



FORSKNINGSOVERSIGT · MAJ 2026

Kunstig intelligens i grundskolen

*Hvad vestlig forskning faktisk siger om AI til 1.-9. klasse,
og hvad en dansk lærer kan stille op med det*

1.-9. KLASSE

NORDEN & VESTLIGE LANDE

2022-2026

© 2026 Martin Hasbirk Brandt

OM DENNE OVERSIGT

Det her er en gennemgang af den vestlige forskning om kunstig intelligens i grundskolen.

Rapporten dækker randomiserede forsøg, metaanalyser, implementeringsstudier og politiske dokumenter fra Norden, EU, Storbritannien, USA og OECD-niveau. Tidsmæssigt ligger tyngden på 2022–2026, altså perioden efter ChatGPT, men ældre fundamentale studier af intelligente tutorsystemer er taget med, hvor det giver mening.

Rapporten er ikke et neutralt resumé. Den er en filtrering. Jeg har valgt det stærkeste evidensgrundlag frem for det hyppigst citerede, og jeg har vægtet nordiske kilder, når de findes. Mange påstande, du møder i debatten, er bygget på meget tynde data. Det forsøger jeg at sige direkte.

Udarbejdet: Maj 2026

Geografisk fokus: Norden, EU, Storbritannien, USA, OECD

Klassetrin: 1.–9. klasse (cirka 6–15 år)

Bemærk: Forskningsfeltet bevæger sig hurtigt. Tjek datoerne på de enkelte studier, før du citerer dem videre.

© 2026 Martin Hasbirk Brandt. Alle rettigheder forbeholdes.

Indhold

—	Sådan læser du tallene	04
—	Hovedpointer	06
01	Effekter på læringsudbytte	08
02	Motivation, engagement og specialundervisning	14
03	Effekter på lærerne	16
04	Kognitive risici	18
05	Politiske rammer i Norden og Vesten	20
06	Implementering og kompetenceudvikling	25
07	Sammenligning af AI-værktøjstyper	26
08	Centrale studier	28
09	Anbefalinger	31
10	Forbehold	34

Sådan læser du tallene

Rapporten henviser til en del studier, og de bruger nogle ord og forkortelser, der ikke nødvendigvis siger sig selv. Her er en kort oversigt, du kan vende tilbage til.

RCT (randomiseret kontrolleret forsøg). Den stærkeste form for evidens i pædagogisk forskning. Eleverne (eller lærerne) bliver fordelt tilfældigt mellem to grupper: én, der bruger AI, og én, der ikke gør. Bagefter sammenligner man resultaterne. Når fordelingen er tilfældig, kan forskellen mellem grupperne ikke skyldes, at den ene gruppe var dygtigere eller mere motiveret fra start. Det er det, der gør et RCT til guldstandarden.

n. Antal deltagere i et studie. Et studie med $n=1.000$ er typisk mere troværdigt end et med $n=20$, fordi tilfældige udsving udjævnes ved store tal.

P-værdi. Sandsynligheden for, at det resultat, man har målt, kunne være opstået ved tilfældigheder alene. Lave p-værdier betyder, at det er usandsynligt, at resultatet er et tilfælde. p mindre end 0,05 betragtes typisk som "statistisk signifikant". p mindre end 10^{-8} betyder mindre end én ud af 100 millioner. Det er ikke til diskussion.

Effektstørrelse (Hedges' g, Cohen's d). Et tal, der viser, hvor stor en forskel der er mellem to grupper, korregeret for naturlig spredning i målingerne. Tommelfingerregel: 0,2 er en lille effekt, 0,5 er moderat, 0,8 er stor. Når en metaanalyse rapporterer $g = 0,46$, betyder det, at AI-gruppen i gennemsnit klarer sig moderat bedre end kontrolgruppen.

Metaanalyse. Et studie af studier. Forskerne samler resultaterne fra mange enkeltstudier og regner en samlet effekt ud. Det giver et mere robust billede end ét enkelt studie, men det er stadig kun så godt som de underliggende studier.

Scoping review og systematisk gennemgang. To slags litteraturoversigter. En scoping review kortlægger, hvad der findes af forskning på et område. En systematisk gennemgang er strengere og bedømmer kvaliteten af de fundne studier efter faste kriterier.

Peer-reviewed kontra preprint. Et peer-reviewed studie er blevet vurderet og godkendt af andre forskere før udgivelse. Et preprint (typisk på arXiv eller SSRN) er offentliggjort før den vurdering. Preprints kan godt være korrekte, men tallene bør altid behandles som foreløbige.

Title I-skoler. Amerikansk betegnelse for offentlige skoler med en høj andel elever fra økonomisk dårligt stillede familier. De modtager ekstra føderal støtte. Når et studie er gennemført på Title I-skoler, betyder det, at resultaterne kommer fra et udsat skolemiljø.

Kort fortalt

Forskningsbilledet er reelt blandet, og det er *designet* af AI-værktøjet, ikke selve tilstedeværelsen, der afgør udfaldet. De studier, der lykkes, er dem, hvor AI er bygget med pædagogisk tøjle.

TRE TING AT TAGE MED

- **Veldesignede AI-tutorer virker, fri ChatGPT-adgang gør det modsatte.** AI-tutorer med pædagogisk stilladsering (Kestin/Harvard 2024–25, Stanford Tutor CoPilot 2024, Eedi/LearnLM 2025, Verdensbankens Nigeria-forsøg 2024) giver moderate til store positive effekter på læring. Ufiltreret adgang til ChatGPT kan derimod direkte skade indlæringen af færdigheder (Bastani et al., PNAS 2025; Kosmyna et al., MIT 2025). Metaanalyser viser samlet en moderat til stor effekt af ChatGPT (Hedges' g cirka 0,45 til 0,87, hvor 0,5 typisk regnes som moderat og 0,8 som stor). Men evidensgrundlaget for grundskolen er tyndt, korte forløb dominerer, og næsten ingen stringente randomiserede forsøg findes for indskoling og mellemtrin (1.–6. klasse).
- **I dansk grundskole spredt AI sig hurtigt, men ujævnt og foran evidensen.** Norske tal fra 2025 viser, at 65 procent af lederne på barnetrinnet og 89 procent på ungdomstrinnet rapporterer, at eleverne bruger AI. Svenske Skolverket finder, at kun cirka 25 procent af lærerne på de yngste klassetrin og omkring 50 procent af lærerne på de ældste klassetrin selv bruger AI, og blot én ud af seks tillader, at eleverne gør det. Danmark er længst fremme med pædagogisk infrastruktur (SkoleGPT, BUVM's syv anbefalinger og STIL's syv-trinsvejledning fra 2025). Sverige er strammest på prøveafvikling, mens Finland og Norge fokuserer på AI-forståelse og aldersgraduering.
- **De største praktiske risici er kognitiv aflastning, GDPR-eksponering og en voksende implementeringskløft.** EU's AI-forordning klassificerer flere skole-AI-anvendelser som højrisiko. UNESCO anbefaler en aldersgrænse på 13 år. Tal fra RAND og Skolverket viser, at lærerkompetencer og retningslinjer halter 12 til 24 måneder efter den faktiske brug. Den forsvarlige position lige nu er en varsom, alderssvarende integration, hvor AI-anvendelsen er designet, så eleverne stadig skal yde en kognitiv anstrengelse selv. Brug GDPR-kompatible værktøjer som SkoleGPT, og gør AI-kompetenceudvikling for lærerne obligatorisk.

Hvad siger den stærkeste evidens?

Den korte version: AI virker, hvis værktøjet er sat op til at stille spørgsmål i stedet for at servere svar. Når det ikke er, lærer eleverne mindre. Det er det enkleste mønster i hele litteraturen, og det er overraskende konsistent på tværs af studier.

POSITIVE EFFEKTER VED OMHYGGELIGT DESIGN

Kestin, Miller, Klaes et al. (Harvard, Scientific Reports, juni 2025)

RCT · N=194 UNIVERSITETSSTUDERENDE · FYSIK

Et skræddersyet GPT-4-baseret tutorforløb (PS2 Pal) med stram pædagogisk styring (ét trin ad gangen, ingen komplette løsninger, kun korte svar) slog traditionel klasseundervisning. De studerende lærte mere end dobbelt så meget på kortere tid (median 49 mod 60 minutter), og engagementet steg. Resultatet var så markant, at sandsynligheden for at det skyldes tilfældigheder er mindre end én ud af 100 millioner. Vigtigt forbehold: deltagerne er universitetsstuderende, ikke grundskoleelever, og der er kun målt på umiddelbar efter-test.

Stanford Tutor CoPilot (Wang et al., 2024)

PRÆREGISTRERET RCT · 900 LEKTIEHJÆLPERE, ~1.800 ELEVER · GRUNDSKOLEMATEMATIK · USA (TITLE I-SKOLER)

Eleverne blev undervist online af menneskelige lektiehjælpere. Halvdelen af lektiehjælperne havde et AI-værktøj kørende ved siden af, der løbende foreslog dem sokratiske spørgsmål, mens de underviste. De elever, hvis lektiehjælper havde værktøjet, scorede 4 procentpoint højere på faglige mestringsmål. Hele 9 procentpoint højere, når lektiehjælperen oprindeligt havde lave bedømmelser. Pris: omkring 20 dollar per lektiehjælper om året. Mekanismen var, at værktøjet skubbede lektiehjælperen i retning af at stille vejledende spørgsmål i stedet for at give svaret. Det her er den klareste demonstration af, at AI brugt som hjælpemiddel for læreren kan løfte de svageste lektiehjælpere mest, og det er den model, der lader sig overføre mest direkte til dansk grundskole.

Eedi/LearnLM RCT (Storbritannien, 2025)

RCT · N=165 ELEVER · 13-15 ÅR · 5 SECONDARY SCHOOLS

Menneskelige tutorer overvågede udkast fra LearnLM, Googles pædagogisk finjusterede model. Tutorerne godkendte 76,4 procent af modellens output med ingen eller minimale ændringer. Eleverne, der fik AI-suppleret undervisning, klarede sig bedre end sammenligningsgruppen. Studiet er væsentligt, fordi alderen rammer direkte ind i udskolingen.

Verdensbankens Nigeria-forsøg (De Simone et al., 2024–25)

RCT · 800 ELEVER · 12–17 ÅR · 6 UGERS EFTERMIDDAGSFORLØB

Engelskundervisning med Microsoft Copilot gav læringseffekter svarende til omkring to års almindelig undervisning. Pigerne fik mest ud af det, og effekten smittede af på de almindelige eksamener uden for forløbets eget indhold. Forbeholdet er vigtigt: det er en kontekst med lav- og mellemindkomster, hvor sammenligningsgrundlaget er en svag undervisning fra start. To-årstallet er kalibreret mod et lavt udgangspunkt, ikke mod dansk grundskolestandard.

Carnegie Learning MATHia og Cognitive Tutor

LANG RCT-HISTORIE · GRUNDSKOLEMATEMATIK · USA

Et amerikansk RAND-finansieret studie af 18.000 elever på 147 skoler fandt næsten en fordobling af elevernes fremgang på standardiserede tests i andet år af forsøget, hvor systemet blev brugt som supplement til almindelig klasseundervisning. Det er fundamental evidens for, at *intelligente tutorsystemer*, en ældre og distinkt klasse af AI sammenlignet med de generative chatbots, leverer robuste effekter på mellemtrin og i udskolingen i matematik.

NEGATIVE ELLER INGEN EFFEKT UDEN STYRING

Eleverne brugte ChatGPT som krykke. De kopierede løsninger, og de mærkede ikke selv, at de havde lært mindre.

BASTANI ET AL., PNAS 2025 · 1.000 ELEVER I 9.–11. KLASSE, TYRKIET

Bastani, Bastani, Sungu, Ge, Kabakçı og Mariman (PNAS 2025)

RCT · ~1.000 ELEVER · 9.–11. KLASSE · TYRKISK GYMNASIUM

Tre forsøgsarme: kontrol, GPT-Base og GPT-Tutor med sikkerhedsskinner. Med AI-adgang under øvelserne steg præcisionen 48 procent (Base) og 127 procent (Tutor). Men på den efterfølgende eksamen *uden* AI scorede GPT-Base-eleverne 17 procent *lavere* end kontrolgruppen. GPT-Tutor lå nogenlunde på linje med kontrolgruppen. Mekanismen var krykke-effekten: eleverne kopierede løsninger uden at lære, og de bemærkede ikke selv noget tab. Det her er det vigtigste enkeltstudie at læne sig op ad i dansk grundskolepolitik. Det viser med tal, at rå ChatGPT-adgang aktivt skader indlæringen af færdigheder.

Kosmyna et al., "Your Brain on ChatGPT," MIT Media Lab (2025)

EEG-STUDIE · N=54 (18 I SESSION 4) · 18–39 ÅR · AKADEMISKE ESSAYS

Forskerne målte hjerneaktivitet med EEG, mens deltagerne skrev essays. ChatGPT-brugerne udviste den svageste hjerneaktivitet og det laveste netværk af forbindelser mellem hjerneområder. De producerede sjælløse og homogene essays og aflastede gradvist mere af opgaven til værktøjet. Forbehold: lille n, ikke peer-reviewed (jævnfør Stanković et al.'s kritik på arXiv 2026), og deltagerne var voksne snarere end børn. Men mekanismen, kognitiv gæld, er gentaget i mange andre studier.

Mindre eksperimenter peger samme vej: Niloy et al. (2023), Yang et al. (2025) og Tong et al. (2025) finder alle, at ChatGPT-assisterede elever klarer sig dårligere, når AI fjernes ved bedømmelsen. En samling i Hechinger Report fra 2024–25 opgør, at af de seks mest stringente RCT'er på grundskoleområdet, vi har, har tre positive effekter, to ingen effekt, og ét er decideret skadeligt. Værktøjsdesignet forklarer størstedelen af variationen.

HVAD METAANALYSERNE SIGER SAMLET

Wang og Fan med flere (Liu, JCAL 2025; Nature HSSC 2025/26) har lavet metaanalyser af henholdsvis 35 og 51 ChatGPT-studier fra 2022–2025. De rapporterer en samlet effekt på læringspræstation, der ligger i moderat til stor størrelsesorden (Hedges' g cirka 0,46 til 0,87), med mindre effekter på kritisk tænkning og tænkning på højere niveauer. Næsten alle primærstudier er korte (otte uger eller derunder), tunge på universitetsniveau og ofte gennemført af AI-entusiaster.

Zhang og Tur (*European Journal of Education*, 2024) har lavet den eneste systematiske gennemgang specifikt af ChatGPT i grundskolen og fandt kun 13 studier på verdensplan. De peger på fordele i personalisering og idégenerering, og på trusler mod prøveafvikling og produktkvalitet.

For ældre intelligente tutorsystemer (Cognitive Tutor, MATHia, ASSISTments, Carnegie Learning) ligger den typiske effekt på 0,27 til 0,66 i grundskolen sammenlignet med konventionel undervisning. Det er små til moderate effekter. En ny amerikansk metaanalyse (arXiv 2511.04997, 2025) med 18 studier finder mere beskedne 0,27, altså en lille effekt. Et robust, men mere afdæmpet billede end de ældre tal.

TIL DEN DANSKE KONTEKST

Læg mærke til, at stort set ingen af de største RCT'er er gennemført i indskoling eller mellemtrin (1.–6. klasse). Det betyder, at hvis du står med en kommune, der vil indføre ChatGPT i 4. klasse, så har du faktisk ingen stringent forskning at læne dig op ad. Du kan kun antage, at det, der virker for ældre elever, også virker for de yngre. Det kalder på en større forsigtighed, end debatten lige nu lægger op til.

Effekter på engagement og elever med særlige behov

MOTIVATION OG ENGAGEMENT

En scoping review i *International Journal of Educational Technology in Higher Education* (Springer, 2025) gennemgik 31 studier, der eksplicit brugte motivationsteoretiske rammer. Mønsteret er rimelig klart: chatbots løfter *engagement*, *interesse* og *indre motivation* mest konsistent. Effekten på *vedholdenhed* er mindre og mere afhængig af konteksten.

Det er en interessant nuance, fordi den hænger fint sammen med Bastani-fundet. Eleverne *oplever*, at de er engagerede, når de bruger AI, men det betyder ikke nødvendigvis, at deres faglige robusthed vokser. Engagement og læring kan i nogle situationer trække i hver sin retning.

NEURODIVERGENTE ELEVER (ADHD, ORDBLINDHED, AUTISME, HØJTBEGAVEDE)

En scoping review i *Computers and Education: AI* (2025) gennemgik 21 kilder om generativ AI til neurodivergente elever. Litteraturen koncentrerer sig om tre områder: personaliseret læring, individuelle undervisningsplaner og lærerens administrative byrde. Den empiriske effektevidens er sparsom. Fordelene er teoretiske eller på pilotniveau.

Reviewets vigtigste advarsel handler om over-automatisering af støtten uden specialistens overblik. AI-drevet ordblindeskreening (DIBELS-8 ML-validering, øjentrackings-klassifikatorer i PMC 2024) er et mere modent felt end AI-baseret tutoring for ordblinde. FACET-multiagentstudiet (arXiv 2601.22788, 2026) bekræfter behovet for differentiering, men understreger samtidig, at lærerens pædagogiske autonomi skal bevares.

EN ADVARSEL OM ULIGHEDEN, DER IKKE KAN IGNORERES

RAND's amerikanske lærerpaneldata (Kaufman et al., 2024; Diliberti et al., 2024–25) viser, at AI-brugen koncentrerer sig i ressourcestærke og overvejende hvide skoledistrikter. Kun 39 procent af de distrikter, der primært betjener farvede elever, planlagde lærer-AI-træning i 2023–24, mod 65 procent af de overvejende hvide distrikter. Det er en strukturel risiko: AI kan ende med at udvide kløften, selv når den gennemsnitlige effekt er positiv. For en dansk kommune med en høj andel af elever fra ressourcevage hjem eller med tosproget baggrund er det her en konkret bekymring at have med, ikke en abstrakt overvejelse.

Arbejdsbyrde, pædagogik og kompetenceudvikling

Det her er måske det område, hvor evidensen er mest entydigt positiv. AI som lærerværktøj sparer tid uden at koste på kvaliteten, hvis det implementeres med både redskab og oplæring.

Education Endowment Foundation x NFER (Storbritannien, 2024)

RCT · 259 KS3-NATURFAGSLÆRERE · 68 SECONDARY SCHOOLS

Lærere, der brugte ChatGPT med en ledsagende værktøjskasse fra Bain og Hg Foundation, reducerede planlægningstiden med **31 procent (25,3 minutter om ugen)**. Et uafhængigt panel kunne ikke se forskel på undervisningskvaliteten. Forbehold: kun en lille stikprøve af lektionerne blev kvalitetsvurderet, og lærerne var selvselekterede. En opfølgende EEF-RCT med Aila fra Oak National Academy er i gang i grundskolen 2024–26 og bliver det første store britiske RCT på AI-lektionsplanlægning på de yngste klassetrin.

53%

AMERIKANSKE LÆRERE
BRUGER AI I 2024–25 (OP
FRA 25% ÅRET FØR)

31%

TIDSBESPARELSE PÅ
LEKTIONSPLANLÆGNING
VED CHATGPT-BRUG (EEF,
2024)

61%

FORÆLDRE BEKYMREDE FOR
AI'S EFFEKT PÅ ELEVERS
KRITISKE TÆNKNING (RAND)

RAND-RAPPORTERNE (2023–25)

Lærernes brug voksede fra 25 procent i 2023–24 til 53 procent i 2024–25. Skoleledere ligger på omkring 60 procent. I foråret 2025 rapporterede 54 procent af amerikanske elever, at deres skole bruger AI. Distriktstræningen halter bagefter: 23 procent i 2023, 48 procent i 2024 og forventet 74 procent i 2025. To andre tal at lægge mærke til: 61 procent af forældrene og 22 procent af distriktslederne er bekymrede for, at AI vil skade den kritiske tænkning.

NATURFAGSLÆRERE BRUGER AI TIL FORBEREDELSE, IKKE DIREKTE UNDERVISNING

RAND's analyse af STEM-lærere i grundskolen (arXiv 2509.10747, 2025) viser, at generativ AI overvejende bruges til forberedelse, lektionsplaner, prøveudkast og feedback, snarere end til direkte undervisning i klassen. Lærerne holder fast i deres faglige autoritet over indholdet. Det er et vigtigt fund, fordi det dæmper fortællingen om, at "AI overtager klasseværelset". Det gør den ikke. Den overtager forberedelsen.

TYSK-DANSK LÆRERSAMMENLIGNING

Et komparativt studie i *Journal of Research on Technology in Education* (2025, n=532, heraf 276 danske og 256 tyske lærere) viser, at de tyske lærere har større tiltro til egne AI-evner og en stærkere oplevelse af, at AI er nyttigt til individualiseret læring, end de danske. En latent klasseanalyse identificerede fire lærertyper: "rather optimistic", "rather critical", "critically reflected" og "ambiguous". De fordeler sig forskelligt mellem landene, hvilket peger på, at karakteristika ved det danske system, decentralisering og stærk professionel autonomi, præger måden, AI bliver taget i brug.

SKOLVERKETS TAL FOR GRUNDSKOLEN

Skolverkets rapport "Artificiell intelligens i undervisningen, grundskolan, förskoleklass och fritidshem" (december 2024, n=529 lärare) viser, at omkring 40 procent af lærerne i grundskolan brugte AI-tjenester i 2024. Højest i högstadiet (cirka 50 procent). Kun én ud af seks tillader, at eleverne bruger AI. Den primære anvendelse er lektionsplanlægning. Lærere på de yngste klassetrin bruger markant mindre AI end deres kolleger på de ældste klassetrin. Skolverket peger eksplicit på et efteruddannelsesbehov, lærerne efterspørger selv mere kompetenceudvikling.

Aflastning, afhængighed og kritisk tænkning

Det er her, jeg synes debatten ofte bliver for sort-hvid. Forskningen siger ikke "AI skader kognition". Den siger noget mere præcist og mere brugbart: AI flytter den kognitive belastning, og hvor den lander, afhænger af, hvordan opgaven er bygget.

DEN FÆLLES KERNE I TI FORSKELLIGE STUDIER

Litteraturen om kognitiv aflastning er ved at finde sin form: Gerlich 2025, Lee et al. 2025, Stadler, Bannert og Sailer 2024, et RCT om "ChatGPT som kognitiv krykke" i *Computers in Human Behavior Reports* 2025, og et review i *Computers and Education: AI* 2026. Det konsistente fund er, at AI *reducerer den umiddelbare kognitive belastning*, men på bekostning af *indlæringsdybde, fastholdelse af viden og argumentkvalitet*. Effekten er stærkest under ustruktureret, ubegrænset adgang, og den mildnes, når AI tvinges til at stilladsere snarere end at svare.

BJORKLUND ET AL. (BJET 2024): "PAS PÅ METAKOGNITIV DOVENSKAB"

Et laboratoriestudie med 117 sprogstuderende viste, at ChatGPT-brugerne skrev og rettede hurtigere, men udviste reduceret selvreguleret læringsadfærd i efterfølgende opgaver uden hjælp. Det er præcis det mønster, danske lærere genkender, når de møder elever, der "har glemt at tjekke" eller "ikke ved, hvor det kom fra".

ANTHROPICS INTERNE ANALYSE 2025

Opgaver på højere taksonomiske niveauer (analysere, evaluere, skabe) bliver i uforholdsmæssig grad aflastet til Claude sammenlignet med opgaver på lavere niveauer. Det rejser en pinlig pointe: AI håndterer netop det kognitive arbejde, som skolen er sat i verden for at udvikle.

DEN MODGÅENDE EVIDENS

Avello et al. (2024), Essel et al. (Ghana, 2024) og MDPI *Education Sciences* 2025 viser, at når AI positioneres som *sokratisk samtalepartner* med vejledende prompts, ses gevinster i kritisk tænkning. Polariteten afhænger af det instruktionelle design. Det er ikke en lille detalje, det er hele pointen.



Eleverne oplever ikke selv tabet. Det er det mest urovækkende ved aflastningsmønsteret. De tror, de lærer, mens evnen falder.

SAMMENFATTET PÅ TVÆRS AF BASTANI 2025, KOSMYNA 2025, BJORKLUND 2024

Norden og Vesten

Danmark

Danmark er længst fremme med pædagogisk infrastruktur. Vi er ikke længst fremme med RCT'er, men vi har bygget de mest gennemarbejdede praktiske rammer i Norden.

Børne- og Undervisningsministeriets syv anbefalinger (juni 2025) dækker lokale rammer, GDPR-compliance, pædagogisk formål, kritisk refleksion, prøveintegritet, lærerkompetence og dialog med forældre.

STIL's vejledning om lovlig brug af kunstig intelligens for uddannelsesinstitutionerne (maj 2025, udarbejdet sammen med Kammeradvokaten) og den efterfølgende mini-guide (januar 2026) opstiller en proces i syv trin til efterlevelse af reglerne: kortlægning, kompetenceudvikling, risikovurdering, afklaring af dataansvar, samtykke, tilsyn og dokumentation.

Det danske marked for AI-platforme til skolen består lige nu af tre væsentlige aktører, der hver dækker forskellige behov.

SkoleGPT (Future Classroom Lab sammen med CFU/Københavns Professionshøjskole og DBC Digital) er en open source-chatbot bygget på Mixtral-8x7B, hostet på dansk infrastruktur drevet af grøn energi, uden logning, GDPR-kompatibel, med justerbare parametre og forhåndsfiltre. Den vandt Sprogprisen 2023 og er nomineret til Digitaliseringsprisen 2025. Det er den åbne, offentligt udviklede platform i feltet, og dens særlige styrke ligger i åbenheden: åbne modelvægte, ingen kommerciel binding og en infrastruktur, der reelt kan hostes uafhængigt. Bachelorprojekter fra 2024 (Honoré og Wæhrens) dokumenterer syv pædagogiske potentialer og fem udfordringer i et engelsk-skriveforsøg i 8. klasse.

Miira (Pathfindr ApS, Roskilde) er det kommercielle danske modstykke. Den er bygget op om en eksplicit pædagogisk tutor-funktion, der er trænet til at stilladsere og stille spørgsmål i stedet for at give færdige svar, præcis det designprincip, forskningen (Kestin, Stanford CoPilot, Eedi/LearnLM) peger på som det, der virker. Skolen vælger selv, hvilken underliggende model der skal drive Miira, og leverandøraftalerne kører med 0-retention, så data ikke gemmes hos modeludbyderen. Lærerne har en separat assistent til planlægning og differentiering, og platformen tilpasses skolens fag og bekendtgørelser. Data lagres i EU, persondata bruges aldrig til træning, og GDPR-arbejdet er omfattende: ISAE 3000 Type II-revision, dansk databehandleraftale og færdige DPIA og TIA. Tutoren har en klassevælger til folkeskolen og understøtter ordblinde med tekst-til-tale. Forbeholdet er kundebasen: den er overvejende ungdomsuddannelser og enkelte kommuner (blandt andet Roskilde), så grundskoleerfaringerne er stadig under opbygning.

Teacherspace (stiftet af Jesper Bergmann og Frederik Voss, finansieret af Innovationsfonden) er en lærer-rettet platform, der hjælper med forberedelse, differentiering, forældrekommunikation og

administration. Platformen er bygget specifikt til danske folkeskolelærere, integrerer med Unilogin, Aula og TRIO/Skema og kører på OpenAI-modeller med EU-only setup. Lærere har gratis adgang til grundversionen, og finansieringen af fuldversionen håndteres af skolen eller kommunen. På Kildebjergskolen, der har taget platformen i brug siden maj 2025, rapporteres tidsbesparelser i størrelsesordenen reduceret rettetid fra 30 til 10 minutter per opgave. I modsætning til Miira og SkoleGPT har Teacherspace ikke en elev-rettet tutor-funktion, det er rent et lærerværktøj.

Hverken SkoleGPT, Miira eller Teacherspace har offentligt tilgængelige effektstudier endnu. Tilgængelig dokumentation hviler på pilotforløb, brugerinterviews og selvrapporterede tidsbesparelser.

Folkeskolens prøver (FP9, FP10): Generativ AI er forbudt ved skriftlige og mundtlige prøver under de gældende UVM-regler. Ekspertgruppen under Vedersø anbefalede en prøvereform i april 2024 for at gøre bedømmelsen mere robust over for AI, med tidligst implementering fra foråret 2025.

DANSKE FORSKNINGSMILJØER

Aarhus Universitets Center for Computational Thinking and Design (Iversen, Dindler, Smith med flere) er det førende nordiske forskningsmiljø om AI i grundskolen. CHI 2024-papiret "From Primary Education to Premium Workforce" udvider tilgangene fra grundskolen omkring *computational empowerment* til AI-forståelse mere bredt. Københavns Universitets forskning i digitalt instinkt (Boie, Dalsgaard, Caviglia, BJET 2024) og **Dalsgaard og Prilop (2025) "Partnerskaber mellem elever og AI"** (n=546 gymnasieelever plus 18 interviews) dokumenterer, hvordan elever pendler mellem *kognitivt partnerskab* og *outsourcing*. Det er det mest citerede danske empiriske studie om AI i skolen lige nu.

EVA's "Elevers brug af AI i gymnasiet" (2025, n=1.411 3.g-elever): To tredjedele rapporterer, at de mindst én gang har brugt AI på en måde, de selv kategoriserer som snyd, i individuelle opgaver. Næsten ni ud af ti i gruppearbejde. Bemærk: det er gymnasieniveau, ikke grundskole, men det er det mest stringente danske datasæt om elevernes AI-adfærd. EVA lukker per finanslov 2026 (effektivt 1. juli 2026), hvilket efterlader et betydeligt evalueringshul, lige når der er mest brug for evidens til skolepolitikken.

Sverige

Det strammeste land i Norden, og samtidig det med den laveste udbredelse i grundskolen. Skolverket udsendte tidlige råd i 2023, opdateret i 2024-25, der fraråder højrisikobedømmelse af afleverede opgaver, opfordrer til tydelige retningslinjer og advarer mod uovervåget elevadgang til generativ AI.

AI blev tilføjet som gymnasiefag i efteråret 2024, men *ikke* som obligatorisk indhold i grundskolan. Sveriges nationale AI-strategi (februar 2026, 479 millioner SEK) inkluderer en opgave til Skolverket om at udvikle en plan for AI i undervisningen fra grundskola til universitet. Sverige er det eneste nordiske land med formelle restriktioner på AI i lektier i OECD's styringsrapport fra 2023.

Norge

Utdanningsdirektoratets "Råd om kunstig intelligens i skolen" og medfølgende kompetencepakke giver ni ikke-bindende anbefalinger, der lægger vægt på lokal pædagogisk, juridisk og etisk vurdering, alderssvar, kildekritik og personvern.

"Spørgsmål til skole-Norge 2025" giver de nyeste nationale tal: 65 procent af lederne på barnetrinnet (1.-7. trin) og 89 procent på ungdomstrinnet rapporterer, at eleverne bruger KI. Kun én ud af fire grundskoler har en skriftlig KI-plan. Kunnskapsminister Nordtun signalerede i 2025, at nye forslag om aldersrestriktion er under overvejelse. Oslo kommune har indkøbt 92.000 ChatGPT-licenser, hvilket har skabt offentlig debat.

Finland

Undervisnings- og kulturministeriet udgiver sammen med Finnish National Agency for Education AI-vejledninger for ECEC, grundskole, ungdomstrin, gymnasium, erhvervsuddannelse og voksenuddannelse i 2024-25 på finsk, svensk og engelsk. Finland har den stærkeste integration af mediekompetence og AI-forståelse fra førskolen, bygget oven på landets allerede førende undervisning i at håndtere misinformation.

Generation AI-projektet (University of Eastern Finland sammen med Helsinki og Oulu, et seksårigt nationalt strategisk projekt) har empirisk undersøgt mere end 200 elever i 4. og 7. klasse, der designede ML-apps og udforskede algoritmisk bias. Børnenes forståelse og forklaringer af AI forbedredes målbart gennem co-design i praksis. Det er et af de få stringente studier af AI-forståelse blandt de yngste elever i Norden.

Tyskland og Holland

Tyskland kører decentralt via KMK med et nationalt AI-kursus i seks moduler (Micheuz 2020). Lærerne har, som nævnt, større tiltro til egne AI-evner end de danske.

Holland har NOLAI (Nationaal Onderwijslab AI), grundlagt 2020 ved Radboud Universitet. Det er det offentlige flagskibs-AI-laboratorium for grundskolen og ungdomsuddannelsen og medudvikler værktøjer som "Happy Readers" (talegenkendelse til læsefærdighed) og en AI-Playground til naturfag. Kennisnet udgav en AI-guide i 2024 og et toolkit i 2025 med diskussionskort, eksempler på retningslinjer og skabeloner. SLO har integreret AI i den nationale læseplan for digital dannelse i 2022-24.

Storbritannien

Department for Education har finansieret EEF-forsøgene (omtalt ovenfor) og Oak National Academys Aila-lektionsplanlægger med over to millioner pund i 2024-25.

National Literacy Trusts undersøgelse fra 2024 (n=53.169, heraf 15.830 13–18-årige og 1.228 lærere) viser, at elevernes brug af generativ AI *steg fra 37 procent i 2023 til 77 procent i 2024*. 49 procent af lærerne er bekymrede for negativ effekt på skrivefærdigheder, 75 procent ønsker mere kompetenceudvikling, og 82 procent støtter undervisning i kritisk omgang med AI.

USA

Office of Educational Technology udgav i maj 2023 grundlagsdokumentet "Artificial Intelligence and the Future of Teaching and Learning" med syv anbefalinger centreret om mennesker-i-loopet, lighed og tillid. Mere end 20 stater (Californien, Oregon, Connecticut, North Carolina med flere) har formelle vejledninger om AI i skolen.

Internationalt

OECD Digital Education Outlook 2023 og 2026 (seneste udgave januar 2026) sporer national governance, brugen af generativ AI per skoletrin (i Schweiz 2024-data: cirka 8 procent af elever i indskoling og mellemtrin og 30 procent på ungdomstrin brugte GenAI ugentligt) samt rammerne for skolepolitik. Kun Sverige forbyder formelt generativ AI i lektier. Japan har midlertidige retningslinjer fra 2023.

UNESCO's "Guidance for Generative AI in Education and Research" (Miao og Holmes, 2023) foreslår en aldersgrænse på 13 år for selvstændig brug af generativ AI, obligatorisk databeskyttelsesregulering og institutionel etisk og pædagogisk validering. Mindre end 10 procent af skoler og universiteter havde formelle AI-retningslinjer ved offentliggørelsen.

UNESCO's **Global Education Monitoring Report 2023** "Technology in Education: A Tool on Whose Terms?" har en forsigtig tone. Hovedbudskabet er, at læringsfordelene forsvinder, når teknologien bruges i overskud eller uden en kvalificeret lærer. Kun 16 procent af landene garanterer beskyttelse af pædagogiske data ved lov, og 89 procent af 163 undersøgte EdTech-produkter kunne overvåge børn.

EU's AI-forordning (Forordning 2024/1689, gældende 2024 med fuld anvendelse i etaper frem mod 2026–27) klassificerer som *højrisiko* AI-systemer, der bruges til at (a) bestemme adgang eller optagelse, (b) vurdere læringsudbytte, (c) bedømme uddannelsesniveau eller (d) detektere forbudt prøveadfærd. Det er Bilag III, punkt 3. **Følelsesgenkendelse i uddannelsessystemet er forbudt** (artikel 5, stk. 1, litra f). For grundskolen betyder det, at mange AI-bedømmelsesværktøjer og tilsynsværktøjer, samt adaptive læringssystemer der får reel betydning for elevens videre uddannelse, falder ind under højrisikoforpligtelserne: kvalitetsstyring, risikovurdering, menneskelig kontrol, post-market-monitorering og gennemsigtighed. Generelle chatbots til lektiehjælp gør det typisk ikke, *medmindre* de bruges til bedømmelse, hvorefter de bliver højrisiko.

EU JRC (Joint Research Centre) har udgivet det grundlæggende "Impact of AI on Learning, Teaching and Education" (Tuomi 2018, opdateret), "On the Futures of Technology in Education"

(2023) og driver AI-Watch-portalen.

Council of Europe Compass for AI and Education (2025) er en ikke-bindende værdiramme for medlemslandene.

"AI in European Schools" sammenlignende rapport (Skills Upload JR, 2024–25, n=7.000 elever på tværs af syv lande) viser, at ChatGPT bruges af 48 procent personligt og 47 procent under læreranvisning. Storbritannien fører på adgang, Grækenland halter. Cirka 30 procent af eleverne i nogle lande angiver, at deres skole ikke har fælles regler for AI-brug.

Kompetenceudvikling og infrastruktur

Det mest konsistente fund i alle de store undersøgelser er det her: brugen kører 12–24 måneder foran (a) lærerkompetencer, (b) skolepolitik, (c) GDPR- og AI-Act-arbejdet, (d) integration i læseplanen og (e) redesign af prøverne.

Det her gælder Norden i særdeleshed. Den fælles nordiske survey (Tekstognode/Kantar, juni 2025) viser, at norske og danske elever begynder at bruge AI tidligere og mere udbredt end deres svenske og finske jævnaldrende. Meget få lærere har bygget interne AI-værktøjer.

OECD Digital Education Outlook 2026 bekræfter, at brugen accelererer, mens effektevalueringen halter. Skolverkets 2024-rapport for grundskolan peger eksplicit på et efteruddannelsesbehov.

Det praktiske svar på den her implementeringskløft er ikke at vente. Det er at indsnævre risikoen ved at vælge værktøjer, der er bygget GDPR-sikre fra bunden (SkoleGPT, fremtidige NOLAI-output), kombinere indkøbet med *obligatorisk* kompetenceudvikling, og udsætte de højrisikoanvendelser, der kræver fuld AI-Act-compliance, til den nødvendige infrastruktur er på plads.

Hvad vi ved virker for de 6–15-årige

VÆRKTØJSTYPE	BEDSTE EVIDENS	EFFEKT	FORBEHOLD
Ældre intelligente tutorsystemer (MATHia, ASSISTments, Cognitive Tutor)	RAND RCT 2010, Carnegie Learning 2018+, Kulik & Fletcher 2016 metaanalyse	ES 0,27–0,66 i grundskolens matematik og læsning	Begrænset i nordisk kontekst, primært fra 6. klasse og opefter
Adaptive matematikplatforme (DreamBox, IXL, i-Ready, Prodigy, ALEKS)	NCTM/RAND blandet evidens; ACM 2024 hybridstudie menneske-AI (3 sites)	Beskedne, dosisafhængige	Effekterne forsvinder ofte ved lav brugstid
Pædagogisk finjusterede LLM-tutorer (LearnLM, Khanmigo, Tutor CoPilot)	Eedi/LearnLM 2025, Stanford 2024, Khan Academy 2024	Små til moderate +; mere sikre end rå chatbots	Ny evidens, behov for replikation uden for forsøgsbetingelser
Generiske chatbots (ChatGPT, Gemini, Claude)	Bastani 2025, Kosmyrna 2025; metaanalyser	Blandet til skadelig uden styring	Stor risiko for kognitiv aflastning og alvorlig GDPR-risiko for elever under 13 år
Skriveassistenter (Grammarly, Brisk, MagicSchool)	Steiss et al., <i>Learning and Instruction</i> , juni 2024 (200 historieessays, 6.–12. klasse)	ChatGPT-feedback bedømt 3,6/5 mod menneske 4,0/5	Lærerfeedbackens fortrin ligger især i udviklingsmæssigt passende næste skridt
Danske pædagogiske platforme (SkoleGPT, Miira, Teacherspace)	Kun pilotstudier og kvalitative studier	GDPR-overholdelse, pædagogisk sikkerhed, men mangler effektstudier	SkoleGPT (åben, gratis, eksperimentel), Miira (kommerciel, lærer + elev-tutor), Teacherspace (kommerciel, kun lærerværktøj). Ingen målte data på læringsudbytte endnu.

Kompakt referenceliste

RCT'ER OG KVASI-EKSPERIMENTER (TÆTTEST PÅ GRUNDSKOLE-ALDERSGRUPPEN)

Bastani et al., PNAS 2025. 9.–11. klasse, tyrkisk matematik, ~1.000 elever.
 Wang/Stanford et al., arXiv 2410.03017. Tutor CoPilot, K-12 USA, n=1.800.
 Eedi/LearnLM, arXiv 2512.23633 (2025). UK secondary, 13–15 år, n=165.
 De Simone et al., Verdensbanken. Edo State Nigeria, 12–17 år, n=800, 6 uger.
 EEF/NFER ChatGPT lektionsplanlægning 2024. UK KS3 naturfag, 259 lærere.
 Kestin et al., *Scientific Reports* 2025. Harvard PS2, n=194 (universitet).
 Carnegie Learning blended approach (RAND). USA mellem- og udskoling, 18.000 elever.
 Dalsgaard og Prilop 2025. Dansk gymnasium, n=546 plus 18 interviews.

METAANALYSER OG SYSTEMATISKE REVIEWS

Kulik og Fletcher, *Review of Educational Research*, 2016. ITS, ES 0,66.
 Steenbergen-Hu og Cooper, 2014; Ma et al., 2014; VanLehn 2011.
 Xu et al., BJET 2019. ITS læseforståelse, ES 0,60.
 Zhang og Tur, *European Journal of Education* 2024. ChatGPT i K-12 (PRISMA, 13 papers).
 Liu et al., *Journal of Computer Assisted Learning* 2025. ChatGPT-metaanalyse.
 Wang og Fan 2025; "ChatGPT student performance" *HSSC* 2025/26. 35 og 51 studier.
 "Meta-analysis on AI agents on K-12 cognitive performance," *Computers and Education: AI* 2026.
 Scoping review om generativ AI for neurodivergente K-12-elever, *C&E: AI* 2025 (21 kilder).
 USA-only K-12 ITS metaanalyse, arXiv 2511.04997 (2025). $g = 0,27$.

KOGNITIV PÅVIRKNING OG AFLASTNING

Kosmyna et al., MIT Media Lab arXiv 2506.08872 (juni 2025).
 Stadler, Bannert og Sailer 2024 (tysk eksperiment).
 Gerlich 2025; Lee et al. 2025; Niloy et al. 2023.
 "Beware of metacognitive laziness," BJET december 2024.
 "ChatGPT as a cognitive crutch," *Computers in Human Behavior Reports* 2025.

SKRIVE-FEEDBACK

Steiss et al., *Learning and Instruction* 2024. Sammenligning af menneske- og ChatGPT-feedback på 6.–12. klasse historie-essays.
 AI-Mediated Feedback (FeedbackWriter), arXiv 2602.16820, 2026. Stort RCT på universitet.

POLICY-RAPPORTER (MEST RELEVANTE FOR EN DANSK KONSULENT)

Børne- og Undervisningsministeriet "Syv anbefalinger" (juni 2025).
 STIL "Vejledning om lovlig brug af kunstig intelligens" (maj 2025) og mini-guide (januar 2026).

Skolverket "Artificiell intelligens i undervisningen, grundskolan" (december 2024).
Skolverket "Råd om AI, chattbottar och liknande verktyg" (2024–25 opdateret).
Utdanningsdirektoratet "Råd om kunstig intelligens i skolen" plus kompetencepakke (2024–25).
Finnish National Agency for Education AI-vejledninger (2024–25, fasevis udgivelse).
Kennisnet AI Toolkit og Guide (2024, 2025).
OECD Digital Education Outlook 2023 og 2026.
UNESCO GenAI Guidance (2023) og GEM Report 2023.
US OET "Artificial Intelligence and the Future of Teaching and Learning" (maj 2023).
EU JRC-rapporter samt EU AI-forordningen 2024/1689 (Bilag III, betragtning 56).
Council of Europe Compass (2025).
European Schools internt rammeværk (2025-01-D-65/66).

LÆRER- OG ELEVSURVEYS

RAND ASDP/AEP/ASLP-serien (Diliberti, Kaufman, Woo et al., 2024–26).
National Literacy Trust UK 2024 (n=53.169).
HEPI Student GenAI Surveys 2024, 2025, 2026 (UK undergraduates).
AI in European Schools (EU/Skills Upload JR, 7 lande, n=7.000).
Nordic joint survey on GenAI/Education/Copyright (Kantar, 2025).
Spørringen "Spørsmål til skole-Norge 2025."
EVA "Elevers brug af AI i gymnasiet" 2025 (n=1.411).
Tandfonline 2025 Tysk-dansk lærer-sammenligning (n=532).

Beslutningsklare, fordelt på alderstrin

Det her er ikke universalmedicin. Det er det, jeg ville lægge til grund, hvis jeg sad i en dansk skoleledelse eller en kommune i 2026 og skulle træffe valg ud fra den eksisterende evidens.

1.–4. klasse (cirka 6–10 år)

Som udgangspunkt: kun lærerformidlet AI. Giv ikke eleverne selvstændig adgang til generativ AI. Brug AI til *lærerens* opgaver alene: lektionsforberedelse, differentiering, læsestilladsering og udkast til individuelle handleplaner, og kun under forudsætning af, at GDPR og STIL's vejledning i syv trin er overholdt.

Begrundelsen er tredelt: udviklingspsykologiske bekymringer (UNESCO's 13-årsgrænse, Nordtuns advarsel om grundlæggende færdigheder), ingen reel RCT-evidens for AI-tutoring i indskoling, og højeste risiko for kognitiv aflastning, mens de grundlæggende færdigheder er under dannelse.

Tærskel for revurdering: hvis der dukker store RCT'er op for de yngste klassetrin. Hold især øje med UK EEF's Aila-forsøg på de yngste klassetrin i 2024–26.

5.–6. klasse

Indfør superviseret, stilladseret arbejde med AI-forståelse. Brug et dansk værktøj med pædagogisk stilladsering indbygget. SkoleGPT (åbent og gratis) eller Miira (kommercielt og produktmodent) er de to oplagte valg. Begge er bygget med GDPR i bunden. SkoleGPT er stærkest, hvis du vægter åbenhed og fuld kontrol over infrastrukturen, Miira er stærkest, hvis du vægter en produktmoden tutor med fleksibelt modelvalg og færdig DPIA/TIA. Fremstil AI som *kognitiv samarbejdspartner*, ikke som svarmaskine, efter designprincipperne fra Kestin/Stanford/Eedi: korte svar, ingen komplette løsninger, sokratiske prompts, og lærersynlighed på chatlogs. Byg på Generation AI (Finland) og Aarhus CCT-pædagogiske materialer.

7.–9. klasse

Integrér AI på en måde, der bevarer den kognitive anstrengelse hos eleven. Indfør Bastani et al.'s GPT-Tutor-mønster: eleverne må bruge AI under øvelser, men bedømmes *altid* uden. Lærerne diskuterer eksplicit aflastningsmekanismen med eleverne, og opgavedesignet antager, at AI er tilgængelig, så vægten skifter mod mundtlige forsvare i klassen, procesporteføljer og superviseret skrivning som ved FP9.

Skolverkets anbefaling om at undgå ikke-superviserede hjemmeopgaver er forsvarlig her. Undervis eksplicit i AI-forståelse: algoritmisk bias, hallucination, prompt-design og kildekritik. Det er på linje med BUVM's nye anbefalinger, Udirs vejledning og den finske læseplan.

For skoleledelse og kommune

- Kør STIL's risikovurdering i syv trin for hvert AI-værktøj *før* brug i klasselokalet. Dokumentér dataansvar.
- Behandl ethvert AI-værktøj brugt til bedømmelse som *højrisiko* efter EU AI-forordningens Bilag III. Implementér kvalitetsstyringssystem, post-market-monitorering og menneskelig override nu, fristen for fuld compliance lukker 2026–27.
- Gør lærernes AI-forståelse til obligatorisk efteruddannelse. RAND-data viser, at træning er den enkeltstærkeste prædiktør for effektiv brug.
- Budgettér til kompetencetid, ikke kun til licenser. EEF og RAND-evidensen viser konsistent, at *værktøj plus kompetenceudvikling* betyder langt mere end værktøjsvalg alene.
- Undgå værktøjer til følelsesgenkendelse fuldstændigt. De er forbudt under EU AI-forordningen.
- Foretræk offentlige open source-alternativer, der er bygget GDPR-sikre fra bunden, som SkoleGPT eller fremtidige NOLAI-output, frem for kommercielle chatbots, hvor det er muligt. Det beskytter dig juridisk og pædagogisk.

For forskere og konsulenter

Det nordiske evidensgrundlag for de yngste klassetrin er stort set tomt på RCT-niveau. Det mest værdifulde, du kunne planlægge eller finansiere, er et RCT i 5.–9. klasse, der sammenligner stilladseret AI med en kontrolgruppe, modelleret efter Bastanis design med tre forsøgsarme, men i en dansk eller nordisk kontekst og med udfaldsmål justeret efter FP9. Det er det største enkeltstående evidenshul i den vestlige forskningslitteratur om AI i grundskolen lige nu.

Hvor evidensgrundlaget knirker

Jeg er nødt til at være ærlig om grænserne for det her. En del af den retorik, du møder fra både entusiaster og skeptikere, hviler på et tyndere fundament, end de selv er klar over.

1. **Evidensgrundlaget for grundskolen er ærlig talt tyndt.** De stærkeste studier (Kestin, Bastani, Stanford CoPilot, Eedi, Nigeria) klumper sig om alderen 13–17 eller universitetsniveau. Næsten ingen stringente RCT'er findes for 1.–6. klasse. At antage at det, der virker for udskoling og videre, også gælder for indskolingen, er fagligt risikabelt. Udviklingsmæssige, kognitive og selvreguleringsmæssige forskelle er præcis det, der bør gøre os forsigtige.
2. **De fleste positive studier er kortvarige** (otte uger eller derunder). Langtidsfastholdelse, transfer og varig færdighedsdannelse er stort set utestet. Bastani-studiet er det vigtigste enkeltmodestykke, fordi det tester, hvad der sker, *når AI fjernes*.
3. **Metaanalysernes effektstørrelser er oppustede** af publikationsbias (positive resultater bliver lettere udgivet end negative) og korte efter-tests. De moderate til store effekter på 0,46 til 0,87, der ofte refereres for ChatGPT, er ikke direkte sammenlignelige med de mere konservative 0,27 fra den stringente amerikanske grundskole-metaanalyse fra 2025.
4. **Værktøjsversionerne bevæger sig hurtigere end evidensen.** De fleste publicerede RCT'er bruger GPT-3.5 eller GPT-4 (2023–24). Aktuelle frontier-modeller opfører sig anderledes. Replikér før du implementerer.
5. **Selvrapporterede snydedata** (EVA, Skolverket, NLT) bruger elevernes egne definitioner af snyd, hvilket både oppuster tallene og afslører normforvirring. Den rette tolkning er, at *eleverne selv* er forvirrede, ikke at to tredjedele af eleverne er forsætligt uærlige.
6. **Flere ofte citerede fund er endnu ikke peer-reviewed.** Kosmyna et al. (MIT) og Eedi/LearnLM-RCT'et er arXiv-preprints. Bastani et al. er nu peer-reviewed i PNAS, men det er stadig SSRN-preprintet, der citeres mest af pressen. Behandl tallene som foreløbige.
7. **Den danske evalueringsinfrastruktur svækkes.** EVA's planlagte lukning (1. juli 2026) fjerner den mest autoritative uafhængige danske uddannelsesevaluator, netop når der er mest brug for evidens til skolepolitikken. Det er en betydelig institutionel risiko at have med i regnestykket.
8. **"AI" er ikke én teknologi.** Ældre intelligente tutorsystemer (MATHia, Cognitive Tutor) har et 30-årigt evidensgrundlag og er på mange måder bedre valideret end de generative chatbots. Sammenblanding af kategorierne er en almindelig fejl, når metaanalyser læses overfladisk.
9. **Advarsel om ulighed.** RAND, Kaufman et al. og den europæiske survey i syv lande viser alle, at udbredelsen af AI koncentrerer sig i ressourcestærke skoler og hjem. Uden en aktiv lighedspolitik kan AI udvide kløften, selv når den gennemsnitlige effekt er positiv. En strukturel bekymring særligt relevant for kommuner med en høj andel af elever fra ressourcetsvage hjem.

10. Den nordiske evidens er domineret af surveys og kvalitative studier, ikke RCT'er. Brug Skolverket, EVA, Udir og Tekstognode/Kantar-Nordic-surveyen til *adoption og holdninger*. Brug internationale RCT'er til *effekter*. Og behandl broen mellem dem med eksplicit ydmyghed.

• • •

Slut på rapporten