

KI og eksamen – løser vi riktig problem?

Stefan Schaubert

Seksjon for helsevitenskapelig
pedagogikk (HELP)
Det medisinske fakultetet

Gunnar Rye Bergersen

Software Engineering
Institutt for Informatikk
Det matematisk-naturvitenskapelige fakultet

«ATLAS»



Utvikling med PhD
og MSc studenter



FormativGPT



Jørgen Afseth Sugar
Universitetet i Oslo



Kåre-Olav Stensløkken
Universitetet i Oslo



Stefan Schauber
Universitetet i Oslo



Sjannie Lefevre
Universitetet i Oslo



Ergi Xhemajli
Universitetet i Oslo



Anh Le
Universitetet i Oslo

Finansiering

Prosjektet er finansiert av Universitetet i Oslo (IT i utdanning 2025)

	Før	Under	Etter
Studenter			
Undervisere			

EKSAMEN

	Før	Under	Etter
Studenter	FormativGPT, <i>EMBLA</i>	---	ATLAS: Individualisert tilbakemelding
Undervisere	«MCQ + GPT = SANT»	Automatisk retting av VSAQs	ATLAS: Psykometriske analyser

	Før	Under	Etter
Studenter	FormativGPT, EMBLA Formativ tilbakemelding på fritekst eksamensoppg.	--- (KI svar på oppgaver som student deretter må vurdere)	ATLAS: Individualisert tilbakemelding Frivillig KI tilbakemelding
Undervisere	«MCQ + GPT = SANT» Evaluering av forslag til oppgaver fra kursmateriell (RAG)	Automatisk retting av VSAQs UiO GPT som sensor nr. 2 evt. 3 i store kurs → Bør man evaluere prosess eller produkt?	ATLAS: Psykometriske analyser Synteser av svar på fritekst oppgaver

Malthe-Sørenssen-utvalget

- «De fremste fordelene ligger i å kunne få **individuelle tilbakemeldinger** og **tilpassede forklaringer** i en helt ny skala» (s. 2)
- «KI-systemer basert på store språkmodeller har et potensial for å gi pålitelig og effektiv gjennomføring av både **sensur** og **karakterbegrunnelse**» (s. 3)
- «**Sensur har generelt større betydning** for studenter enn karakterbegrunnelse, og krever dermed større varsomhet [...] Utvalget anbefaler universiteter og høyskoler å gjøre **forsøk med KI-systemer til sensur og karakterbegrunnelse**, både som automatisering og som vurderingsstøtte.» (s. 3)
- «Dagens system med menneskelig sensur er heller ikke fullt ut pålitelig (reliabel) fordi **ulike sensorer kan legge ulike vurderinger til grunn**» (s. 32)
- «Riktig **kompetanse hos ansatte** vil bidra til bedre løsninger i møte med hele spekteret av problemstillinger knyttet til læring, vurdering og utforming av studietilbud» (s. 45)

MEN...

**Vi når ikke
studentene
som trenger
det mest.**

«[...] bra for videre læring for oss studenter» — «[...] viser [...] hva kan jeg forbedre meg på.» — «Ganske fornøyd, gir oversikt over ting man må jobbe videre med.» — «Visuell fremstilling er veldig fint.»

**«Det var
vanskelig å
forstå den»**

**Det spennende er ikke de
som ble med i et KI-
undervisnings-opplegg.**

**Utfordringer ligger i å nå
de som aldri svarte på
invitasjonen.**

**Hvordan kan vi nå de
studentene som
trenger det?**

**Og hvilken rolle kan KI
spille her?**