsamlingen består derudover af relevante artikler og undersøgelser, som er lokaliseret pba. referencelisterne i den allerede identificerede litteratur.

2.3.3 Afdækning af erfaringer fra universiteterne

Kortlægningens afdækning af erfaringer med brug og bedømmelse af generativ AI til prøver på universiteterne bidrager til at identificere foreløbige erfaringer og relevante overvejelser, som kan være til inspiration i udviklingen af retningslinjer for brug af generativ AI til prøver i gymnasiet.

Vi har gennemført semistrukturerede interview á ca. 30-40 minutters varighed med repræsentanter for syv danske universiteter (se figur B.1 i appendiks B for en oversigt over universiteterne). De interviewede repræsentanter er udvalgt som informanter, fordi de har været del af processen med at udvikle retningslinjer for brug af generativ AI til prøver på det pågældende universitet. Interviewene har belyst repræsentanternes oplevelser og perspektiver på potentialer og udfordringer ved at til-lade brug af generativ AI ifm. prøver. Interviewene har derudover sat fokus på repræsentanternes beskrivelse af deres universitets konkrete erfaringer med at udarbejde og anvende retningslinjerne for brug af generativ AI til udvalgte prøver.

Ud over interviewene baserer afdækningen sig desuden på dokumenter og materialer fra de syv universiteter, herunder de udviklede retningslinjer og projektbeskrivelser.

2.3.4 Interview med ledere og lærere fra gymnasiale uddannelser

For at opnå indblik i erfaringerne med brug og bedømmelse af generativ AI på tværs af de gymnasiale uddannelser har vi foretaget interview med ledere og lærere fra i alt 12 gymnasier (se figur B.2 i appendiks B for en oversigt over gymnasierne). Gymnasierne er udvalgt, fordi de har gjort sig erfaringer og igangsat særlige initiativer, hvor de aktivt arbejder med at udvikle en praksis, der tager højde for fremkomsten af generativ AI.

Vi har udvalgt de 12 gymnasier pba. dialog med centrale aktører inden for det gymnasiale område: Gymnasieskolernes Lærforening (GL), Danske Gymnasier, Danske Erhvervsskoler og -Gymnasier

(DEG), Kontor for Prøver, Eksamen og Test samt lærings- og fagkonsulenterne inden for det gymnasiale område i STUK. Med afsæt i denne dialog har vi udarbejdet en bruttoliste med 22 gymnasier, der har søsat særlige initiativer med fokus på brug af generativ AI i bedømmelse af produkter, hvor eleverne har brugt teknologien som del af deres arbejdsproces. Med afsæt i denne bruttoliste udvalgte vi 12 gymnasier til kvalitative interview ud fra et hensyn til, at de skulle repræsentere en variation ift. gymnasietype, geografisk placering og størrelse.

Vi har efterfølgende gennemført virtuelle, semistrukturerede interview á ca. 45 minutters varighed med ledelsesrepræsentanter og semistrukturerede gruppeinterview á ca. 60 minutters varighed med lærere fra de 12 gymnasier. Lederne har bidraget med overordnede perspektiver på kortlægningens genstandsfelt, herunder beskrivelser af de igangsatte initiativer relateret til generativ AI på deres gymnasium. Lærerne har bidraget med beskrivelser af deres erfaringer med at stille og bedømme opgaver, der integrerer brug af generativ AI. Ud over interviewene med ledere og lærere baserer kortlægningen sig desuden på dokumenter og materialer fra de 12 gymnasier, herunder projektbeskrivelser samt beskrivelser af opgaver og undervisningsforløb.

2.4 Afgrænsning

Kortlægningen afgrænser sig til at fokusere på erfaringer forbundet med *bedømmelsespraksisser* i lyset af generativ AI, som kan anvendes i det videre arbejde med at udvikle prøver i gymnasiet, der tager højde for teknologiens fremkomst.

Det er dog væsentligt for forståelsen af kortlægningens pointer, at bedømmelse og prøver i gymnasiet hænger tæt sammen med de gymnasiale uddannelsers *formål* og fagenes *læreplaner*, idet prøver og eksaminer skal sikre ensartethed og sammenhæng mellem netop uddannelsens formål, læreplaner og bedømmelse (BUVM, 2024b). Med andre ord har rammerne for og bedømmeskriterierne til prøver og eksaminer betydning for lærernes til- og fravalg i undervisningen ift. at skabe sammenhæng mellem, hvad eleverne lærer, og hvad de forventes at kunne og bliver bedømt ud fra til eksaminerne i de enkelte fag.

Selvom kortlægningen først og fremmest fokuserer på gymnasieledere og -læreres erfaringer med at afprøve valide bedømmelsespraksisser i lyset af generativ AI, og hvilke overvejelser det giver anledning til ift. prøver i gymnasiet, tager ledernes og lærernes erfaringer udgangspunkt i bedømmelsessituationer i undervisningen. Det skyldes som beskrevet, at der indtil videre kun findes sparsomme erfaringer med bedømmelse af generativ AI til prøver i gymnasiet, hvor det med enkelte undtagelser ikke er tilladt at benytte sig af teknologien.

2.5 Begrebsafklaring

I kortlægningen bruger vi gennemgående udtrykket 'generativ AI', der er en paraplybetegnelse for generative sprogmodeller, der kan læse og forstå tekst og på denne baggrund producere tekst, billeder, videoer eller andet data med afsæt i kunstig intelligens. Det kan fx dreje sig om ChatGPT, Copilot eller Midjourney. I kortlægningen benytter vi os desuden af udtrykket 'prompts', der dækker over den tekst, man fodrer sin generative AI-model med mhp. at få løst en specifik opgave.

Kortlægningen læner sig derudover op ad et begrebsapparat for elevers brug af generativ AI, som Christian Dalsgaard, lektor ved DPU på Aarhus Universitet, og Christopher Neil Prilop, lektor ved Centre for Educational Development (CED) på Aarhus Universitet, har udviklet som del af deres følgeligforskning til et projekt på Aarhus Business College, Herning Gymnasium og Uddannelsescenter Ringkøbing-Skjern vedrørende skriftlighed og læring i lyset af generativ AI. Her identificerer de seks arbejdsmetoder, som eleverne i projektet benytter sig af ifm. deres brug af generativ AI, i form af forklaring, feedback, revision, inspiration, forberedelse og udarbejdelse, som kan placere sig i et udfaldsrum mellem hhv. *udlicitering til* og *partnerskab med* generativ AI (Dalsgaard & Prilop, 2024).

Med dette begrebsapparat argumenterer Dalsgaard og Prilop for, at ingen af de seks AI-medierede arbejdsformer i sig selv er læringshæmmende eller -fremmende. Dalsgaard og Prilop peger således på, at alle seks arbejdsformer er læringshæmmende, hvis de bruges til at udlicitere læring til generativ AI, hvor en elev fx kopierer et svar direkte fra en sprogmodel ind i sin egen opgave. Omvendt peger de også på, at alle seks arbejdsformer kan være læringsfremmende, hvis de udspringer af et partnerskab mellem eleven og generativ AI, hvor teknologien bidrager med eksempler, viden eller feedback, som eleven forholder sig kritisk og selvstændigt til (Dalsgaard & Prilop, 2024).

2.6 Læsevejledning

Kortlægningen indeholder foruden resumé og dette indledende kapitel yderligere tre kapitler:

Kapitel 3 præsenterer litteraturopsamlingen og giver et overblik over viden om udfordringer og potentialer ved at integrere og bedømme brug af generativ AI til prøver i uddannelsessektoren.

Kapitel 4 præsenterer afdækningen af eksisterende erfaringer med at tillade brug af generativ AI til prøver på de lange videregående uddannelser og giver indblik i syv universiteters retningslinjer og overvejelser i den sammenhæng.

Kapitel 5 præsenterer bedømmelsespraksisser, lærere i gymnasiet benytter sig af for at bevare deres føling med og mulighed for at bedømme elevernes læring og selvstændige faglige niveau i lyset af generativ AI. Kapitlet fremlægger derudover de udfordringer, gymnasieledere og -lærere oplever, at fremkomsten af generativ AI har afstedkommet ift. bedømmelse af elevernes faglighed, og de perspektiver, som deres hidtidige erfaringer med bedømmelse i lyset af generativ AI giver anledning til ift. prøver i gymnasiet.

3 Litteraturopsamling – Bedømmelsespraksis i lyset af AI

I dette kapitel sammenfatter vi relevant litteratur vedrørende bedømmelsespraksisser i lyset af generativ AI, og hvad det betyder for udviklingen af prøver. Litteraturen er identificeret gennem en kombination af en overordnet litteratursøgning med fokus på forskning og anden systematisk indsamlet viden om kortlægningens genstandsfelt, anbefalinger fra projektets sparringsgruppe og relevante artikler og øvrige udgivelser fra de udvalgte kilders referencelister.

Litteraturopsamlingen viser, at det på nuværende tidspunkt er begrænset, hvad der findes af studier, som specifikt beskæftiger sig med bedømmeskriterier og prøver i lyset af udbredelsen af generativ AI på de gymnasiale uddannelser i en dansk kontekst. Litteraturopsamlingen er derfor udvidet til at dække uddannelsesområdet bredt set, ligesom den også inkluderer internationale studier. Selv med denne udvidelse er feltet af relevant litteratur med et bedømmelses- og prøvefokus imidlertid begrænset.

Det fremgår af kapitlet, at litteraturen samlet set peger på såvel udfordringer som potentialer knyttet til bedømmelsespraksisser i lyset af generativ AI. Hvad angår udfordringsbilledet, bekræfter litteraturen gymnasielæreres oplevelse af, at det kan være vanskeligt at skelne mellem AI- og elevgenererede tekster. Det kan ifølge litteraturen medføre en latent tvivl hos lærerne om elevernes selvstændige indsats, der skaber usikkerhed i bedømmelsessituationen og en risiko for hårdere bedømmelse af elevens arbejde. Litteraturen peger desuden på, at der er en risiko for, at elevers brug af generativ AI kan føre til dårligere bedømmelser i prøvesituationer, hvis ikke de er klædt på til at benytte sig af teknologien som et læringsunderstøttende hjælpemiddel (se afsnit 3.1 for referencer til den specifikke litteratur).

Kapitlet beskriver i forlængelse heraf, hvordan usikkerheden i bedømmelsessituationen ifølge litteraturen kan imødekommes ved at lade mundtlighed, herunder mundtlig opfølgning på skriftlige prøver, prøver uden hjælpemidler og opgavebeskrivelser, der forebygger uselvstændig brug af generativ AI, fylde mere i udprøvningen og bedømmelsen af elever, end tilfældet er i dag (se afsnit 3.2 for referencer til den specifikke litteratur).

Endelig fremgår det af kapitlet, at litteraturen peger på en række etiske og datamæssige udfordringer ved at tillade brug af generativ AI til prøver, herunder dataejerskab og risiko for bias i algoritmerne (se afsnit 3.3 for referencer til den specifikke litteratur).

3.1 Udfordringsbilledet

Den identificerede litteratur peger på en række udfordringer forbundet med fremkomsten af generativ AI ifm. undervisere, eksaminatorer og censorers bedømmelsespraksis. Overordnet set består udfordringsbilledet i, at generativ AI udfordrer hidtidige tilgange til at bedømme elevernes selvstændige faglige niveau. Udfordringerne med bedømmelsen af elevernes selvstændige faglige niveau handler i litteraturen om tre ting, der vil blive udfoldet i de følgende afsnit:

1. Det kan være svært at skelne mellem AI- og elevgenererede tekster
2. Tvivl udfordrer en fair bedømmelse
3. Brug af generativ AI kan undergrave elevens læreproces og resultere i dårligere bedømmelse.

3.1.1 Svært at skelne mellem AI- og elevgenererede tekster

En del af den identificerede litteratur handler om, hvorvidt og hvordan undervisere, eksaminatorer og censorer kan identificere, om en tekst er genereret af generativ AI eller en elev (Scarfe, Watcham, Clarke & Roesch, 2024; Blom, Holsting & Svendsen, 2024; Farazouli, Cerratto-Pargman, Bolander-Lakskov & McGrath, 2023). Det kan i yderste konsekvens indebære en vurdering af, om elevens arbejde er et resultat af snyd eller et udtryk for deres selvstændige faglige niveau. Den overordnede konklusion på tværs af studierne er, at denne skelnen er særdeles vanskelig.

En dansk undersøgelse, der gennemfører en sproglig ophavsanalyse af tekster skrevet af hhv. ChatGPT og danskstuderende, viser, at studerende på den ene side kan bruge chatbotter til at skrive opgaver, uden at det kan opdages, og at de studerende på den anden side også kan risikere at blive mistænkt for at have brugt chatbotter, selvom det ikke er tilfældet (Blom, Holsting & Svendsen, 2024). Undersøgelsen finder endvidere, at der er visse særtræk ved ChatGPT's skrivestil, men at det er muligt at sløre disse, fx ved bevidst brug af prompts, hvilket gør det svært at spore snyd (Blom, Holsting & Svendsen, 2024).

Et andet studie viser tilsvarende, at det kan være noget nær umuligt at udpege AI-genererede produkter. Undersøgelsen havde til formål at afdække, hvorvidt eksamensbesvarelser udarbejdet af generativ AI blev opdaget, når de blev indleveret sammen med selvstændigt udarbejdede eksamensbesvarelser blandt bachelorstuderende i psykologi ved et britisk universitet (Scarfe, Watcham, Clarke & Roesch, 2024). Her viser resultaterne, at 94 % af de indleverede AI-genererede besvarelser forblev uopdagede, og at besvarelserne i gennemsnit scorede højere karakterer end de studerendes selvstændige besvarelser (Scarfe, Watcham, Clarke & Roesch, 2024).

Resultaterne i disse studier illustrerer vanskeligheden med at skelne mellem AI- og elevgenererede tekster. Fundene understreger dermed de udfordringer, der kan være forbundet med at bedømme skriftlige produkter, der står alene, efter fremkomsten af generativ AI.

3.1.2 Tvivl udfordrer en fair bedømmelse

I litteraturopsamlingen indgår også en undersøgelse fra 2023, der ser nærmere på, hvordan udfordringerne med at skelne mellem AI- og elevgenererede tekster påvirker læreres læsning og bedømmelse af produkter, der potentielt kan være udarbejdet af generativ AI. Det fremgår af undersøgelse, at den latente tvivl kan føre til en hårdere bedømmelse af produkterne (Farazouli, Cerratto-Pargman, Bolander-Lakskov & McGrath, 2023).

Undersøgelsen vedrører bedømmelsen af eksamensopgaver på bachelorniveau. Underviserne i undersøgelsen har blindt bedømt besvarelserne, som er udarbejdet af hhv. studerende eller ChatGPT. Efterfølgende har underviserne deltaget i fokusgruppeinterview om kvaliteten af de eksamensopgaver, de har bedømt. Undersøgelsen finder, at underviserne generelt fattede mistanke om brug af ChatGPT under bedømmelsen. Det fik dem til at vurdere besvarelserne mere kritisk end normalt. Mistanken om brug af ChatGPT bevirkede desuden, at underviserne tolkede nogle meneskelige sproglige fejl som udtryk for, at teksten var skrevet af generativ AI. Underviserne vished om, at nogle eksamensopgaver kunne være udarbejdet gennem generativ AI, ændrede således deres måde at læse og bedømme eksamensbesvarelserne på (Farazouli, Cerratto-Pargman, Bolander-Lakskov & McGrath, 2023).

Den identificerede litteratur peger altså på, at den svære skelnen mellem AI- og elevgenererede tekster grundlæggende kan udfordre en fair bedømmelse af elevernes arbejde.

3.1.3 Brug af generativ AI kan resultere i dårligere bedømmelse

Litteraturen peger desuden på, at brugen af generativ AI kan undergrave elevernes læreproces, hvis de udliciterer deres læring til generativ AI (Dalsgaard & Prilop, 2024). Det kan resultere i, at de får en dårligere bedømmelse til prøver end elever, der evner at benytte teknologien som et lærings-understøttende hjælpemiddel.

En undersøgelse sammenligner high school-elevs læring i matematik ved brug af to forskellige sprogmodeller designet ud fra ChatGPT (Bastani, Bastani, Sungu, Ge, Kabaci & Mariman, 2024). Den ene sprogmodel ligner den almindelige ChatGPT, mens den anden har prompts udviklet til at understøtte elevernes læring. Undersøgelsen viser, at de elever, der har haft adgang til de to sprogmodeller, klarer sig væsentligt bedre, end elever, der ikke har haft adgang til de samme hjælpemidler (Bastani, Bastani, Sungu, Ge, Kabaci & Mariman, 2024, s. 1-2). Det fremgår dog også af undersøgelsen, at de elever, der har haft adgang til den ordinære ChatGPT-lignende sprogmodel, klarer sig dårligere end de øvrige elever, hvis de bliver frataget deres adgang til sprogmodellen. Det gør sig imidlertid ikke gældende i samme grad for de elever, der har haft adgang til den version af sprogmodellen, der er designet med prompts mhp. at understøtte elevernes læreproces. Undersøgelsen peger således på, at elevernes varige læringsudbytte svækkes, hvis de benytter sig ureflekteret af generativ AI (Bastani, Bastani, Sungu, Ge, Kabaci & Mariman, 2024).

I et andet studie sammenligner forfatterne eksamensresultater i et kursus af et semesters varighed blandt to grupper af studerende (Weeks, Voshaar, Plate & Zimmermann, 2024). Den ene gruppe har

brugt generativ AI til at skrive opgaver i løbet af kurset, mens den anden gruppe ikke har. Undersøgelsen viser, at de studerende, der løbende har gjort brug af generativ AI ifm. opgaveskrivning, typisk scorer lavere i selve eksamenssituationen, hvor de studerende ikke må bruge generativ AI, end studerende, der ikke har brugt generativ AI i løbet af kurset (Wecks, Voshaar, Plate & Zimmermann, 2024).

Litteraturen peger således på, at der er en risiko for, at brugen af generativ AI kan ske på bekostning af elevernes læring og resultere i en dårligere bedømmelse af eleverne, hvis ikke brugen af teknologien er didaktisk tilrettelagt til at understøtte deres læring.

3.2 Greb til at håndtere generativ AI i bedømmelsessituationer

På tværs af den identificerede litteratur kan der sammenfattes tre primære greb, som kan være med til at håndtere det beskrevne udfordringsbillede. Grebene, der kan hjælpe undervisere, eksaminatorer og censorer med at få blik for og bedømme elevernes selvstændige faglige niveau, handler i litteraturen om tre forhold, der udfoldes i de følgende afsnit. Det knytter sig til at:

1. Sætte mere fokus på mundtlighed som supplement til skriftlige produkter
2. Skabe bedømmelsessituationer uden brug af hjælpemidler
3. Udarbejde opgavebeskrivelser, der forebygger uselvstændig brug af generativ AI.

3.2.1 Fokus på mundtlighed

Det er en gennemgående pointe i litteraturen, at et øget fokus på mundtlighed som supplement til skriftlige produkter i udprøvningen af elever kan være et nyttigt greb til at bevare blikket for og bedømme elevernes selvstændige faglige niveau efter fremkomsten af generativ AI. Netop fordi det kan være svært at skelne AI-genererede tekster fra elevgenererede tekster, hvis man kun har det skriftlige produkt at forholde sig til.

Et studie argumenterer fx for at lade mundtlighed fylde endnu mere i sprogfag, da en integration mellem mundtlig og skriftlig bedømmelse vil gøre det muligt at tillade brug af generativ AI til prøver, samtidig med at man som underviser, eksaminator eller censor kan bedømme elevernes selvstændige faglige niveau (Caviglia, Dalsgaard, Boie & Thomsen, 2021, s. 30, 37).

Et andet studie argumenterer tilsvarende for et øget fokus på mundtlighed og mindre fokus på skriftligt hjemmearbejde som modsvar til fremkomsten af generativ AI (Chiu, 2023, s. 11). Studiet er baseret på en spørgeskemaundersøgelse og fokusgruppeinterview blandt lærere og ledere om deres holdninger til ChatGPT og Midjourney. Med afsæt i analysen af disse data anbefaler studiet, at elever kan bruge generativ AI til at få skriftlig information og sparring, men at de samtidig må være i stand til mundtligt at forklare det, de har lært gennem deres brug af teknologien (Chiu, 2023).

I et studie fra 2023 peger forfatterne derudover på, at et øget fokus på mundtlig udprøvning, både som prøveform i sig selv og som opfølgning på skriftligt arbejde, kan være en vej til at få blik for og bedømme elevers evne til kritisk tænkning i lyset af generativ AI. Den mundtlige prøvesituation forudsætter således, at eleverne er i stand til at reflektere over og forholde sig kritisk til pensum og/eller valg og fravalg i deres opgaveproces (Mao, Chen & Lui, 2023, s. 62).

3.2.2 Bedømmelsessituationer uden brug af hjælpemidler

En del af den identificerede litteratur peger også på, at klassiske prøveformater som skriftlige og mundtlige stedprøver uden hjælpemidler kan være et greb til at skabe bedømmelsessituationer, der muliggør bedømmelse af elevernes selvstændige faglige niveau i lyset af generativ AI (Blom, Holsting & Svendsen, 2024; Scarfe, Watcham, Clarke & Roesch, 2024; Lo, 2023; Milano, McGrane & Leonelli, 2023).

Flere af studierne argumenterer imidlertid også for, at disse prøveformer uden hjælpemidler ikke er meningsfulde i alle situationer. Studierne peger fx på, at stedprøver uden hjælpemidler ikke er hensigtsmæssige ifm. større, skriftlige opgaver, der løber over en længere periode, eller i situationer, hvor man ønsker at bedømme arbejdet ud fra mere komplekse læringsmål, der ikke kan demonstreres mundtligt eller ”på stedet” (Blom, Holsting & Svendsen, 2024; Milano, McGrane & Leonelli, 2023).

Perkins, Furze, Roe og McVaugh (2024) foreslår i en artikel en overordnet model for, hvordan generativ AI kan integreres i bedømmelsessituationer på en etisk forsvarlig måde. Modellen består af fem forskellige niveauer, hvor brug af generativ AI i stigende grad tillades og inkorporeres. Det første niveau dækker over bedømmelsessituationer, hvor AI er fuldstændig forbudt. Artiklens forfattere beskriver, at det kan være en passende eksamensform, når det er afgørende at bedømme, om de studerende kun gør brug af deres egen forståelse og evner, eller når brugen af generativ AI ikke kan lade sig gøre. Artiklen foreslår, at de konkrete prøveformer, der ikke tillader brug af generativ AI, fx kan være skriftlige stedprøver og mundtlige eksaminer (Perkins, Furze, Roe & McVaugh, 2024).

Artiklen peger derudover på, at prøver, hvor brug af generativ AI ikke er tilladt, bør finde sted under grundig supervision, da det selv med redskaber udviklet til at identificere brug af generativ AI kan være svært at spore brug af teknologien (Perkins, Furze, Roe & McVaugh, 2024).

3.2.3 Opgavebeskrivelser, der forebygger uselvstændig brug af generativ AI

Endelig peger litteraturopsamlingen også på, at opgavebeskrivelser, der forebygger uselvstændig og ureflekteret brug af generativ AI, kan være et greb til at bevare blikket for og bedømme elevers selvstændige faglige niveau i lyset af teknologiens fremkomst.

Et litteraturstudie om ChatGPT’s indflydelse på uddannelse identificerer en række tiltag, der kan være med til at begrænse snyd gennem brug af ChatGPT og lignende sprogmodeller (Lo, 2023). De foreslåede tiltag knytter sig til opgavebeskrivelser, identifikation af tekst udarbejdet af sprogmodeller og institutionelle politikker (Lo, 2023, s. 9). Ift. opgavebeskrivelser (task design) kan det handle om at stille opgaver, der kræver analyse frem for genkaldelse af tillært viden. Et andet greb

kan være at inkludere billeder eller andre multimedieressourcer i eksamensspørgsmålene, hvilket gør det sværere at udarbejde en uselvstændig besvarelse gennem brug af generativ AI. Endelig kan opgavedesignet handle om at afholde stedprøver, hvor det ikke er tilladt at bruge ChatGPT i tråd med pointerne i afsnit 3.2.2 (Lo, 2023).

3.3 Persondata og etiske hensyn

Den identificerede litteratur peger ligeledes på en række etiske og datamæssige udfordringer forbundet med at bruge generativ AI til prøver. Det drejer sig bl.a. om spørgsmål vedrørende dataejerskab, aktørers mulighed for samtykke, retten til privatliv samt risikoen for, at bias i algoritmerne, eksempelvis ift. køn og etnicitet, kan underminere basale menneskerettigheder (UNESCO, 2021, s. 20; Mao, Chen & Liu, 2023).

I situationer, hvor man tillader brug af generativ AI til prøver, er det derfor nødvendigt med overvejelser om behandling af den data, brugen involverer.

4 Brug og bedømmelse af AI til prøver på universitetet

I dette kapitel udfolder vi erfaringer med at tillade brug af generativ AI til prøver på landets universiteter. Formålet med kapitlet er at formidle potentialer og udfordringer, som man kan hente inspiration fra ifm. drøftelser om at tillade brug af generativ AI til udvalgte prøver i gymnasiet. Kapitlet bygger på kvalitative interview med repræsentanter for syv universiteter, der alle har været involveret i universiteternes arbejde med at udvikle retningslinjer for brug af generativ AI til prøver.

Kapitlet beskriver først de syv universiteters retningslinjer for brug af generativ AI til prøver samt deres erfaringer med de gældende retningslinjer. Dernæst præsenterer kapitlet universiteternes perspektiver på, hvad fremkomsten af generativ AI betyder for den måde, de tænker eksaminer på.

Det fremgår af kapitlet, at universiteterne har vedtaget retningslinjer for de studerendes brug af generativ AI ifm. eksaminer, som alle giver mulighed for, at de enkelte studieområder kan tilpasse retningslinjerne til deres specifikke fagområde og eksamensformer. Det er desuden fælles for universiteternes retningslinjer, at de studerendes brug af generativ AI ifm. skriftlige eksaminer skal indebære en form for deklARATION af, hvor teknologien er brugt. Universiteterne benytter sig derudover fortsat af skriftlige stedprøver i skærmet form eller uden hjælpemidler. Universiteternes foreløbige erfaringer viser, at det er en stor opgave at inkorporere generativ AI ifm. eksaminer, men at de til gengæld bruger færre ressourcer på at kontrollere for snyd med brug af AI-genereret materiale.

I fortolkningen af disse pointer er det imidlertid væsentligt at holde sig for øje, at erfaringer fra universiteterne ikke kan overføres 1:1 til det gymnasiale område. Det handler bl.a. om, at det er nogle andre, mere avancerede læringsmål, studerende skal lære og bliver bedømt ud fra på universiteterne, hvilket har betydning for, hvordan man kan udprøve deres tilegnelse af disse kompetencer.

4.1 Universiteternes retningslinjer for brug af generativ AI

De følgende afsnit sætter fokus på, hvad de udvalgte universiteter konkret har gjort ift. at udvikle retningslinjer for brug af generativ AI til prøver, og hvilke erfaringer de hidtil har gjort sig med de opstillede retningslinjer.

4.1.1 Hvad har universiteterne gjort?

På tværs af repræsentanterne for de syv universiteter er der en fælles forståelse af, at man er nødt til at forholde sig til generativ AI og integrere teknologien i opgaver og eksaminer. De interviewede

repræsentanter peger på forskellige årsager til dette: Dels skal de som universiteter klæde de studerende på til et arbejdsmarked, hvor teknologien bliver anvendt, og dels er erfaringen, at universiteterne alligevel ikke kan håndhæve et totalforbud. Repræsentanterne oplever, at mange studerende allerede anvender generativ AI i udarbejdelsen af hjemmeeksamensopgaver, og at grupper af studerende har presset på for, at universiteterne åbner bredere op for brug af generativ AI. Repræsentanterne peger også på, at et forbud, der ikke kan håndhæves, vil ende med at straffe de pligtopfyldende studerende, som overholder reglerne, og som dermed bliver stillet dårligere end de studerende, der anvender generativ AI på trods af et forbud.

I de følgende afsnit udfoldes det, hvordan universiteterne er gået til arbejdet med at udvikle retningslinjer for brug af generativ AI til prøver blandt de studerende.

Universiteterne har forskellige retningslinjer for brug af generativ AI

Universiteterne har inden for de seneste år vedtaget forskellige retningslinjer for de studendes brug af generativ AI til opgaver og eksaminer. Retningslinjerne kan deles op i to overordnede tilgange, som er hhv. *tilladt, medmindre andet er angivet*, og *ikke tilladt, medmindre andet er angivet*. På alle universiteterne er der dermed mulighed for, at de enkelte studieområder og kursusansvarlige kan tilpasse retningslinjerne til deres specifikke fagområde, undervisning og eksamensformer.

Det er fælles for de syv universiteters retningslinjer, at de studerende tydeligt skal angive, hvor og hvordan de har benyttet sig af generativ AI, hvis de har brugt teknologien til at generere hele eller dele af tekstpassager. Universiteterne sidestiller typisk deklaration af anvendelse af generativ AI med almindelig akademisk referencepraksis.

Tabel 4.1 viser de syv universiteters overordnede retningslinjer for studendes brug af generativ AI.

Tabel 4.1

Universiteternes overordnede retningslinjer for brug af generativ AI

Institution	Overordnede retningslinjer
DPU, AU	<ul style="list-style-type: none">• <u>Generativ AI tilladt</u> til alle opgaver og eksaminer, medmindre andet er angivet.• Krav om tydelig angivelse af anvendelse.
DTU	<ul style="list-style-type: none">• <u>Generativ AI tilladt</u> til opgaver og eksaminer med åbent internet, medmindre andet er angivet.• Krav om tydelig angivelse af anvendelse.
AAU	<ul style="list-style-type: none">• <u>Generativ AI tilladt</u> til opgaver og eksaminer, medmindre andet er angivet.• Krav om tydelig angivelse af anvendelse.
CBS	<ul style="list-style-type: none">• <u>Generativ AI tilladt</u> ved skriftlige hjemmeeksamensopgaver, bachelorprojekter og specialer.• <u>Specifikke retningslinjer</u> for andre prøveformer, herunder skriftlige stedprøver, varierer mellem studieområder og fag.• Krav om tydelig angivelse af anvendelse.
SDU	<ul style="list-style-type: none">• <u>Generativ AI ikke tilladt</u> til stedprøver eller mundtlige prøver, men <u>generativ AI tilladt</u> til hjemmeeksamensopgaver, bachelorprojekter og specialer, medmindre andet er angivet.• Krav om tydelig angivelse af anvendelse.
RUC	<ul style="list-style-type: none">• <u>Generativ AI ikke tilladt</u> til hjemmeeksamensopgaver, stedprøver eller mundtlige prøver, men <u>generativ AI tilladt</u> til projektrapporter, bachelorprojekter og specialer, medmindre andet er angivet.

Institution	Overordnede retningslinjer
	<ul style="list-style-type: none">• Krav om tydelig angivelse af anvendelse.
KU	<ul style="list-style-type: none">• <u>Generativ AI ikke tilladt</u>, medmindre de enkelte studieområder har tilladt brugen.• Fra efterårsemesteret 2025: <u>Generativ AI tilladt</u> til alle opgaver og eksaminer, medmindre andet er angivet.• Krav om tydelig angivelse af anvendelse.

Som det fremgår af tabellen ovenfor, tillader DPU og AAU i videst udstrækning de studerende at bruge generativ AI ifm. prøver. Her er det tilladt som hjælpemiddel til alle opgaver og eksaminer, medmindre andet er angivet. Det er op til den enkelte kursusansvarlige at beslutte, hvilke værktøjer der tillades, og hvordan værktøjerne må bruges ifm. opgaver og eksaminer.

På alle universiteter gælder det imidlertid, at der forekommer skærmede stedprøver eller stedprøver uden hjælpemidler.

På DTU er generativ AI tilladt til opgaver og eksaminer med åbent internet, medmindre andet er angivet. CBS tillader de studerende at anvende generativ AI til skriftlige hjemmeeksamensopgaver, bachelorprojekter og specialer, mens der findes specifikke retningslinjer for andre eksamensformer, som kan variere inden for det enkelte studieområde eller fag. SDU tillader som udgangspunkt brug af generativ AI til hjemmeeksamensopgaver, bachelorprojekter og specialer, men ikke ved stedprøver eller mundtlige prøver. RUC tillader brug af generativ AI ifm. projektrapporter, bachelorprojekter og specialer, men ikke til hjemmeeksamensopgaver eller andre eksaminer, medmindre andet er angivet. Både CBS og RUC har dog en målsætning om at udvikle nye eksamensformer, der i højere grad tillader brug af generativ AI.

KU har, som det eneste af de syv universiteter, fortsat forbud mod brug af generativ AI ifm. eksaminer – dog med mulighed for, at de enkelte studieområder kan tillade brug af teknologien til prøver lokalt. Fra efteråret 2025 ændres retningslinjerne på KU, så generativ AI som udgangspunkt må bruges til alle opgaver og eksaminer, medmindre andet er angivet. Denne ændring indebærer dog ikke, at de studerende i praksis får tilladelse til at bruge generativ AI til alle prøver på tværs af KU, men at de enkelte studieområder får frihed til selv at udvikle lokale retningslinjer for, hvornår teknologien skal holdes ude af prøvesituationer på det specifikke studie.

Løbende udvikling og tilpasning af retningslinjerne

Universiteterne har valgt forskellige retningslinjer for generativ AI, men det er fælles for universiteternes tilgang, at de løbende har udviklet og tilpasset retningslinjerne. Det hænger bl.a. sammen med, at teknologien er blevet udbredt på ganske kort tid, og at den fortsat er i hastig udvikling.

SDU er et eksempel på et universitet, som ændrede kurs med kort varsel efter fremkomsten af generativ AI. Her havde man oprindeligt besluttet at forbyde brug af generativ AI til prøver, men man endte med at tillade brug af teknologien ved skriftlige hjemmeeksamensopgaver fra september 2023. Ifølge den interviewede repræsentant for SDU anså man de nye retningslinjer som et eksperiment og samlede op på erfaringerne i løbet af semesteret. Det resulterede i et sæt reviderede retningslinjer det efterfølgende år, dog uden markante ændringer, da de foreløbige erfaringer ikke gav anledning til dette.

På DPU har man ligeledes løbende lært af processen med at tillade anvendelse af generativ AI til udvalgte eksaminer. I foråret 2024 åbnede DPU op for brug af generativ AI ifm. bachelorprojekter og specialer, under forudsætning af at de studerende deklarerede deres brug af teknologien. Ifølge den interviewede repræsentant for DPU erfarede de herefter, at der var behov for en skabelon, der kan skabe en fælles standard for de studerendes deklaration, når de har anvendt generativ AI i deres bachelorprojekter eller specialer. Den er nu udviklet, og man har en intention om på sigt at indarbejde deklarationen i opgavernes metodeafsnit.

KU har hidtil haft en mindre eksperimenterende overordnet tilgang til brug af generativ AI til prøver sammenlignet med de andre universiteter, og man ophæver som nævnt først det generelle forbud fra og med efterårssemesteret 2025. Ifølge den interviewede repræsentant for KU var det oprindelige formål med at formulere et forbud, at de ville undersøge potentialer og udfordringer, før de åbnede bredt op for brug af teknologien. Repræsentanten for KU forklarer desuden, at de enkelte studieområder har fået et år til at forberede overgangen til de nye retningslinjer.

Retningslinjerne kan ikke stå alene

Det fremgår af interviewene, at universiteterne har grebet udarbejdelsen og understøttelsen af retningslinjerne an på forskellig vis. På hovedparten af de syv universiteter har man nedsat interne arbejdsgrupper, som dels understøtter ledelsen i at udvikle og udbrede retningslinjer og andre støttematerialer og dels understøtter studieområderne i at inkorporere generativ AI i faglige mål og eksaminer. På RUC har biblioteket desuden spillet en stor rolle i at understøtte og systematisere hele universitetets tilgang til den nye teknologi. Biblioteket underviser fx alle studerende på første år af bacheloren og kandidaten i at bruge generativ AI ifm. litteratursøgning.

På nogle af universiteterne har de igangsat kompetenceudvikling for underviserne, som skal klæde dem på til at forberede de studerende til og bedømme dem ud fra deres brug af generativ AI til eksaminer. Kompetenceudviklingen kan også dreje sig om at begrænse de studerendes brug af generativ AI, fx ved at formulere opgaver og eksamensspørgsmål, som er svære at få direkte hjælp til ved brug af generativ AI. Nogle universiteter har også afholdt interne workshops, hvor medarbejderne har drøftet potentialer og barrierer ved brug og bedømmelse af generativ AI ifm. prøver.

Det fremgår desuden af interviewene med universitetsrepræsentanterne, at de generelt er optagede af løbende at følge med i, hvordan generativ AI bliver anvendt inden for de forskellige studieområder og ifm. forskellige prøveformer. På RUC er man fx ved at afdække anvendelsen af generativ AI på tværs af universitetet, så denne viden kan tænkes ind i universitetets retningslinjer.

4.1.2 Hvad er de første erfaringer med de opstillede retningslinjer?

Når repræsentanterne for de syv universiteter bliver bedt om at fortælle om erfaringerne med de opstillede retningslinjer, fylder udfordringerne mest. Repræsentanterne forklarer dog, at de oplevede udfordringer ikke er ensbetydende med, at de fortryder at have åbnet op for brug af generativ AI til udvalgte prøver på de enkelte universiteter. Det er snarere et udtryk for, at generativ AI er en ny teknologi for både undervisere og studerende, der fortsat er under hastig udvikling.

I de følgende afsnit udfoldes universiteternes indledende erfaringer med deres retningslinjer for brug af generativ AI til prøver. Det fremgår af afsnittene, at der er behov for tilpasning af retningslinjerne til de enkelte studieområder, men at behovet for at kontrollere for snyd ifm. eksamensopgaver generelt er ændret. Universiteternes erfaringer peger derudover på, at det kan være udfordrende at finde den rigtige form for de studerendes deklaration af, hvor og hvordan de har benyttet sig af teknologien, hvis det skal være muligt at inddrage som del af bedømmelsesgrundlaget.

Retningslinjer er nødvendige, men ikke ”one size fits all”

De interviewede repræsentanter for universiteterne oplever det generelt som positivt, at man har fået udviklet eller er i gang med at udvikle retningslinjer for brug af generativ AI til prøver. Det er således en generel opfattelse på tværs af de syv universiteter, at de finder det nødvendigt at forholde sig til teknologien, da den allerede bliver anvendt af de studerende og kun vil blive mere udbredt i de kommende år. Det er også et perspektiv blandt universitetsrepræsentanterne, at retningslinjerne kan fjerne noget af den tvivl omkring brug af teknologien, som har hersket blandt undervisere, eksaminatorer, censorer og studerende. Retningslinjerne kan dermed flytte fokus fra, *om* man bruger teknologien, til *hvordan* den bruges.

Repræsentanterne for universiteterne oplever, at måden, hvorpå undervisere og studerende arbejder med generativ AI, i høj grad skal tilpasses det enkelte fag og den enkelte eksamensform. Mens nogle fagområder og eksamensformer forholdsvis nemt kan forenes med generativ AI, kræver det større ændringer for andre fagområder og eksamensformer, ligesom enkelte fagområder og eksamensformer slet ikke vurderes som egnede til at integrere brug af generativ AI.

Repræsentanten for AAU nævner fx, at det kan fungere fint at integrere generativ AI som hjælpemiddel ved eksaminer, hvor der arbejdes med problembaseret læring, men at tekstproducerende fag, fremmedsprog og humanistiske discipliner generelt er udfordret af generativ AI. Repræsentanten for DTU fortæller også, at det fx ikke giver mening at integrere generativ AI ifm. matematikkurser i forudsætningsviden, hvor de studerende nemt kan anvende generativ AI til at finde svaret på en opgave, uden at det bidrager til deres grundlæggende forståelse af kursets kernestof.

Behovet for at kontrollere for snyd er ændret

Det er et udbredt perspektiv blandt de interviewede repræsentanter, at universiteterne i mindre grad skal bruge ressourcer på at kontrollere for snyd med brug af generativ AI, efter at de har tilladt de studerende at anvende teknologien til udvalgte prøver. Repræsentanterne forklarer, at de især har haft svært ved at håndhæve et forbud og bevise brug af generativ AI ifm. skriftlige hjemmeeksamensopgaver, og at der tidligere er blevet brugt meget energi og mange ressourcer på mistanker om snyd med generativ AI i bedømmelsen af eksamensopgaver, som sjældent kunne påvises. Repræsentanten for CBS fortæller fx, at de bliver kontaktet af færre bekymrede eksaminatorer og censorer, efter at generativ AI er blevet tilladt ved flere eksaminer.

Behovet for at kontrollere for snyd eksisterer dog fortsat – både i tilfælde hvor universiteternes retningslinjer begrænser brugen af generativ AI, og i tilfælde hvor brugen er tilladt, og de studerende skal angive, hvor og hvordan de har anvendt teknologien i opgaveløsningen. Repræsentanterne for universiteterne peger i den sammenhæng på, at det fortsat er svært at vurdere snyd med generativ

AI i skriftlige opgaver, der står alene, da det ofte kræver et godt kendskab til de studerendes faglige niveau og skrivestil at skille deres selvstændige præstation fra AI-genereret materiale.

Udgangspunkt i kendt referencepraksis skaber fordele, men også udfordringer

Det fremgår af interviewene, at universiteterne generelt har gode erfaringer med, at de studerende skal deklare deres anvendelse af generativ AI ifm. skriftlige eksamensprojekter med udgangspunkt i de allerede gældende retningslinjer for referencehåndtering. Disse retningslinjer er kendt af både undervisere og studerende og kan generelt overføres direkte til brug af generativ AI.

Det er dog også en pointe i interviewene, at det kan være svært at finde den rigtige form, når de studerende skal redegøre for deres anvendelse af generativ AI. Det skyldes, at en komplet redegørelse for, hvor og hvordan den studerende har brugt generativ AI, kan fylde mange sider, som bedømmerne ikke har tid til at gå i dybden med som bedømmelsesgrundlag. På flere af universiteterne forlanger de derfor kun, at de studerende deklarerer deres brug af teknologien, hvis de har brugt generativ AI til at producere tekst, der indgår ordret eller parafraseret i deres opgave.

4.2 Prøveformer på universiteterne i lyset af generativ AI

I interviewene er repræsentanterne for universiteterne blevet bedt om at forholde sig til, hvad fremkomsten af generativ AI betyder for tilrettelæggelsen af eksaminer på deres universitet. Det følgende afsnit udfolder i forlængelse heraf repræsentanternes beskrivelser af, hvad generativ AI betyder for eksamensformer og bedømmelse af de studerendes præstationer.

4.2.1 Justerede eksamensformer i lyset af generativ AI

Det er en udbredt pointe i interviewene, at fremkomsten af generativ AI har givet anledning til drøftelser af, hvordan man kan justere og sammensætte eksaminerne på universiteterne, så det er muligt at fastholde en valid bedømmelse af de studerendes selvstændige faglige niveau.

Universitetsrepræsentanternes eksempler på eksamensformer, universiteterne benytter sig af, der er tilpasset fremkomsten af generativ AI, kan opdeles i to grupper:

1. **Eksamensformer, der forhindrer brug af generativ AI.** Her afholdes eksamen som en stedbunden skriftlig eller mundtlig eksamen, hvor de studerende ikke har adgang til internettet og dermed heller ikke til generativ AI.
2. **Eksamensformer, der tillader brug af generativ AI.** Det indebærer både mundtlige og skriftlige eksaminer eller en kombination af forskellige eksamensformer. Det væsentligste er, at selve indholdet i eksamen tilpasses muligheden for at bruge generativ AI. Ifm. skriftlige eksaminer betyder det, at de studerende skal redegøre for og evt. reflektere over deres brug af generativ AI, hvilket indgår som en del af bedømmelsesgrundlaget. Der nævnes ligeledes eksempler på skriftlige eksaminer, der tillader brug af generativ AI, hvor bedømmerne stiller eksamensspørgsmål, som ikke udelukkende kan besvares gennem brug af generativ AI, fx ved at inddrage cases og specifikke eksempler fra undervisningen. Derudover består eksamensformerne, der tillader brug af generativ AI, også i kombinerede skriftlige og mundtlige eksaminer, fx ifm.

bachelorprojekter og specialer eller andre eksaminer, hvor de studerende skal forsvare deres skriftlige projekt mundtligt og får en samlet bedømmelse på den baggrund.

De interviewede repræsentanter peger på, at universiteterne generelt har opmærksomhed omkring at sammensætte den samlede mængde af prøver i løbet af de studerendes studietid, så de både udprøves gennem eksaminer med og uden brug af hjælpemidler, herunder generativ AI. Det indebærer også at kombinere skriftlige og mundtlige eksaminer, så det er muligt for eksaminator og censor at få blik for og vurdere den studerende ud fra, om personen har lært og kan forholde sig selvstændigt til kursets centrale faglige mål. Universitetsrepræsentanterne peger også på løbende vejledning ifm. større skriftlige eksamensopgaver som en måde, hvorpå vejlederen kan få indblik i og en fornemmelse for den studerendes faglige niveau, der kan danne baggrund for den endelige bedømmelse af opgaven.

Universitetsrepræsentanterne forklarer, at den ovenstående variation i sammensætningen af de eksamensformer, studerende møder i løbet af deres studietid, kan være med til at forebygge, at studerende kan klare sig gennem deres studie uden selvstændigt at have arbejdet med og tilegnet sig studiets centrale faglige mål og kompetencer.

4.2.2 Metakompetencer bliver vigtigere i lyset af generativ AI

Det er en udbredt pointe i interviewene, at fremkomsten af generativ AI har nødvendiggjort et større fokus på metakompetencer som bedømmelseskriterium på universiteterne. Repræsentanterne for universiteterne har dog forskellige holdninger til, om dette fokus skal være det primære bedømmelseskriterium, eller om det blot skal være en del af et bredere bedømmelsesgrundlag.

Den generelle opfattelse på tværs af repræsentanterne er, at kritisk tænkning, kildekritik og referencehåndtering er helt centrale metakompetencer, som de studerende skal besidde og udvikle gennem deres uddannelse. Det gælder, uanset om de må bruge generativ AI eller ej. Det er derfor en udbredt pointe, at fremkomsten af generativ AI ikke skaber behov for nye bedømmelseskriterier til prøverne, men at det kan være relevant at placere større fokus på de førnævnte kompetencer.

Nogle af de interviewede repræsentanter for universiteterne mener dog, at fremkomsten af generativ AI bør medføre større krav til de studerende. De pågældende repræsentanter forklarer det med, at generativ AI kan hjælpe de studerende med at besvare de mest basale spørgsmål og med at skabe struktur i deres eksamensbesvarelser. De mener derfor, at universiteterne bør stille større krav til den individuelle præstation og lade eksaminationen og bedømmelsen lægge endnu større vægt på metakompetencer som refleksionsevne, herunder om de studerende formår at forholde sig reflekteret til deres arbejdsproces frem mod det endelige produkt.

5 Bedømmelsespraksis og prøver i gymnasiet i lyset af AI

I dette kapitel fokuserer vi på, hvilke greb lærere benytter sig af for at tage højde for fremkomsten af generativ AI i deres bedømmelse af elevernes selvstændige faglige niveau. Pointerne i kapitlet bygger på kvalitative interview med ledere og lærere fra 12 udvalgte gymnasier på tværs af htx, hhx, stx og hf, som alle har igangsat konkrete initiativer med fokus på brug og bedømmelse af generativ AI.

Det fremgår af kapitlet, at lærerne benytter sig af en række greb som reaktion på udbredelsen af generativ AI for at kunne bevare deres føling med og mulighed for at bedømme elevernes læring og selvstændige faglige niveau. Grebene består i et øget fokus på arbejdsformer, der kan træne elevernes grundfaglighed, en fornyet opmærksomhed på at forskyde fokus fra resultat til proces og en endnu større betoning af mundtlighed som opfølgning på skriftligt arbejde.

Grebene fremlægges som en reaktion på udfordringer med validiteten i nogle af de eksisterende bedømmelsespraksisser, som knytter sig til elevernes skriftlige fremstilling i lyset af generativ AI. Det er således fælles for grebene, at de enten skærmer eleverne fra at bruge hjælpemidler i deres skriftlige arbejde eller giver lærerne mulighed for at følge op på og bedømme, hvorvidt og hvordan eleverne arbejder selvstændigt i deres arbejdsproces med skriftlige produkter, når de må benytte sig af forskellige hjælpemidler, herunder generativ AI.

Det fremgår afslutningsvis af kapitlet, at der kan opsummeres to potentielle veje for skriftlige prøver i gymnasiet pba. de interviewede ledere og læreres erfaringer med bedømmelse i lyset af generativ AI. Det drejer sig for det første om skriftlige prøveformer med mundtligt forsvar og for det andet om skriftlige prøveformer, der inkluderer elementer uden hjælpemidler. Det flugter med erfaringerne fra litteraturopsamlingen og universiteterne samt ekspertgruppen for ChatGPT og andre digitale hjælpemidlers anbefalinger om prøver, der skal gå på to ben og udprøve både med og uden digitale hjælpemidler.

5.1 Konkrete initiativer

På tværs af de 12 gymnasier, der indgår i kortlægningen, er der stor forskel på, hvor formaliserede de initiativer, de hver især arbejder med, er. Initiativerne spænder over velbeskrevne, regionsfinansierede indsatser og samarbejder med forskningsmiljøer til mere lokale aktiviteter, hvor generativ AI udgør et strategisk fokuspunkt snarere end en detaljeret og afgrænset indsats. Der er derudover

gymnasier repræsenteret i kortlægningen, som har oprettet særlige studieretninger, der centrerer sig omkring brug af generativ AI (se tabel B.1 i appendiks B for en samlet oversigt over de 12 gymnasiers initiativer).

5.1.1 Fællestræk ved de 12 initiativer

Selvom de 12 gymnasiers initiativer er forskellige, går der elementer igen på tværs af initiativerne, som fokuserer på bedømmelse og prøver i lyset af generativ AI.

Det knytter sig især til lærernes kompetencer og elevernes oplæring i teknologien. Selvom disse erfaringer tager afsæt i undervisningen, har de bidraget med viden om, hvilke kompetencer lærerne skal have, hvis generativ AI bliver tilladt til prøver i deres fag, dels for at kunne klæde eleverne på til denne prøvesituation og dels for at kunne tage højde for elevernes brug af teknologien som del af bedømmelsesgrundlaget til prøven.

Kompetenceudvikling og aktionslæring som greb til at klæde lærerne på

Flere af gymnasiernes initiativer er bygget op omkring, hvordan lærerne kan tilrettelægge undervisning og stille opgaver, der inddrager teknologien på en måde, hvor lærerne fortsat kan bevare blikket for og vurdere elevernes læring.

Derudover centrerer initiativerne sig om en række aktionslæringsforløb, typisk forankret i faggrupper, hvor lærerne udvikler og gennemfører undervisningsforløb og opgaver, der inddrager generativ AI, mhp. at gøre sig erfaringer med, hvornår brug af teknologien fungerer godt og mindre godt inden for deres fagområde.

Elever klædes på til læringsunderstøttende brug af generativ AI

På tværs af de 12 gymnasiers initiativer indgår der ligeledes elementer, som har til formål at klæde eleverne på til at anvende generativ AI på en læringsunderstøttende måde mhp. at forebygge, at de anvender teknologien som et redskab til snyd. Det indebærer også en afdækning af, hvordan eleverne anvender teknologien og deres holdninger til og forståelse af plagiat, fx gennem snakke med elever om, hvornår generativ AI er læringsfremmende og -hæmmende, og hvornår man udliciterer læring til generativ AI eller indgår i en læringsproces i samarbejde med teknologien.

5.1.2 Eksempler på faglige mål og bedømmelseskriterier ved brug af generativ AI

Som del af initiativerne på de 12 gymnasier er der desuden udviklet en række forskellige undervisningsforløb, hvor lærerne har arbejdet med at inddrage og opstille læringsmål for elevernes brug af generativ AI på forskellig vis og i forskellige fag (se tabel B.2 i appendiks B for seks eksempler på dette).

Eksemplerne giver et indblik i, hvad det er for nogle formål, lærerne opstiller for elevernes læring, når de må benytte sig af generativ AI i en opgaveproces, og dermed også hvad det er for nogle kriterier, som lærerne bedømmer elevernes indsats i forløbene ud fra. Det fremgår på tværs af de seks eksempler, at det drejer sig om tre overordnede fokuspunkter ift. læring og bedømmelse:

1. **Elevernes grundfaglighed:** Eksempelvis elevernes genreforståelse, begrebslæring og evne til at formulere sig sprogligt korrekt både skriftligt og mundtligt på fremmedsprog.
2. **Elevernes metakompetencer:** Eksempelvis elevernes evne til at kunne forholde sig kritisk til det output, de får gennem deres brug af generativ AI som del af opgaveprocessen, og have en reflekterende tilgang til egen læreproces, ift. hvornår det giver mening at benytte sig af teknologien, og hvornår det ikke gør.
3. **Elevernes interaktion med generativ AI:** Eksempelvis elevernes evne til at benytte sig af læringsunderstøttende prompts i det specifikke fag.

I afsnit 5.3 udfoldes det, hvorfor lærerne på tværs af de 12 gymnasier vurderer, at ovenstående fokuspunkter er relevante bedømmelseskriterier i lyset af generativ AI's udbredelse i gymnasiet.

5.2 Udfordringer med bedømmelse

Det fremgår af interviewene med ledere og lærere fra de 12 gymnasier, at gymnasiernes initiativer er udsprunget af et akut behov for viden og handling for at kunne navigere i en ny virkelighed. Det knytter sig ikke mindst til de implikationer, som fremkomsten af generativ AI har for elevernes tilgang til skriftlige produkter og lærernes mulighed for at foretage en retvisende bedømmelse af disse.

Ifølge lederne og lærerne udgør generativ AI på den måde et paradigmeskifte, der rækker ved den måde, gymnasierne tænker og bedriver uddannelse på. Det gælder både gymnasiets formål og undervisningens form og indhold, og dermed også hvordan gymnasierne i sidste ende udprøver og bedømmer, om eleverne har lært det, de skal.

De følgende afsnit beskriver, hvordan de interviewede ledere og lærere på tværs af htx, hhx, stx og hf oplever generativ AI som et uomgængeligt vilkår, der sætter elevernes læring og lærernes mulighed for at have føling med og bedømme denne under forandring. Det knytter sig til, at:

1. Elevernes brug af generativ AI pendulerer fra læringsunderstøttelse til snyd
2. Tvivl om elevernes brug af generativ AI skaber usikkerhed i bedømmelsessituationer.

5.2.1 Elevernes brug af generativ AI pendulerer fra understøttelse til snyd

Det fremgår af interviewene, at elevernes brug af generativ AI fylder meget på gymnasierne, og at eleverne anvender teknologien til en lang række forskellige aktiviteter ifm. deres uddannelse.

I interviewene kommer lærerne med forskellige eksempler på, hvad eleverne anvender generativ AI til. Det kan bl.a. være til at få informationer om et nyt emne eller begreb eller til at få forklaret vanskeligt stof i en tekst. Elevernes brug af generativ AI sammenlignes i den forbindelse med, hvordan eleverne tidligere har brugt Google eller Wikipedia. Lærerne giver også eksempler på, at eleverne

anvender generativ AI som sparringspartner til at skærpe deres projektidé eller få inspiration undervejs i deres arbejdsproces ifm. skriftlige opgaver eller større projekter. Lærerne forklarer desuden, at nogle elever anvender generativ AI til at komme i gang med en opgave gennem en disposition eller et første udkast, som de kan arbejde videre med, eller til at rette fejl, når de har udarbejdet et udkast eller en opgavebesvarelse. Endelig er der også eksempler på, at elever helt eller delvist udliciterer udarbejdelsen af skriftlige opgaver til generativ AI.

Elevernes brug af generativ AI spænder altså vidt og placerer sig på et kontinuum fra at være et hjælpemiddel, der kvalificerer deres opgaveløsning og understøtter deres læreproces, til et redskab, der medfører tab af selvstændighed, udlicitering af læring og direkte snyd.

Lærerne oplever, at der kan være gråzoner på det beskrevne spektrum, hvor det kan være svært at gennemskue, hvilken rolle generativ AI har spillet i elevernes opgaveløsning. Her bliver lærernes kendskab til generativ AI på den ene side og elevernes selvstændige faglige niveau på den anden side afgørende for, om lærerne kan få øje på generiske formuleringer eller uoverensstemmelser mellem en elevs normale faglige niveau og niveauet i en opgave ift. disposition, brug af begreber eller formidling.

Som det også fremgår af litteraturopsamlingens afsnit 3.1.1, kan det dog være en vanskelig opgave at skelne mellem elev- og AI-genereret tekst. Lærerne forklarer i overensstemmelse med denne pointe, at de bruger meget tid og energi på at vurdere, om elevernes præstationer er et udtryk for deres eget selvstændige faglige niveau eller for snyd gennem brug af generativ AI, når de bedømmer skriftlige opgaver.

5.2.2 Tvivl om elevers brug af generativ AI skaber usikkerhed i bedømmelsessituationer

Det fremgår af litteraturopsamlingens afsnit 3.1.2, at underviseres grundlæggende tvivl om, hvorvidt elevernes arbejde er et udtryk for en selvstændig indsats eller snyd, indfører en mistillid i lærer-elev-relationen, der i sidste ende kan føre til en hårdere bedømmelse af elevernes produkter.

Denne pointe går igen i interviewene med lærerne fra de 12 gymnasier, som giver udtryk for, at tilstedeværelsen af generativ AI medfører en generel mistillid til eleverne og en konstant bevågenhed ift. at bedømme, hvorvidt og hvordan generativ AI har spillet en rolle for elevernes præstationer.

Lærerne er bevidste omkring, at det i nogle tilfælde kan medføre, at elever, der har arbejdet selvstændigt med en opgave, bliver mistænkt for at have snydt gennem brug af generativ AI. Lærerne oplever ligeledes, at det kan føre til, at de fagligt dygtigste elever i nogle tilfælde indfører bevidste fejl i deres opgaver, fordi de er bange for at blive uretmæssigt mistænkt for at have snydt.

Det fremgår af interviewene, at det netop er blandt hovedårsagerne til, at de 12 gymnasier har igangsat initiativer med generativ AI, at de ønsker at komme denne grundlæggende mistillid til livs, dels ved at klæde lærere og elever på til at bruge teknologien som læringsunderstøttende hjælpemiddel og dels ved at klæde lærerne på til at udvikle nye undervisnings- og bedømmelsespraksisser, der gør det muligt for dem at bevare blikket for og bedømme elevernes læring og selvstændige faglige niveau efter fremkomsten af generativ AI.

5.3 Bedømmelseskriterier

Det fremgår af interviewene med lærerne, at der generelt er en række greb, de har skruet op for som reaktion på udbredelsen af generativ AI for at kunne bevare deres føling med og mulighed for at bedømme elevernes læring og selvstændige faglige niveau.

Grebene er især knyttet til et øget fokus på arbejdsformer, der kan træne elevernes grundfaglighed; en styrket opmærksomhed på elevernes arbejdsproces, herunder hvordan eleverne bruger generativ AI og reflekterer over dette undervejs i arbejdet med en opgave, og en endnu større betoning af mundtlighed.

På baggrund af lærernes pointer fremstår de fire fokuspunkter i boksen nedenfor som centrale bedømmelseskriterier i lyset af generativ AI, der understøtter, at lærerne kan udvikle og bedømme elevernes selvstændige faglige niveau. Det er en pointe blandt lærerne, at de tre første fokuspunkter er velkendte greb, som de i forvejen har benyttet sig af for at få blik for og bedømme elevernes selvstændige faglige niveau, men som lærerne har fundet det nødvendigt at have større fokus på som reaktion på fremkomsten af generativ AI. Boksens pointer udfoldes i de følgende afsnit.

Bedømmelseskriterier i lyset af generativ AI

1. Fokus på elevernes grundfaglighed

- *Bedømmelsesgrundlag:* Skriveøvelser med pen og papir eller test uden hjælpemidler samt opgavebeskrivelser, der bevidst er udformet, så opgaven er vanskelig at besvare gennem brug af generativ AI.
- *Bedømmelsesfokus:* Basale faglige kompetencer.

2. Fokus på elevernes arbejdsproces

- *Bedømmelsesgrundlag:* Projektarbejde og åbne opgaver.
- *Bedømmelsesfokus:* Elevernes ageren i en opgaveproces med fokus på mere avancerede faglige kompetencer, herunder metakompetencer såsom kritisk tænkning, nysgerrighed og refleksion over egen læring.

3. Fokus på mundtlig formidling

- *Bedømmelsesgrundlag:* Mundtlige præsentationer og gruppedrøftelser eller videoer og podcasts som supplement til skriftlige produkter.
- *Bedømmelsesfokus:* Basale faglige kompetencer ift. mundtlig formidling, og om elevernes mundtlige præsentation indikerer, at de har arbejdet selvstændigt med deres skriftlige produkt.

4. Fokus på elevernes interaktion med generativ AI

- *Bedømmelsesgrundlag:* Bilag ifm. skriftlige opgaver, hvor eleverne vedlægger deres prompts og evt. deres refleksion over, hvordan de har brugt dem til at forbedre deres opgave.
- *Bedømmelsesfokus:* Elevernes brug af læringsunderstøttende prompts, herunder om elevernes prompts indikerer, at de har benyttet sig af generativ AI som en konstruktiv sparringspartner i udarbejdelsen af deres opgave eller som et redskab til at udlicitere udarbejdelsen af dele af eller hele opgaven.

5.3.1 Generativ AI nødvendiggør fokus på grundfaglighed

Det er en pointe blandt både ledere og lærere, at lærerne har set sig nødsaget til at skrue op for arbejdsformer, der træner elevernes grundfaglighed i de enkelte fag, som respons på udbredelsen af generativ AI. Det handler for det første om, at der ifølge lærerne er nogle grundlæggende fagspecifikke kompetencer, eleverne skal kunne uafhængig af diverse hjælpemidler, og som samtidig er nødvendige at tilegne sig for at kunne indgå i en meningsfuld faglig dialog med generativ AI. Lærerne forklarer derudover, at arbejdet med grundfaglighed kan give dem en føling med og et grundlag for at bedømme elevernes selvstændige tilegnelse af fagets centrale faglige mål uafhængigt af generativ AI og andre digitale hjælpemidler.

I interviewene viser der sig især tre tilgange til at udvikle og bedømme elevernes grundfaglighed. De tre tilgange supplerer eller opstiller benspænd for elevernes arbejde med skriftlige hjemmeopgaver, som lærerne ikke nødvendigvis oplever giver et repræsentativt indblik i elevernes læring og selvstændige faglige niveau.

De tre tilgange består i:

1. Skriveøvelser med blyant og papir
2. Løbende test uden hjælpemidler
3. Opgavebeskrivelser, der aktivt tager højde for generativ AI's svagheder i faget.

De tre tilgange uddybes i de følgende afsnit. Som ramme for forståelsen af de to første tilgange er det vigtigt at understrege, at det hverken er en generel pointe blandt ledere eller lærere, at den analoge undervisning, som de to tilgange repræsenterer, skal stå alene. De to tilgange skal ses som elementer, der kan understøtte elevernes mulighed for faglig fordybelse og lærernes mulighed for at vurdere elevernes læring i faget.

Træning af grundfaglighed gennem skriveøvelser med blyant og papir

Det er en pointe blandt lærerne, at de har set sig nødsaget til at skrue op for skriveøvelser med blyant og papir som reaktion på udbredelsen af generativ AI. Ifølge lærerne er det en arbejdsform, som kan fastholde eleverne i at træne kompetencer i fagene, hvor de ellers ville kunne benytte sig af generativ AI til at springe over, hvor gærdet er lavest. Samtidig forklarer lærerne, at det analoge

arbejde giver dem en mulighed for at bedømme elevernes faglige niveau uafhængigt af digitale hjælpemidler.

I dansk og fremmedsprog kan det fx dreje sig om grammatikøvelser eller tidsafgrænset friskrivning om umiddelbare refleksioner pba. læsningen af en tekst. Sproglærerne finder det især vigtigt at prioritere de analoge grammatikøvelser i undervisningen for at understøtte elevernes læring af basale sproglige kompetencer, som de nemt ville kunne få hjælp til at finpudse ved brug af generativ AI.

En dansk- og engelsklærer fra htx sætter eksempelvis følgende ord på sin prioritering af de analoge grammatikøvelser:

”Jeg er begyndt at finde grammatikopgaverne frem igen, som eleverne skal lave med papir og blyant uden hjælpemidler. Det gør det muligt for mig at få et indblik i deres reelle faglige niveau, som jeg så kan sammenligne med deres hjemmeafleveringer, hvor de har måttet bruge AI. På den måde bliver en del af undervisningen lidt mere old school. Men eleverne kan egentlig godt lide at skrive på papir. Måske de godt kan lide det taktile som et lille afbræk.

LÆRER, DANSK OG ENGELSK, HTX

Det fremgår desuden af interviewene, at de analoge skriveøvelser i samfundsfag, biologi og bioteknologi samt profilfagene på erhvervs gymnasierne kan bestå i udkast til begrebsafklaringer eller redegørelse for teorier, som man så sidenhen kan arbejde videre med og præcisere ved brug af forskellige hjælpemidler.

I matematik, fysik og kemi samt de mere matematiske profilfag på erhvervs gymnasierne kan det analoge element også bestå i udregning af opgaver uden digitale hjælpemidler. Det er her en pointe blandt lærerne, der underviser i matematik A på stx, at det giver god mening at prioritere den analoge opgaveløsning i undervisningen, i lyset af at prøveformerne i faget netop er blevet ændret til at bestå af en tre timers prøve uden digitale hjælpemidler og en to timers prøve med digitale hjælpemidler.

Træning af grundfaglighed gennem test uden hjælpemidler

I forlængelse af det øgede fokus på analoge skriveøvelser er der også en gruppe lærere, der fortæller, at de er begyndt at indføre test uden hjælpemidler. De pågældende lærere forklarer, at det dels handler om løbende at kunne bedømme elevernes selvstændige tilegnelse af grundlæggende faglige mål og dels om at skabe et incitament blandt eleverne til at fordybe sig og lægge sig i selen for at lære centrale kompetencer i faget, selvom de i princippet ville kunne udlicitere denne læring gennem brug af generativ AI.

Hvor lærere i fag med matematiske elementer fortæller, at det er et greb, de også har benyttet sig af før udbredelsen af generativ AI, er det en udbredt pointe blandt lærere i sprogfag, at de har indført testene som modsvar til den sproglige hjælp, deres elever kan få fra generativ AI. De pågældende sproglærere forklarer, at de generelt ville foretrække ikke at skulle bruge tid på denne type

test, men at det er nødvendigt, hvis de skal have et sikkert grundlag for at vurdere elevernes sproglige og skriftlige kompetencer uafhængigt af generativ AI.

En lærer, der underviser i dansk og engelsk på hhx, sætter som eksempel følgende ord på sine overvejelser bag at lægge test uden hjælpemidler ind i sin engelskundervisning:

”Jeg tror, at der er rigtig mange sproglærere, som efterhånden arbejder med forskellige former for test uden hjælpemidler i deres undervisning. Jeg giver heller ikke karakterer for opgaver, de laver derhjemme, hvor jeg ikke har noget indblik i, hvad de har brugt af hjælpemidler. Jeg supplerer det med test i timen, for der har jeg mulighed for at være over dem. [...] Jeg har aldrig før lavet test på papir. Man tænker, at det er fuldstændig oldnordisk, men man kan simpelthen være så meget i grundtvivl om, hvad eleverne rent faktisk kan med alle de lækre hjælpemidler, der findes.

LÆRER, DANSK OG ENGELSK, HHX

Det er ligeledes en pointe blandt sproglærerne, at det også handler om professionel integritet ift. at føle, at man som lærer har gjort sit for at klæde sine elever på til de vilkår, de bliver bedømt ud fra til en eventuel eksamen.

Træning af grundfaglighed gennem opgavebeskrivelser, der udfordrer generativ AI

Endelig fremgår det af lærerinterviewene, at de også har fokus på at udvikle og bedømme elevernes grundfaglighed ved at stille opgaver, som det kan være svært at få kvalificeret hjælp til at løse gennem brug af generativ AI. Det kan fx være gennem fokus på autenticitet i opgavebeskrivelserne eller gennem genrespecifik skrivning – især på andre sprog end engelsk, som generativ AI ikke er lige så godt trænet i endnu.

Det fremgår af interviewene med samfundsfagslærere og lærere fra profilmag på hhx, at man kan forebygge uhensigtsmæssig brug af generativ AI i opgaveløsningen ved at have fokus på autenticitet i opgavebeskrivelsen, såsom cases fra gymnasiets lokalsamfund. Ved eksempelvis at bede eleverne om at tage udgangspunkt i en specifik case fra en virksomhed, der ligger i gymnasiets nærområde, kan man ifølge lærerne gøre det sværere for eleverne at udlicitere opgaven ved brug af generativ AI. Det skyldes, at teknologien ikke nødvendigvis er trænet ud fra data, der tager højde for den specifikke kontekst, som den lokale virksomhed befinder sig i.

De interviewede dansklærere peger ligeledes på, at fremkomsten af generativ AI har fået dem til at skruer op for deres fokus på genrebevidsthed som et greb til at bevare blikket for og bedømme elevernes selvstændige faglige niveau. Det er således en pointe blandt dansklærere, at generativ AI på nuværende tidspunkt ikke er trænet tilstrækkeligt i det danske sprog og genretraditioner til fx at kunne skrive et fyldestgørende essay. På den måde tvinger genrespecifikke opgaver eleverne til selv at tage stilling til genreens stilistiske træk og enten selv skrive opgaven ud fra denne ramme eller have tilstrækkeligt kendskab til genre til at kunne omskrive eventuel hjælp fra generativ AI ud fra genreens karakteristika.

En dansklærer fra stx beskriver denne pointe på følgende måde:

”**Genreforståelse er et vigtigt bedømmelseskriterium i dansk, hvor eleverne ikke nødvendigvis kan få kvalificeret hjælp ved brug af AI. De forskellige chatbotter er fx ikke trænet særlig godt i at skrive essays på dansk. Så på den måde slipper eleverne ikke for at skulle tage stilling til, om det output, botten kommer med, også lever op til genrekriterierne for et essay. Eleverne skal derfor vide, hvad et essay i dansk består i, og hvordan man bygger det op, før de kan bruge den hjælp, de evt. kan få ved at få AI til at skrive det.**

LÆRER, DANSK, STX

Det er samtidig en pointe blandt dansklærerne, at det også gør det nemt at få øje på afleveringer, hvor eleverne ukritisk har benyttet sig af generativ AI uden at have forståelse for, at den tekst, der er kommet ud af det, ikke lever op til de genrekrav, der var i fokus i opgavebeskrivelsen.

5.3.2 Lærere fokuserer på elevers arbejdsproces i lyset af generativ AI

Ledere og lærere peger på, at udbredelsen af generativ AI har været katalysator for et perspektivskifte fra resultat til arbejdsproces. Lærerne forklarer, at det forstærkede fokus på elevernes ageren i en opgaveproces først og fremmest handler om, at det giver et indblik i og mulighed for at bedømme elevernes selvstændige faglige niveau. Det skyldes ifølge lærerne, at fokus hermed flyttes fra den endelige præstation, som eleverne i princippet kan udlicitere gennem brug af generativ AI, til *hvordan* eleverne har arbejdet med opgaven.

Det fremgår af interviewene, at lærerne især sætter fokus på elevernes arbejdsproces som didaktisk greb og bedømmelseskriterium ved at:

1. Have fokus på elevernes arbejdsproces
2. Bedømme eleverne ud fra deres interaktion med generativ AI, når brug af teknologien har været tilladt i opgaveprocessen.

Lærerne sætter fokus på arbejdsproces

Det fremgår af interviewene, at lærerne i endnu højere grad bruger undervisningstid på elevernes arbejdsproces ifm. opgaveløsninger efter fremkomsten af generativ AI og også bruger elevernes ageren i denne proces som et bedømmelseskriterium i faget.

Det er således en pointe blandt lærerne, at en konstruktiv tilgang til en opgaveproces forudsætter, at eleven har centrale kompetencer, der afspejler elevens faglige niveau, såsom tekstlæsning, udarbejdelse af en disposition, kritisk tænkning, refleksion over egen læreproces, evnen til at opstille relevante og nysgerrige spørgsmål til opgaven og forholde sig kritisk og reflekterende til kilder og outputtet af de hjælpemidler, eleven benytter sig af. Lærerne forklarer derfor, at de er begyndt at bruge mere tid på at dvæle ved elevernes arbejdsproces, så de i højere grad kan få indblik i og bedømme elevernes indsats ud fra, hvorvidt og hvordan de agerer selvstændigt i den forbindelse.

En lærer i afsætning fra hhx sætter eksempelvis følgende ord på, hvordan generativ AI har fået personen til primært at fokusere på elevernes ageren i opgaveprocesser frem for det endelige resultat i sin bedømmelse af skriftlige produkter:

”Jeg har stort set kun fokus på processen: Elevernes forståelsesevne, hvordan de reflekterer over grundfagligheden, og hvordan man kan kombinere forskellige teorier ved inddragelse af en sprogmodel. Det er processen, jeg bedømmer dem på, og ikke det endelige resultat. Allerede i den første lektion siger jeg med et smil om munden, at det endelige resultat jo kan være perfekt, selvom de ikke har forstået en skid. Det kan være formuleret perfekt, selvom de ikke kan forklare det. Det hele handler om proces, og det gør jeg meget ud af at forklare dem. Det kan de godt have lidt svært ved at forstå. De vil meget gerne have en karakter for det produkt, de afleverer.

LÆRER, AFSÆTNING, HHX

Fællesnævneren for disse tilgange er ifølge lærerne, at de træner eleverne i at bruge generativ AI som et konstruktivt hjælpemiddel i en opgaveproces, hvor brugen af teknologien ikke går ud over elevens selvstændige faglighed i opgaveløsningen og lærerens mulighed for at bedømme denne.

Lærerne sætter fokus på proces gennem elevers interaktion med generativ AI

Som det fremgår af afsnit 4.1.1, er det en udbredt praksis på universiteterne at tillade brug af generativ AI ifm. skriftlige projekter med kravet om, at de studerende skal vedlægge deres prompts som bilag, evt. suppleret af refleksioner over, hvordan de har brugt outputtet i deres opgave. Denne praksis går igen i interviewene med lærere fra de 12 gymnasier, hvor flere forklarer, at de ligeledes har tilladt brug af generativ AI ifm. opgaver i den løbende undervisning med krav om, at eleverne skal dokumentere deres brug af teknologien som bilag.

Lærerne forklarer, at de på den måde kan få indblik i og bedømme opgaver ud fra elevernes interaktion med generativ AI. Det inkluderer et bedømmelsesfokus på, hvorvidt og hvordan eleverne har forholdt sig kritisk til outputtet af deres brug af generativ AI i relation til den specifikke opgave, herunder hvordan de er kommet frem til det konkrete output og har brugt det konstruktivt i deres opgave. Det er en pointe blandt lærerne, at det er et tidskrævende tiltag, fordi det forudsætter, at lærerne både bruger tid på at læse elevernes skriftlige produkt og de dertilhørende bilag.

En lærer i historie og samfundsfag fra stx sætter dog følgende ord på sin oplevelse af, at kravet om, at eleverne skal vedlægge deres prompts som bilag, i sig selv kan få dem til at bruge generativ AI mere velovervejet ud fra visheden om, at deres lærer potentielt kommer til at læse dem igennem:

”Jeg kan selvfølgelig ikke tjekke, om de vedlægger alle deres prompts som bilag, og jeg kan heller ikke læse dem alle sammen igennem. [...] Men det virker lidt som et panoptikon, at eleverne ved, at jeg kan finde på at læse deres prompts igennem. Eleverne antager, at jeg bruger bilagene til at kontrollere, om de har snydt, og det alene gør, at de ikke bare kopierer tekstbidder direkte fra sprogmodellen.

LÆRER, HISTORIE OG SAMFUNDSFAG, STX

Det er således en pointe blandt lærerne, at klare retningslinjer for, hvordan eleverne skal redegøre for og henvise til deres prompts ifm. en skriftlig opgave, dels kan forebygge, at de bruger generativ AI til at snyde, og dels kan give lærerne indblik i elevernes opgaveproces og hermed deres selvstændige aftryk på den endelige opgave.

5.3.3 Generativ AI får lærere til at skrue op for mundtlig formidling

Som beskrevet i afsnit 5.1.1 oplever ledere og lærere, at bedømmelsen af skriftlige produkter er særligt udfordret med udbredelsen af generativ AI. Fremkomsten af generativ AI har således accelereret en mistillid til, hvorvidt skriftlige hjemmeopgaver egentlig giver et repræsentativt indblik i elevernes selvstændige faglige niveau, der allerede begyndte at manifestere sig med introduktionen af digitale hjælpemidler som Wikipedia og Google Translate. Lærerne forklarer i forlængelse heraf, at de har fundet det nødvendigt at skrue op for mundtlig formidling for på den måde at kunne bedømme, om eleverne arbejder selvstændigt med skriftlige opgaver og har den ønskede faglige progression i faget.

Det fremgår af interviewene, at det øgede fokus på mundtlighed består i:

1. Mundtlig opfølgning på skriftlige opgaver
2. Fordybelsestid bruges på opgaver med mundtlig og skriftlig formidling.

Disse pointer udfoldes i de følgende afsnit.

Mundtlig opfølgning på skriftlige opgaver

Det er især en udbredt pointe blandt lærere i sprog- og samfundsvidenskabelige fag, at de ser sig nødsaget til at inkludere noget af det skriftlige arbejde i undervisningen for at forebygge uhensigtsmæssig brug af generativ AI og få indblik i elevernes skriveproces. Nogle lærere beskriver ligeledes, at de reelt kan være i tvivl om, hvorvidt det er elevens selvstændige skriftlige arbejde, de giver feedback på.

Det er derfor en pointe blandt lærerne, at de i højere grad er begyndt at prioritere undervisningstid til at følge mundtligt op på elevernes skriftlige arbejde. Det kan fx bestå i mundtlige præsentationer i undervisningen pba. skriftlige opgaver, hvor eleverne efterfølgende får feedback og en eventuel karakter ud fra en helhedsvurdering af deres mundtlige og skriftlige præstation. Det kan også bestå i mundtlige test pba. en opgave, eleverne har arbejdet med som lektie, eller dreje sig om peer-to-

peer-feedback i undervisningen, hvor eleverne diskuterer og giver feedback på hinandens opgaver i grupper, hvorefter de får mulighed for at gennemskrive deres opgave i undervisningen.

En uddannelsesleder fra hhx sætter fx følgende ord på behovet for at supplere skriftligt arbejde med en mundtlig dimension i lyset af generativ AI:

”Et greb, mange af vores lærere bruger, er at kombinere skriftlighed med mundtlighed. [Med udbredelsen af AI] skal skriftlighed altid følges op af noget mundtligt, så man som lærer har mulighed for at vurdere, om eleverne har arbejdet selvstændigt med opgaven.

UDDANNELSESLIDER, HHX

Det er altså en pointe blandt både ledere og lærere, at mundtlig opfølgning på skriftlige opgaver er et greb, der handler om at få et retvisende indblik i og mulighed for at bedømme elevernes læring og selvstændige faglige niveau i faget. En lærer i engelsk på hhx forklarer, at det også handler om at forebygge allerede erfarede eksamensoplevelser efter fremkomsten af generativ AI, hvor nogle elever er faldet flere karakterer til den skriftlige eksamen, hvor de ikke må benytte sig af teknologien, sammenholdt med deres skriftlige standpunktskarakter i faget.

Fordybelsestid bruges på opgaver med mundtlig og skriftlig formidling

Et andet greb, lærerne benytter sig af for at få indblik i og mulighed for at bedømme elevernes selvstændige arbejde med skriftlige opgaver i lyset af generativ AI, er at tilføje en mundtlig dimension til elevernes skriftlige arbejde i fagene.

Det er en pointe blandt lærerne, at de ikke kan vide sig sikre på, at det skriftlige arbejde, der typisk laves hjemme, bliver brugt til selvstændigt skriftligt arbejde pga. generativ AI, medmindre de supplerer med et mundtligt element. Det fremgår af lærerinterviewene, at det fx kan bestå i, at eleverne optager en video, hvor de i et kort oplæg reflekterer over deres skriveproces eller præsenterer hovedpointerne i deres opgave, eller laver en podcast om et emne ud fra et skriftligt oplæg. Lærerne bedømmer efterfølgende elevernes indsats ud fra en helhedsvurdering af det skriftlige og mundtlige produkt.

En lærer i dansk, mediefag og informatik fra stx og hf sætter følgende ord på, hvad denne multimediale tilgang kan bestå i og bidrage med:

”Opgaverne skal stilles på en måde, så der er et ekstra lag. Det kan fx være ved, at eleverne laver en præsentationsvideo ud fra noget skriftligt arbejde, de har lavet. På den måde kan man hurtigt få en fornemmelse af deres faglighed, og hvor meget de har styr på emnet. Det er en kontrol for mig som lærer, men også et læringsrum for eleverne. [...] Det handler om at få en fornemmelse af, hvad det er for nogle elever, man har, og hvor de ligger fagligt.

LÆRER, DANSK, MEDIEFAG OG INFORMATIK, STX & HF

5.4 Prøveformer

Som det fremgår af afsnit 5.3, betyder fremkomsten af generativ AI, at lærerne oplever at skulle bedømme og udprøve elevernes faglighed på nye måder. På baggrund af de interviewede ledere og læreres erfaringer med bedømmelse i lyset af generativ AI kan der udledes to potentielle prøveformer, der rummer et skriftligt element, uden at det går ud over eksaminator og censors mulighed for at have blik for og bedømme elevernes selvstændige faglige niveau.

De udledte prøveformer knytter sig til de skriftlige eksaminer, da mundtlige prøver med både lang og kort forberedelsestid generelt beskrives som velegnede til at imødekomme udviklingen i generativ AI's udbredelse og kapacitet. Det skyldes, at mundtlige prøvesituationer giver eksaminator og censor mulighed for at spørge ind til og bedømme elevernes selvstændige forståelse af relevante faglige mål.

Prøveformerne, der kan udledes pba. de interviewede ledere og læreres erfaringer med bedømmelse i lyset af generativ AI, afspejler i vid udstrækning tendenserne, der ses i udviklingen af bedømmeskriterier efter udbredelsen af generativ AI, som beskrevet i afsnit 5.3. Derudover stemmer prøveformerne generelt overens med pointerne i litteraturopsamlingen i kapitel 3 og erfaringerne fra universiteterne i kapitel 4, ligesom de afspejler ekspertgruppen for ChatGPT og andre digitale hjælpemidlers anbefaling om prøver, der skal gå på to ben og udprøve både med og uden digitale hjælpemidler.

Konkret fremgår det af kortlægningen, at skriftlige prøveformer med mundtligt forsvar og skriftlige prøveformer, der inkluderer elementer uden hjælpemidler, kan være to mulige veje for skriftlige prøver i lyset af generativ AI, der understøtter en valid bedømmelse af elevernes selvstændige faglige niveau. De to veje udfoldes i de følgende afsnit.

Da afsnittenes pointer bygger på kvalitative interview med ledere og lærere fra 12 forskellige gymnasier, er de ikke repræsentative for erfaringerne hos alle ledere og lærere på landets gymnasier. Afsnittene skal i stedet give et overordnet indblik i, hvad det er for nogle overvejelser, ledere og læreres første erfaringer med bedømmelse efter fremkomsten af generativ AI har givet anledning til ift. den videre udvikling af prøver i gymnasiet.

5.4.1 Skriftlige prøveformer med mundtligt forsvar

Med udbredelsen af generativ AI er det blevet muligt for eleverne at producere store mængder tekst af høj kvalitet, der gør det vanskeligt for lærerne at vurdere og bedømme elevernes selvstændige bidrag i skriftlige opgaver. Mundtlig opfølgning på skriftlige produkter fremstår derfor som en meningsfuld måde at fastholde validiteten i bedømmelsen af elevernes skriftlige arbejde, da det giver eksaminator og censor mulighed for at udprøve og bedømme, om eleverne har arbejdet selvstændigt med indholdet i deres skriftlige produkt.

Det indebærer også, at det skriftlige og mundtlige prøveelement med fordel kan tænkes sammen, så der gives én samlet karakter pba. en helhedsvurdering af elevens skriftlige og mundtlige præstation. Netop fordi det skriftlige produkt i sig selv kan være svært at bedømme retvisende.

Denne type af kombinerede skriftlige og mundtlige prøveformer kan fx dreje sig om:

1. Projekt- eller opgavebaserede prøveformer, der kombinerer et skriftligt produkt med et mundtligt forsvar
2. Portfoliopróver, hvor eleverne bliver udpróvet mundtligt pba. et udvalg af afleveringer, opgaver eller produkter, de har udarbejdet i løbet af undervisningsforløbet i et fag.

Projekt- eller opgavebaserede prøveformer med mundtligt forsvar

Det fremgår af kortlægningen, at gymnasierne har gode erfaringer med prøveformen, de allerede kender fra SRP på stx og SOP på hhx og htx, hvor eleven skal forsvare sit skriftlige produkt til en mundtlig prøve. Denne prøveform muliggør brug af generativ AI som et digitalt hjælpemiddel i arbejdet med den skriftlige opgave, fordi man som bedømmer har mulighed for at udpróve elevernes ejerskab til indholdet i opgaven og refleksioner over valg og fravalg i deres arbejdsproces gennem det mundtlige forsvar.

Det mundtlige forsvar behøver dog ikke at være pba. et større skriftligt projekt, men kan også tage udgangspunkt i skriftlige opgaver, som man kender dem fra opgavesættene til skriftlige stedpróver, hvor hjælpemidler er tilladt.

Kortlægningen viser, at det både er lærere fra forskellige gymnasietyper og fagområder, der oplever, at en projekt- eller opgavebaseret prøveform med mundtligt forsvar ville give god mening ift. de bedømmelseskriterier, som de benytter sig af efter udbredelsen af generativ AI.

Det er netop en projektbaseret prøve med mundtligt forsvar, der ligger til grund for et aktuelt prøveforsøg fra Børne- og Undervisningsministeriet med afsætning på A-niveau på hhx (BUVM, 2024c). De interviewede lærere, der skal indgå i prøveforsøget, beskriver det som et godt bud på fremtidens prøveform i faget, og de pågældende læreres kolleger fra øvrige faggrupper opfatter ligeledes prøveforsøget som interessant og potentielt oplagt til próver i andre fag.

Mundtlig prøve med afsæt i portfolio

Det fremgår ligeledes af kortlægningen, at portfoliopróver kan være en prøveform, der vil gøre det muligt at bruge de bedømmelseskriterier, lærerne fokuserer på i lyset af generativ AI. I interviewene refereres der fx til eksempler på portfoliobaserede prøveformer på htx, hvor eleverne arbejder problembaseret med løbende at identificere, undersøge og skitsere løsninger på et problem, som løbende munder ud i delafleveringer, der udvælges og samles i en portfolio. Undervisningsforløbet afsluttes i forlængelse heraf med en individuel mundtlig prøve med afsæt i elevens portfolio, der skal give indblik i elevens læreproces og udpróve elevens selvstændige forståelse af de faglige mål.

Baseret på erfaringerne hos lærere, der allerede fører elever til portfoliopróver i deres fag, er fordelene ved denne prøveform, at den giver eleverne et incitament til at arbejde seriøst og selvstændigt med de løbende opgaver og afleveringer i undervisningen, fordi de potentielt skal kunne stå på mål for dem til en mundtlig eksamen. På den måde kan denne prøveform også være med til at forbygge udlicitering af opgaver gennem brug af generativ AI.

Som med de projektbaserede prøveformer med mundtligt forsvar bliver portfolioprøver også fremhævet som en meningsfuld prøveform på tværs af gymnasietyper og fagområder, herunder naturvidenskabelige, samfundsvidenskabelige og humanistiske fag.

5.4.2 Skriftlige prøveformer, der inkluderer elementer uden hjælpemidler

Som beskrevet i afsnit 5.3.1 oplever ledere og lærere, at udbredelsen af generativ AI nødvendiggør et forstærket fokus på elevernes grundfaglighed. Det skyldes for det første, at der ifølge lærerne er nogle grundlæggende fagspecifikke kompetencer, eleverne skal kunne uafhængig af diverse hjælpemidler, og som samtidig er nødvendige at tilegne sig for at kunne indgå i en meningsfuld faglig dialog med generativ AI. Lærerne forklarer derudover, at fokuset på grundfaglighed kan give dem en føling med og et grundlag for at bedømme elevernes selvstændige tilegnelse af fagets centrale faglige mål uafhængigt af generativ AI og andre digitale hjælpemidler.

Af samme grund fremgår det af kortlægningen, at skriftlige stedprøver, der inkluderer elementer uden hjælpemidler, er svære at komme udenom i nogle fag som reaktion på udbredelsen af generativ AI. Det skyldes, at skriftlige stedprøver uden hjælpemidler dels giver eksaminator og censor et grundlag for at bedømme elevernes grundlæggende faglige niveau og dels kan skabe et incitament hos eleverne til at lægge sig i selen for selvstændigt at tilegne sig de grundlæggende kompetencer i faget, selvom de i princippet kunne udlicitere denne læring gennem brug af generativ AI.

Det er især i fremmedsprogsfag, at de interviewede læreres erfaringer med bedømmelse i lyset af generativ AI peger i retning af et behov for skriftlige stedprøver, der inkluderer elementer uden hjælpemidler. Det fremgår af kortlægningen, at en sådan prøve kunne starte uden hjælpemidler med fokus på grammatik og sprogforståelse efterfulgt af et prøveelement, hvor eleverne skal udarbejde et skriftligt produkt med hjælpemidler med fokus på mere komplekse faglige mål.

5.5 Rammevilkår

Kortlægningen viser afslutningsvis, at der er nogle rammevilkår, som det er væsentligt at være opmærksom på, hvis generativ AI bliver tilladt til udvalgte prøver i gymnasiet. Det knytter sig til lærernes kompetencer og til GDPR.

5.5.1 Lærernes kompetencer

Hvis man ønsker at tillade generativ AI ved udvalgte prøver, er lærernes kompetencer ift. brug og bedømmelse af teknologien centrale for, at eleverne så vidt muligt kan gå til prøver på lige vilkår.

De interviewede ledere og lærere peger på, at det er afgørende for, at de pågældende lærere dels kan undervise eleverne i den læringsunderstøttende brug af generativ AI, som de vil blive bedømt ud fra til eksamen, og dels kan tage højde for elevernes brug af teknologien som del af bedømmelsesgrundlaget til prøven.

Som det fremgår af afsnit 5.1 om de 12 gymnasiers initiativer og erfaringer, finder der allerede en masse intern kompetenceudvikling sted gennem videndeling i projektnetværk og kollegial oplæring, hvor lokale frontløbere lærer deres kolleger om generativ AI's muligheder og begrænsninger. Der er ligeledes initiativer, hvor lærere har meldt sig som spydspidser og modtager ekstern kompetenceudvikling i fagdidaktisk brug af generativ AI, som de sidenhen skal dele i deres faggruppe.

Det fremgår af interviewene, at de forskellige former for intern kompetenceudvikling og kollegial oplæring kan bidrage til, at alle relevante lærere oplever at være klædt på til at føre deres elever til eksamen i, hvordan de lærer *med* generativ AI i de enkelte fag.

5.5.2 GDPR

Det fremgår ligeledes af leder- og lærerinterviewene, at der er en række GDPR-relaterede udfordringer forbundet med brug af generativ AI til prøver, der kan være komplekse for gymnasierne at navigere i.

De udfordringer, der bliver peget på i interviewene, omfatter, om gymnasierne kan benytte sig af 'open source' løsninger, om eleverne kan tvinges til at oprette en profil for at benytte en sprogmodel, hvorvidt eleverne kan påkræves at gemme og dele chatdialoger, hvordan eleverne kan undgå at uploade rettighedsbeskyttet materiale i deres brug af generativ AI, og om det er etisk forsvarligt, at eleverne benytter sig af sprogmodeller, hvor deres data deles med en ekstern tredjepart.

Ifølge ledere og lærere er det vigtigt at forholde sig til disse overvejelser og udvikle tværinstitutionelle netværk, som kan være med til at adressere dem.

Appendiks A – Litteraturliste

- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). *Generative AI Can Harm Learning*. The Wharton School Research Paper.
- Blom, J. N., Holsting, A., & Svendsen, J. T. (2024). *På sporet af chatbottens sproglige fingeraftryk: En sproglig ophavsanalyse af tekster skrevet af danskstuderende og ChatGPT*. *Nydanske Sprogstudier* 65, 79-110.
- Børne- og Undervisningsministeriet (BUVM). (2024a). *Ekspertgruppe klar med anbefalinger for brug af ChatGPT ved prøver*. Lokaliseret d. 5. november 2024 via: <https://www.uvm.dk/aktuelt/nyheder/uvm/2024/april/240424-ekspertgruppe-klar-med-anbefalinger-for-brug-af-chatgpt-ved-proever>.
- Børne- og Undervisningsministeriet (BUVM) (2024b). *Prøver og eksamen på gymnasiale uddannelser*. Lokaliseret d. 4. december 2024 via: <https://www.uvm.dk/gymnasiale-uddannelser/proever-og-eksamen>.
- Børne- og Undervisningsministeriet (BUVM) (2024c). *Forsøg med prøven i afsætning A, hhx*. Lokaliseret d. 25. oktober 2024 via: <https://www.uvm.dk/gymnasiale-uddannelser/proever-og-eksamen/forsoeg/forsoeg-med-gruppeproeve-i-afsætning-a-paa-hhx>.
- Caviglia, F., Dalsgaard, C., Boie, M. A. K., & Thomsen, M. B. (2021). *Undersøgelse af digitale teknologier i fremmedsprogundervisningen. Sproglærares anvendelse af digitale teknologier I gymnasiet*. Aarhus: Danmarks Institut for Pædagogik og Uddannelse, Aarhus Universitet.
- Chiu, T. K. F. (2023). *The impaCT OF Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney*. *Interactive Learning Environments*.
- Dalsgaard, C. & Prilop, C.N. (2024). *Elevens brug af generativ AI: Partnerskaber mellem elever og AI*. Artikel indleveret til publicering.
- Danske Gymnasier (2024). *Danske Gymnasiers strategi for digital teknologi. Læring om og med digital teknologi*. Lokaliseret d. 29. november 2024 via: <https://danskegymnasier.dk/danske-gymnasier-offentliggør-sin-nye-digitaliseringsstrategi/>.

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 49:3, 363-375.

Gymnasieskolen. (2024a). *AI forandrer undervisning og truer eksamen: Ministeren afviser dog hurtige løsninger*. Lokaliseret d. 5. november 2024 via: <https://gymnasieskolen.dk/articles/ai-forandrer-undervisning-og-truer-eksamen-ministeren-afviser-dog-hurtige-loesninger/>.

Gymnasieskolen. (2024b). *Lærere: Vi kan ikke fortsætte eksamen på denne måde*. Lokaliseret d. 3. december 2024 via: <https://gymnasieskolen.dk/articles/laerere-vi-kan-ikke-fortsaette-med-eksamen-paa-denne-maade/>.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education sciences*, 13, 410

Mao, J., Chen, B., & Liu, C. J. (2023). Generative Artificial Intelligence in Education and Its Implications for Assessment. *TechTrends*, 68:58-66.

Milano, S., McGrane, J.A., & Leonelli, S. (2023). Large language models challenge the future of higher education. *Nature Machine Intelligence* 5, 333–334.

Perkins, M., Furze, L., Roe, J., MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice*, 21(6).

Ritzau. (2024). *Elever kræver handling efter nye AI-anbefalinger*. Lokaliseret d. 5. november 2024 via: <https://via.ritzau.dk/pressemeddelelse/13810816/elever-kræver-handling-efter-nye-ai-anbefalinger?publisherId=13561703&lang=da>.

Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. *PLoS ONE*, 19(6): e0305354.

UNESCO. (2021). *AI and Education. Guidance for policy-makers*. Paris: UNESCO.

Wecks, J. O., Voshaar, J., Plate, B. J., & Zimmermann, J. (2024). Generative AI Usage and Academic Performance. *Electronic Journal*. April 2024

Appendiks B – Figur og tabelbilag

Appendiks-figur B.1

De syv universiteter i kortlægningen

- Aalborg Universitet (AAU)
- Copenhagen Business School (CBS)
- Danmarks Institut for Pædagogik og Uddannelse (DPU), Aarhus Universitet
- Danmarks Tekniske Universitet (DTU)
- Københavns Universitet (KU)
- Roskilde Universitet (RUC)
- Syddansk Universitet (SDU).

Note: EVA har som del af kortlægningen gennemført semistrukturerede interview á ca. 30-40 minutters varighed med repræsentanter for de ovenstående syv danske universiteter.

Appendiks-figur B.2

De 12 gymnasier i kortlægningen

- Aalborg Handelsskole (hhx)
- Aarhus Akademi (Hf & studenterkursus)
- Aarhus Business College (hhx)
- Favrskov Gymnasium (stx)
- Gefion Gymnasium (stx)
- Herning Gymnasium (stx)
- Hjørring Gymnasium (stx & hf)
- Horsens Gymnasium og hf (stx & hf)
- Learnmark Horsens (htx & hhx)
- Odense Tekniske Gymnasium (htx)
- Rosborg Gymnasium og hf (stx & hf)
- UCRS - Uddannelsescenter Ringkøbing-Skjern (htx & hhx).

Note: EVA har gennemført semistrukturerede interview á ca. 45 minutters varighed med ledelsesrepræsentanter og semistrukturerede gruppeinterview á ca. 60 minutters varighed med lærere fra de ovenstående 12 gymnasier.

Appendiks-tabel B.1

Oversigt over gymnasiernes initiativer

Institution	Retning	Fag/fokus	Indsats
Aalborg Handelsskole	Hhx	Sprogfag	"Brug af kunstig intelligens i engelskundervisningen på gymnasiale uddannelser: sprogteknologi, literacy og læringsforløb". Formålet med projektet er at skabe viden og erfaringer, som kan blive til generelle anbefalinger for nutidig og fremtidig brug af kunstig intelligens-baserede sprogteknologier i den gymnasiale engelskundervisning, og som samtidig forventes at kunne udbredes til øvrige fremmedsprogsfag. Engelsklærere modtager undervisning og træning i digital literacy inden for de kunstig intelligens-baserede sprogteknologier gennem en række workshops med forskere fra Aalborg Universitet (AAU). Med udgangspunkt i dette skal lærerne udvikle læringsdesigns, som de skal afprøve i undervisningen. Del af projekt på Aalborg Universitet sammen med Aalborg Katedralskole (stx). Løber fra 2023-2026.
Aarhus Akademi	Hf og studenterkursus	Uspecificeret	"ChatGPT/AI i undervisningen". Aktionslæringsprojekt, der undersøger, hvordan man kan tænke AI og sprogmodeller ind i undervisningen, hvor det er meningsfuldt, og hvordan og hvornår man kan og bør begrænse AI og sprogmodellers indflydelse ift. fx skriftlige opgaver, den personlige stemme, kritisk tænkning og kildekritik og at turde give sig i kast med det, man ikke har svar på. Evt. også fokus på lærerplaner og afprøvning af eksamensformer, som i højere grad afspejler den nye virkelighed med AI og sprogmodeller. I samarbejde med Skanderborg Gymnasium og Randers Statsskole. Projektet løber fra 2023-2025.
Favrskov Gymnasium	Stx		
Horsens Gymnasium og HF	Stx og hf		
Aarhus Business College	Hhx	Skriftlighed (som udgangspunkt)	"Skriftlighed forfra". Formålet med projektet er at kortlægge forståelser af, hvad der udgør plagiat ifm. skriftligt arbejde, og hvordan man kan arbejde med at skabe nye veje (kulturelle, strukturelle og didaktiske) til at modvirke plagiat og sætte fokus på elevernes forståelse af egen læring. Faggrupper skal undersøge hjælpemidlers muligheder og begrænsninger og give forslag til, hvordan man didaktisk kan integrere dem i den daglige undervisning samt tilrettelægge skriftlighed, så man sikrer en større involvering af læreren i elevernes skriveproces. Projektet er samtidig et forsøg på at afdække elevens holdninger til plagiat og skabe en kulturændring, ift. hvad der er etisk og fagligt korrekt. Projektet er støttet af Region Midtjylland og sker i samarbejde med forskere fra Aarhus Universitet. Projektet løber fra 2023-2024.
Herning Gymnasium	Stx		
Uddannelsescenter Ringkøbing-Skjern (UCRS)	Hhx og htx		
Gefion Gymnasium	Stx	Uspecificeret	Strategisk fokus på brug af generativ AI i undervisningen forankret i faggrupper med opstart i skoleåret 2024-25.

Institution	Retning	Fag/fokus	Indsats
Hjørring Gymnasium	Stx og hf	Uspecificeret	"AI – ven eller fjende?". Afprøvning af kunstig intelligens i alle fag. Efteruddannelse udvikles i samarbejde med lærere, så det bliver relevant og tager udgangspunkt i det, lærerne har brug for at vide mhp. meningsfuldt at kunne didaktisere brugen af kunstig intelligens i undervisningen. To halve kursusdage for alle medarbejdere. Desuden er der udpeget 17 pilotlærere med særlig forpligtelse til at afprøve faglige forløb. Pilotlærere får to ekstra halve kursusdage. Projektet er støttet af Region Nordjylland og sker i samarbejde med forskere fra Aalborg Universitet. Projektet løber fra 2024-2026.
Learnmark Horsens	Hhx og htx	Sprogfag	"AI – Didaktik, sprog og kompetencer til fremtiden". Formålet med projektet er at skabe tre AI-didaktiske modeller og en AI-kompetencemodel. Disse skal give svar på, hvordan skoler kan bruge AI i undervisningen til at differentiere, fremme læring og sikre, at eleverne opnår AI-kompetencer, der er relevante på arbejdsmarkedet. Projektet fokuserer særligt på sprogfagene på det erhvervs-gymnasiale område. Projektet er støttet af Region Midtjylland og sker i samarbejde med en række erhvervsskoler, Aarhus Universitet og Center for It i Undervisningen. Projektet løber fra 2024-2026.
Odense Tekniske Gymnasium	Htx	Strategisk fokus og AI-studieretning	Strategisk fokus på brug af generativ AI på uddannelsen og oprettelse af ny studieretning i kunstig Intelligens fra og med skoleåret 2024-25.
Rosborg Gymnasium	Stx og hf	Strategisk fokus og digital studieretning	Strategisk fokus på brug af generativ AI i undervisningen og en eksplicit forventning om, at alle lærere anvender teknologien og deler viden og erfaringer på tværs. Derudover oprettelse af ny digital studieretning, hvor generativ AI indgår i undervisningen i alle fag, fra og med skoleåret 2023-24.

Kilde: Initiativerne er beskrevet pba. projektbeskrivelser og kvalitative interview med ledelsesrepræsentanter fra de 12 gymnasier.

Appendiks-tabel B.2

Eksempler på undervisningsforløb og læringsformål med brug af generativ AI

Fag	Forløb	Formål
Dansk (hhx) – Argumentation og retorik med AI	<p>6 lektioner:</p> <p><i>Lektion 1:</i> Lærer introducerer til emnet (debatten om kongehusets rolle i et demokratisk samfund ifm. dronningens abdicering). Introduktionen efterfølges af brainstorm over centrale begreber, som eleverne tidligere har lært om: Først stilles hv-spørgsmål til ChatGPT om de enkelte begreber. Dernæst bedes ChatGPT komme med argumenter (med påstand og belæg) for og imod kongehuset. Eleverne skal derefter reflektere over, om ChatGPT bidrager til at skabe klarhed over begreberne.</p>	<p>At øge elevernes forståelse af argumentation og retorik i den offentlige debat.</p> <p>At give eleverne erfaringer med vigtigheden af gode prompts.</p> <p>At gøre eleverne mere bevidste om, hvilke muligheder og begrænsninger der er ved brug af AI til forskellige formål.</p>

Fag	Forløb	Formål
	<p><i>Lektion 2-3:</i> Elever inddeler kronprins Christians tale (indledning, midte inkl. underemner, afslutning) og sætter fokus på hans argumenter ift. at være ung i dag. Dernæst skal eleverne gruppevis inddrage ChatGPT i analysen af et afsnit. Eleverne skal efterfølgende reflektere over, hvilke begrænsninger og muligheder ChatGPT har ift. at analysere sammenhængende tekster.</p> <p><i>Lektion 4-5:</i> Elever diskuterer rækkefølgen af temaer og argumenter i kronprinsens tale samt vurdering af gennemslagskraft. Diskussionen efterfølges af fælles læsning af debatindlæg om kongehuset i Politiken. Elever diskuterer dernæst hovedpåstanden i debatindlægget og arbejder fælles med at identificere indlæggets argumenter. Til sidst anvender eleverne ChatGPT til at finde inspiration til argumenter og modargumenter for kongehuset som institution, hvilket skal danne afsæt for en disposition til en aflevering. Eleverne skal reflektere over, hvorledes ChatGPT kan anvendes til at danne sig et overblik over gængse emner for fremadrettet at kunne anvende det som et værktøj til at brainstorme.</p> <p><i>Lektion 6:</i> Elever udformer disposition til et debatindlæg, der skal være et modsvar til det læste i Politiken. Eleverne "omskriver" gruppevis kronprinsens tale ved at instruere ChatGPT i at skrive i en anden stil el. toneleje. Herefter skal eleverne markere og forklare de ord, der anvendes til at ændre tonen. Eleverne definerer dernæst den tone, de gerne vil skrive deres eget indlæg i. Eleverne skal afslutningsvis reflektere over, hvordan ChatGPT kan anvendes til at synliggøre teksters ordvalg, og hvordan ordvalg kan påvirke sprogets påvirkning af modtageren.</p>	
Matematik B (stx) – ChatGPT som kognitiv partner	<p>2 x 70 min.:</p> <p>Eleverne skal indlede en selvstændig dialog med ChatGPT: "Kan du introducere mig for begrebslæring i matematik?" og få ChatGPT til at komme med eksempler.</p> <p>Eleverne skal lave en Powerpoint-præsentation, hvor de forklarer begreberne med egne ord med udgangspunkt i følgende punkter:</p> <ul style="list-style-type: none">• Hvad er en kontinuert funktion?• Hvad er en differentiabel funktion?• Hvad er den afledte funktion?• Hvad er funktionens monotoniforhold?• Hvad er sammenhængen mellem en funktions monotoniforhold og den afledte funktion?	<p>At undersøge, om AI kan forbedre elevens læring i matematik med fokus på den daglige begrebslæring, som tester elevens evne til at anvende AI som en kognitiv partner på en hensigtsmæssig måde og dermed opnå læring inden for en udvalgt fagspecifik teori.</p> <p>At gøre eleverne i stand til at definere, fortolke og anvende en differentialkvotient.</p>
Samfundsfag A (stx) – ChatGPT som kognitiv partner	<p>Spredd over flere moduler:</p> <p><i>Introduktion til AI som sprogmodel (to moduler):</i> Hvad kan AI, og hvad kan AI ikke?</p>	<p>At introducere og repetere samfundsfaglige begreber.</p> <p>At kvalificere elevernes arbejde med problemformuleringer og SRP.</p>

Fag	Forløb	Formål
	<p><i>Øvelse 1:</i> Eleverne skal afprøve, om det er muligt at bruge AI til at lave en samfundsfaglig hypotese, der følger den korrekte samfundsvidenskabelige form. Eleverne skal skrive fem hypoteser ned, der tager afsæt i teorier om ulighed i USA. Derefter skal eleverne stille spørgsmål til ChatGPT inden for hver af hypoteserne og udvælge eksempler på svar. Eleverne skal faktatjekke ChatGPT's besvarelser og undersøge, om der er sammenhæng mellem teorierne/hypoteserne og ChatGPT's besvarelser.</p> <p><i>Øvelse 2a:</i> Eleverne skal bruge AI til at forbedre deres problemformuleringer og ideer til SRP. Eleverne skal bede ChatGPT om at lave tre spørgsmål til det emne, de har valgt. Derefter skal de blive ved med at prompte, indtil den har formuleret undersøgelsesspørgsmål, hvor de taksonomiske niveauer er tydelige. Herefter skal eleverne sammenligne spørgsmålene med de spørgsmål, de selv har lavet. De må også gerne inlæse deres egne undersøgelsesspørgsmål til chatten for at se, om ChatGPT kan forbedre dem. Dernæst skal de overveje, hvilke spørgsmål der fungerer bedst.</p> <p><i>Øvelse 2b:</i> Eleverne skal bruge ChatGPT som inspirationspartner til SRP. De skal give ChatGPT et emne og se, om den kan komme med forslag til, hvad de kan skrive om. De skal stille spørgsmål og bede den uddybe undervejs, så de kommer tættere på noget, de kan bruge til SRP.</p> <p><i>Øvelse 3</i> (del af et repetitionsforløb): Eleverne skal bruge ChatGPT til at teste, hvad de kan huske om økonomisk teori. De skal give ChatGPT følgende prompt: ”Lad os spille et escape room-spil: Jeg lærer om økonomi i samfundsfag i gymnasiet. Jeg går i 3. g. Spillet går ud på, at du forklarer et begreb, og jeg skal gætte, hvad det er. Kun når jeg har gættet rigtigt, må jeg gå til næste rum. Okay?” Eleverne skal parvis diskutere sig frem til de rette svar.</p> <p><i>Øvelse 4</i> (opsamling på et forløb): Eleverne skal bruge AI (DALL.E) til at generere et billede, der består af økonomiske begreber. Herefter skal de se, om de kan genkende og forklare de begreber, billedet er bygget op af. Dernæst skal eleverne bruge AI til at lave en liste – uden forklaringer – over de 20 vigtigste begreber i økonomi i samfundsfag A på gymnasiet. Eleverne forklarer begreberne med afsæt i deres noter. De må gerne bruge AI til at tjekke, om deres forklaring stemmer overens med den AI-genererede forklaring. Begreber, der ikke er gennemgået, springes over.</p> <p><i>Øvelse 5</i> (start på forløb om ligestilling og køn): Eleverne skal starte en diskussion med ChatGPT om køn. Når det første spørgsmål er stillet, skal eleverne forholde sig til det svar, de har fået, og evt. faktatjekke. Eleverne kan også bede ChatGPT om at påtage sig en rolle, den taler ud fra. Til sidst skal klassen i fællesskab reflektere over diskussionen, herunder faldgruber som politisk korrekthed i diskussion med ChatGPT</p>	<p>At give eleverne en forståelse af, hvad AI og en sprogmodel er.</p> <p>At give eleverne erfaringer med vigtigheden af gode prompts.</p> <p>At gøre eleverne mere bevidste om, hvornår det er hensigtsmæssigt at bruge AI.</p>

Fag	Forløb	Formål
Engelsk (stx) – Immerse yourself in a literary or historical character	<p><i>Første del:</i> Eleverne skal bede ChatGPT eller lignende sprogmodel om at have en dialog med dem, hvor eleven skal være en litterær eller historisk person, fx Scout fra ”To Kill a Mockingbird”.</p> <p><i>Anden del:</i> Eleverne skal bede ChatGPT om at stille dem meningsfulde spørgsmål, der knytter sig til den karakter, de har valgt at være. Eleverne skal blive i deres rolle og svare på spørgsmålene i samme stil, som deres karakter ville have gjort.</p> <p><i>Tredje del:</i> Eleverne skal samle fem eksempler fra dialogen, der illustrerer, at de har spillet deres rolle, og dele dem med klassen.</p> <p><i>Efterrefleksion:</i> Eleverne skal overveje, hvordan brugen af ChatGPT henholdsvis har understøttet og hæmmet deres læreproces.</p>	<p>At organisere læring på en måde, så generativ AI bliver en kognitiv samarbejdspartner og ikke et sted, læring udliciteres til.</p> <p>At understøtte fagets formål om at læse litteratur, at sætte litteraturen ind i en litterær, historisk og samfundsmæssig kontekst, at analysere litteratur og sprog og at bruge akademisk sprog i praksis.</p> <p>At gøre eleverne bevidste om, hvilke muligheder og begrænsninger der er forbundet med brugen af generativ AI.</p>
Spansk (hhx) – ChatGPT som uformel chatpartner	<p><i>Intro:</i> Kort lærerpræsentation af forløb og begrebet ”uformel skriftlighed”. Dernæst reflekterer eleverne i par over fordele og ulemper ved brug af ChatGPT med afsæt i refleksionsspørgsmål.</p> <p><i>Forberedelse:</i> Lærerpræsentation af rammer og fremgangsmåde for opgaven, inkl. forslag til prompts.</p> <p><i>Udførelse:</i> Eleverne sidder individuelt og chatter på spansk via ChatGPT med ”deres unge spanier”. Eleverne kan udveksle erfaringer med sidemakker og spørge læreren til råds.</p> <p><i>Afslutning:</i> Eleverne afslutter deres chat og deler den med læreren.</p> <p><i>Efterrefleksion og opsamling:</i> Eleverne sidder i par, og reflekterer over opgaven og brugen af ChatGPT som chatpartner med afsæt i refleksionsspørgsmål. Afslutningsvis er der fælles opsamling og evaluering.</p>	<p>At træne elevernes evne til at korrespondere uformelt skriftligt på spansk.</p> <p>At træne elevernes promptkompetence.</p> <p>At eleverne bliver opmærksomme på fordele og ulemper ved brug af ChatGPT.</p>
Programmering C (htx) – ChatGPT som software udviklingsværktøj	<p><i>Opstart:</i> Læreren præsenterer formålet med dagens lektion og sætter gang i en plenumdiskussion af værdien af at anvende indekserede variable.</p> <p><i>Udførelse:</i> Eleverne skal gruppevis anvende ChatGPT til at få kodeinput til software, der indeholder indekserede variable. Eleverne skal sikre sig, at de forstår alle programlinjerne og den overordnede kodestruktur.</p> <p><i>Efterrefleksion:</i> Gruppediskussion og plenumopsamling om udbyttet af at bruge ChatGPT, og hvad man gør, hvis ChatGPT bidrager med en kode, som man ikke forstår.</p>	<p>At give eleverne forståelse for værdien og nødvendigheden af at anvende indekserede variable i softwareprogrammer.</p> <p>At give eleverne erfaringer med fordele og ulemper ved brug af ChatGPT til softwareudvikling.</p>

Kilde: Undervisningsforløbene er beskrevet pba. materialer fra forløb udviklet af Herning Gymnasium, Aarhus Business College og UddannelsesCenter Ringkøbing Skjern i regi af projektet ”Læring i feltet mellem faglige hjælpemidler og plagiat”, som kan findes via projektbanken hos Region Midtjylland, der har støttet projektet: <https://www.projektbanken-rm.dk/laering-i-feltet-mellem-faglige-hjaelpemidler-og-plagiat>.

Kortlægning af bedømmelsespraksis i gymnasiet i lyset af generativ AI

© 2025 Danmarks Evalueringsinstitut

Citat med kildeangivelse er tilladt

Publikationen er kun udgivet i elektronisk form på: www.eva.dk

Foto: iStock

ISBN (www) 978-87-7182-780-4

Danmarks Evalueringsinstitut
– analyser, redskaber og udvikling, der styrker uddannelser og
dagtilbud.



Danmarks
Evalueringsinstitut

+45 35 55 01 01
eva@eva.dk
www.eva.dk