



IDEA

Workshop



Bruk av KI i
undervisning
Eksamensdesign





IDEA

Workshop



Bruk av KI i undervisning
Eksamensdesign

Prosjektgruppe KI i utdanning

Andreas Lund

Anders Mørch

Ane Røvik Wahlen

Clay Gouin

Greta Bjørk Gudmundsdottir

Haris Temistokleus

Henrik Tjønn

Kennedy Bailey

Louie Lei Wang

Maja Nordtug

Maryum Challoiui (*student*)

Morten Andreas Høglend (*student*)

Ola Jørgen Swang (*student*)

Torgeir Christiansen

Ulrikke Rindal

Workshopserie om KI i utdanning våren 2024



IDEA



Workshop

Bruk av KI i undervisning Eksamensdesign



Tid og sted: 14. feb. 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Studentoppgaver og arbeidsformer

Hvordan legge til rette for at studenter bruker Kunstig intelligens (KI) kritisk til læring og samhandling? Hvilke pedagogiske hensyn må vi ta i møte med KI?

Bli med på IDEA-workshop, og lær sammen med andre undervisere hvordan KI kan tas i bruk for å støtte studenters læring med KI i ulike emner.



Tid og sted: 21. mars 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Vurderingsformer

IDEA sin workshop-serien om bruk av kunstig intelligens i undervisning fortsetter med en workshop på temaet vurderingsformer.

Bli med på IDEA-workshop, og lær sammen med andre undervisere hvordan KI kan tas i bruk for å støtte studenters læring med KI i ulike emner.



Tid og sted: 25. apr. 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Eksamensdesign

IDEA workshop-serien om bruk av kunstig intelligens i undervisning fortsetter med en workshop om tema eksamensdesign.

Bli med på IDEA-workshop, og tenk sammen med andre undervisere om fornyende eksamensformer som tar hensyn til nye utviklinger om generativ KI.

Workshopserie om KI i utdanning våren 2024

IDEA

Workshop

Bruk av KI i undervisning Eksamensdesign



Tid og sted: 14. feb. 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Studentoppgaver og arbeidsformer

Hvordan legge til rette for at studenter bruker Kunstig intelligens (KI) kritisk til læring og samhandling? Hvilke pedagogiske hensyn må vi ta i møte med KI?

Bli med på IDEA-workshop, og lær sammen med andre undervisere hvordan KI kan tas i bruk for å støtte studenters læring med KI i ulike emner.



Tid og sted: 21. mars 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Vurderingsformer

IDEA sin workshop-serien om bruk av kunstig intelligens i undervisning fortsetter med en workshop på temaet vurderingsformer.

Bli med på IDEA-workshop, og lær sammen med andre undervisere hvordan KI kan tas i bruk for å støtte studenters læring med KI i ulike emner.



Tid og sted: 25. apr. 2024 14:15 – 15:45, [HES 241](#)

IDEA Workshop - Bruk av KI i undervisning: Eksamensdesign

IDEA workshop-serien om bruk av kunstig intelligens i undervisning fortsetter med en workshop om tema eksamensdesign.

Bli med på IDEA-workshop, og tenk sammen med andre undervisere om fornyende eksamensformer som tar hensyn til nye utviklinger om generativ KI.

Takk og farvel, masteroppgaven

Det er på tide å ta farvel med masteroppgaven slik vi kjenner den. Det har vært en lang og lærerik reise, men nå er det på tide å omfavne fremtiden.

Statsråden om kunstig intelligens: – Dei største utfordringane er knytte til vurdering og eksamen

– Kunstig intelligens-teknologiar byr på store moglegheiter for samfunnet, men også nokre openberre trugsmål. Frå samfunnet si side krev dette ei fast og trygg regulering av teknologiar, seier forskings- og høgare utdanningsminister Oddmund Hoel.

Av Martin Toft

Publisert 17. apr. 2024 16:55



Felles retningslinjer for fusk kommer i juni

Del på   

UK ditches exam results generated by biased algorithm after student protests

Protesters chanted 'Fuck the algorithm' outside the country's Department for Education

By Jon Porter | @JonPorty | Aug 17, 2020, 12:16pm EDT

f   SHARE



How a computer algorithm caused a grading crisis in British schools
PUBLISHED FRI, AUG 21 2020 7:18 AM EDT | UPDATED FRI, AUG 21 2020 8:45 AM EDT
sam shead @SAM_L_SHEAD



Noen forskjeller mellom statisk og dynamisk vurdering

Statisk vurdering

- **Ingen/Minimal hjelp**
- Undervisning og vurdering holdes atskilt
- Undervisning mot eksamen
- “Test-wise learners”
- Vekt på testens begrepsvaliditet og reliabilitet
- **Perspektiv: fortid + nåtid**
- **Vurdering som forsinket respons**
- Individuelt fokus, få og kontrollerte variabler
- Bruk av tilgjengelige kulturelle verktøy ofte oversett eller misforstått

Dynamisk vurdering

- **Hvordan vi responderer på hjelp**
- Undervisning, læring og vurdering utgjør en enhet
- Undervisning forbi eksamen
- “World-wise learners”
- Vekt på testens økologiske validitet og studentenes utvikling
- **Perspektiv: fremtid, potensial**
- **Vurdering som intervensjon**
- Samarbeidsfokus, mange variabler, (“økologi”)
- Bruk av tilgjengelige kulturelle verktøy som uttrykk for relevant (‘performativ’) kompetanse

Kan kobles både til diagnostisering, og ‘educational data mining’: uendelig med informasjonsbiter kobles mot definerte kompetanser, og over tid; clickstreams, logdata...



IDEA



Workshop

Bruk av KI i undervisning

Eksamensdesign

Mål

- Drøfte muligheter og utfordringer ved generativ kunstig intelligens (GenKI) i eksamensdesign.
- Arbeide med hensiktsmessig formulering av oppgaver der GenKI inngår i eksamensdesign. Hva skjer med kompetansebegrepet?
- Utforske GenKI i eksamensdesign innenfor ditt spesifikke studieområde.
- Identifisere konsekvenser for læringsutbyttebeskrivelser, informasjon til sensorer, rettleidninger til studenter.



IDEA



Workshop

Bruk av KI i undervisning
Eksamensdesign

Program

- 13:45** – Velkommen
- 13:50** – Introduksjon til GenKI
- 14:00** – Mini-lecture:
- 14:15** – Tilrettelagt diskusjon
- 14:30** – Kaffepause
- 14:35** – Gruppearbeid
- 15:25** – Studentperspektiv Q&A
- 15:40** – Tilbakemelding og sluttrefleksjon



IDEA

Workshop



**Anmeldelse:
Introduksjon til GenKI**

Morten Andreas Høglend,
IDEA StudentTeam





GenKI - Innhold

Hva er det?

Hva kan det brukes til?

Ulike modeller

Styrker og svakheter

- **MÅL:** Kunne viktig KI-termer, vite hva det kan brukes til og gi noen tanker om styrker og svakheter



Hvordan fungerer GenKI?

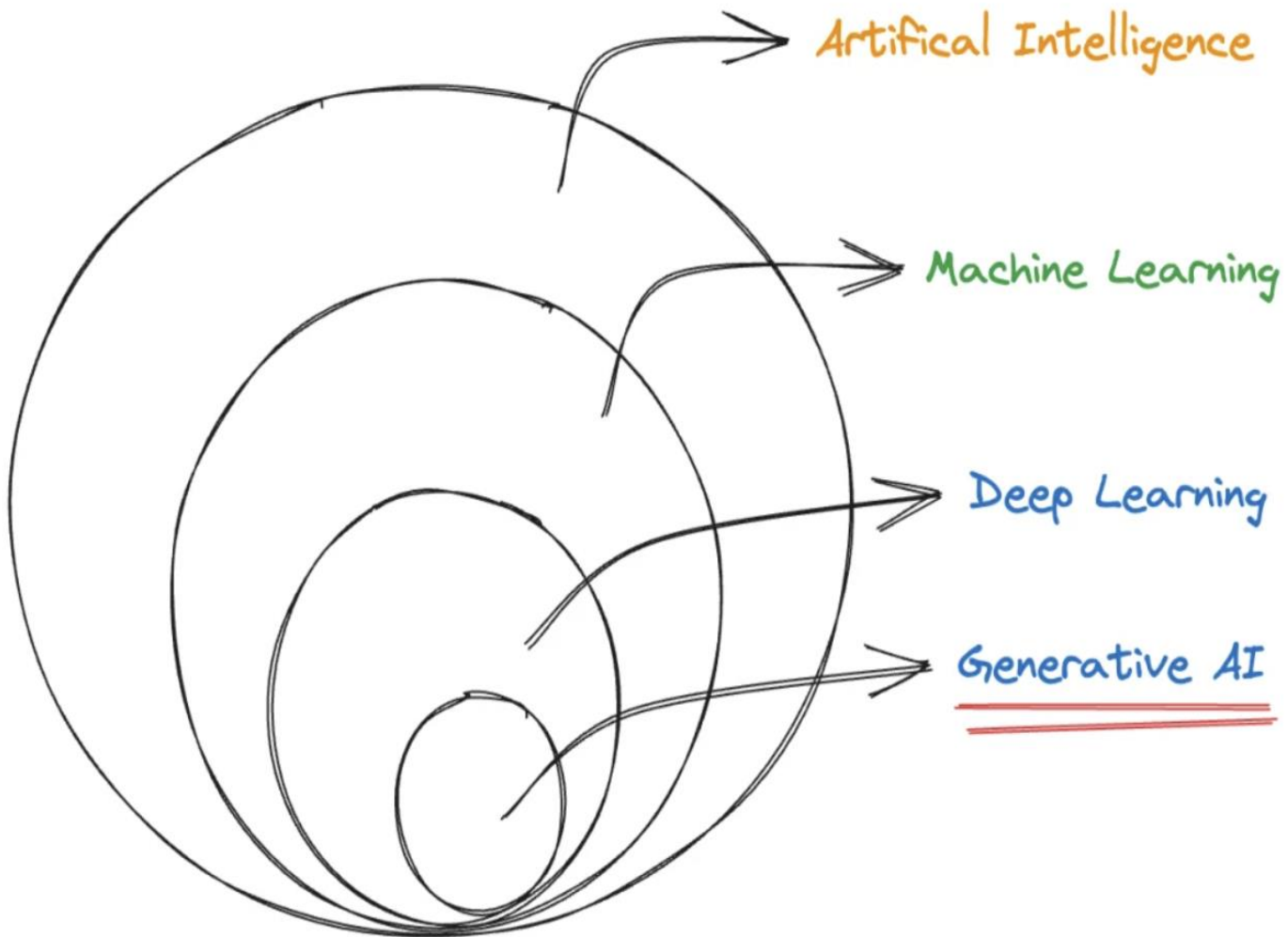
- **GenKI** kan generere: **Tekst**, bilder og lyd.
- **Teknologisk rammeverk for «Chatboter»**
- **LLM** – Large Language model
- **Maskinlæring** □ Transformer algoritmer og nevralt nettverk.
- **Læringsbaserte modeller VS regelstyrte modeller**
- **Hallusinasjoner**
 - **Henvendelse (Inndata/prompt) □ Prediksjon/analyse □ Svar (output).**

IDEA

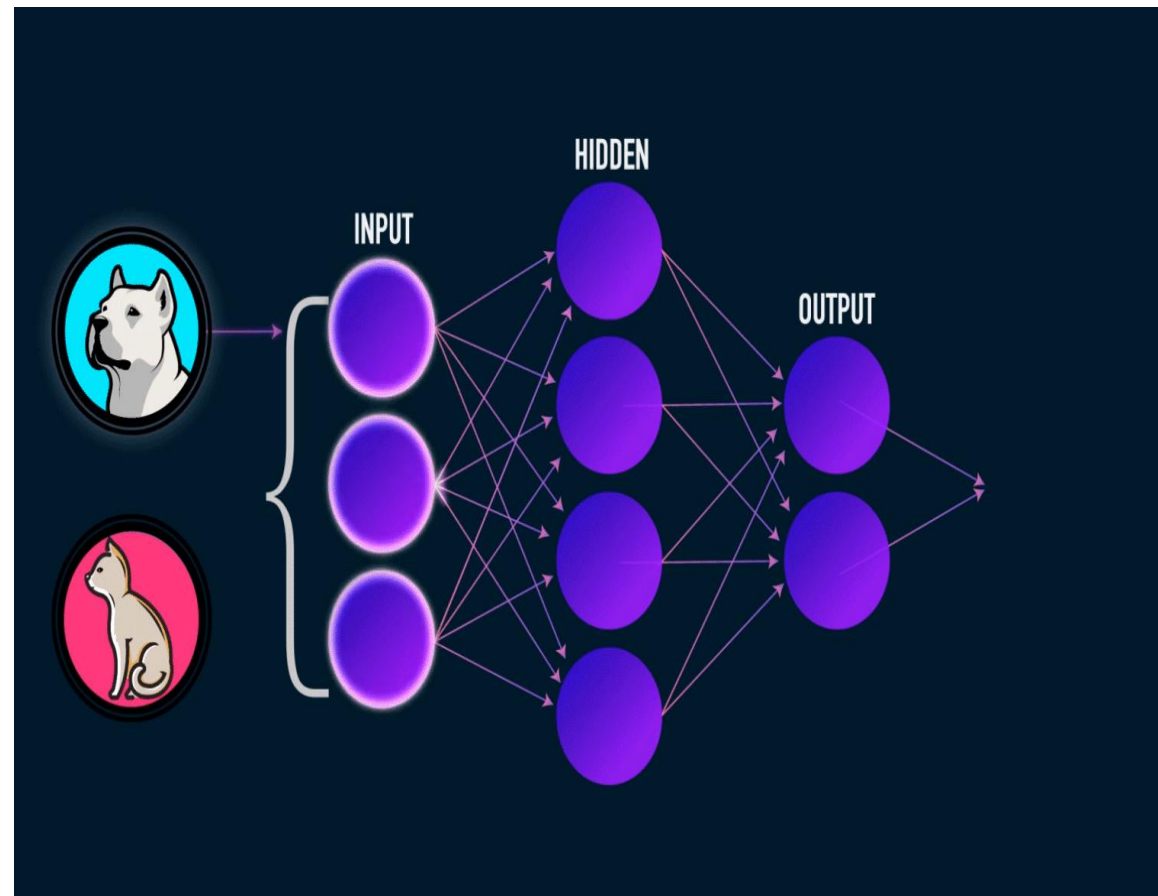
(Online version: presentation notes)

- KI er et felt innenfor datavitenskap, hvor intensjon er å få maskiner til å kunne ha/erme etter intellektuell adferd.
- GenAI er en samlebetegnelse for KI modeller som kan generere naturlig språk i ulike modaliteter. F.eks: Tekst, lyd, bilde. I denne konteksten skal vi fokusere mest på Tekst.
- LLM – er en samlebetegnelse for modeller som bruker språk som semantisk grunnlag for å forstå og generere naturlig språk. --> Disse modellene trenes på enorme mengder tekstdata for å lære språkets struktur og nyanser. Resultatet er en AI som kan utføre en rekke språkrelaterte oppgaver, som å svare på spørsmål, skrive tekst, oversette språk, og mer.
- Maskinlæring er et felt innen kunstig intelligens som fokuserer på å utvikle algoritmer og modeller som lar datamaskiner lære fra og gjøre prediksjoner eller beslutninger basert på data. Dette gjøres uten å være eksplisitt programmert til hver enkelt oppgave. I stedet oppdager og tilpasser modellene seg mønstre i dataene de behandler.
- Transformers – T'en i GPT er et veldig viktig gjennombrudd som skjedde i 2017. Den bruker noe som heter "self attention" til å vekte betydningen av forskjellige deler av inputdataene. F.eks Har man setningen "Jeg var i en ulykke i går" er det Jeg, ulykke og I GÅR som blir vektlagt. Da må man analysere mye mindre, og resultatet blir at man prosesserer raskere og vektlegger de viktigste ordene i en prompt eller setning. Den er kort forklart veldig god til å forstå konteksten i skriftlige (og billedlige) data
- Læringsbaserte modeller

Oppbyggingen av KI



Nevrale nettverk

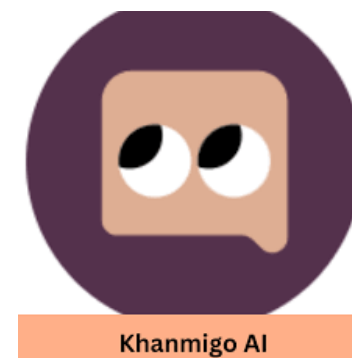


IDEA

(Online version: presentation notes)

- AI teknologi har vært tilgjengelig i mange tiår.
- Venndiagrammet i bildet illustrerer at kunstig intelligens er det bredeste feltet, som omfatter alle teknologier som lar maskiner utføre oppgaver som krever menneskelig intelligens. Innenfor dette feltet finner vi maskinlæring, som er teknikker hvor datamaskiner lærer fra data. Enda mer spesifikt innenfor maskinlæring har vi dyp læring, som involverer nevralt nettverk med mange lag for å tolke kompleks data. Til sist, inni dyp læring, finner vi generativ kunstig intelligens, som er en type AI som kan lage nytt innhold, som bilder eller tekst, som kan virke troverdig for mennesker.
- Nevrale nettverk er selve arkitekturen til modellene – og består av “inndata” – mange lag av skjulte “noder” som blir grunnlaget for prediksjon – altså svaret man får ut. Disse nevralt nettverkene består av utrolig mange lag. En samling av noder fungerer som parametere som blir skrudd på under treningsfasen til modellen. Nettverket består av enorme mengder av data. Flere petabytes av data. For å sette dette i perspektiv: 1 GB tilsvarer ca 178 millioner skrevne ord. 1 PB tilsvarer 1 million GB. Altså 178 millioner ord ganget med 1 million skrevne ord.

GPT UiO



IDEA

(Online version: presentation notes)

- Det finnes et mylder av ulike Chatbots og KI-teknologier med ulike styrker og svakheter. Det blir for snevert å tenke at studenter og forelesere kun bruker ChatGPT og UiO GPT. Mange av de følgende programmene er både spesialiserte og bak betalingsmurer.
- ChatGPT – Originale chatboten som tok verden med storm. Den besto Turing-testen. Retorisk overbevisende tekstproduksjon.
- Claude – Beste chatboten for øyeblikket, basert på Benchmark test.
- Grok – Chatboten posisjonerer seg som «en sannhetsbot» – uten bias, kalde harde fakta... Har mindre sikkerhetsrestriksjoner og forbudte svar enn andre modeller.
- Ernie – 200 millioner unike brukere, beste chatboten i Kina. Er ryktet til å ha tregangeren av parameterne til ChatGPT. Altså hvor mange ventiler man kan skru på for å optimalisere språkmodellen.
- Perplexity og Copilot – Gode chatbotter som inkluderer kilder.
- CustomGPT – Kontekstualiserte og spesialiserte KI programvare som dekker alt fra investeringsbotter til akademiske skrivehjelpere.
- Grammarly – spesialisert på grammatikk og syntaks (brukes som retteprogram av skrevet tekst for mange studenter)
- Khanmigo AI – Mest utviklede adaptive læringshjelperen – basert på Khanacademy – best i STEM-fagene.
- Her har vi ikke inkludert ulike AI teknologi som f.eks går på det å forstå fagartikler og gjøre litteratursøk F.eks Scispace, consensus eller Keenious.

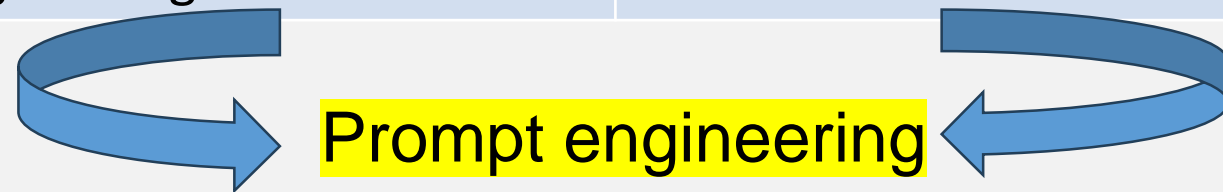


Hva kan det brukes til?

LLM'S brukes allerede til å:

- Automatisere
- Effektivisere
- Avlaste

Styrker	Svakheter
Kreativitet - sparring	Biased
Personalisering – tilpassede forklaringer	The Alignment problem
Effektivitet (kan være negativt også?)	Hallusinasjoner
Data undersøkelse	Sikkerhet
Renskriving, grammatikk og ordvalg.	Kostnad



IDEA

(Online version: presentation notes)

- Hva kan det brukes til?
 - Chatboter kan idag brukes til mange ting. Det er vanskelig å si helt sikkert hva styrkene og svakhetene er, men her er et forsøk på å si noe om funksjonaliteten til de ulike LLM'ene.
- **Styrker:**
 - -Kreativ sparringspartner i kreativt arbeid, spesielt basert på form, layout og oppbygging.
 - -Kan fungere som en samtalepartner for å forklare fenomener, ideer, begreper på et persontilpasset nivå.
 - -Kan avlaste tidkrevende oppgaver som f.eks transkripsjon, litteraturgjennomgang, research arbeid eller skriftproduksjon.
 - - Kan oppsummere og sammenfatte lengre tekster og skriftlige og (biledlige) dokumenter.
 - -Kan hjelpe godt med renskriving tilpasset ulike formal.
- **Svakheter:**
 - -Modellene er trent på historisk data og Big tech selskapenes etiske kompass. Dette kan føre til biaser.
 - -The Alignment problem handler om at vi vet lite om hvilke mønstre KI modellene faktisk har brukt. Dette gjør at man ikke kan være sikre på hvordan de resonnerer og hvorvidt de oppfører seg i tråd med menneskelige verdier og intensjoner. Spesielt når KI-modellene blir mer komplekse og autonome.
- **Det grunnleggende spørsmålet er hvordan man kan kode menneskelige etiske retningslinjer inn i et system som behandler informasjon og tar beslutninger på måter som kan være uforståelige for oss.**
 - -Hallunisasjoner – kan være direkte feil eller retorisk overbevisende tull. Dette er spesielt farlig hvis man er lite kritisk eller har lite domenekunnskap om feltet man prompter rundt.
 - -Personvern og GDPR hensyn og data blir samlet inn av BIG tech selskapene nettopp for å forbedre modellene. AI ACT og andre regulatoriske hensyn blir dog implementert, spesielt i Europa. Men er det nok og hvor effektive er de? Har ballen allerede rullet for langt?
 - -Lisensene til disse programvarene kan være ganske kostbare, og driften er ikke spesielt grønn. 1000 husholdninger. Monopolisering av CPU'er og chips til dette markedet.
- En annen svakhet/styrke, er din personlige evne til å prompt engineere – altså hvor flink du er til å samarbeide med teknologien. Og er helt fundamentalt for bruken av teknologien.
- **Jeg skal ikke prøve å vekke disse styrkene en eller annen vei, det vil opp til dere.**



IDEA



Workshop

**Intelligent Design: Creating
Exams that Measure Human
Intelligence**

Kennedy Bailey,
Rådgiver, IDEA



IDEA



Today's learning goals:

- We can:
 1. Understand why students use AI/GenAI in assessments and exams
 2. Examine the effect of GenAI on examination
 3. Consider ways to address misuse of AI by students during exams
 4. Review guidelines and regulations

IDEA



1. How are students using AI?

- 2023 Student Barometer survey: **61%** of students have used AI in their studies
- Uses include:
 - Explain topic, syllabus, concept, or terminology (70%)
 - Discussion partner (40%)
 - Translation services (30%)
 - Generating text (20%)

Hystad, Joar. "Dette Bruker Studentene Kunstig Intelligens Til." Khrono, 15 Feb. 2024, www.khrono.no/dette-bruker-studentene-kunstig-intelligens-til/849733.

IDEA

1. Why are students using AI?

If I wanted to run a marathon...



I might use a tool, like a running app, to help me achieve my goal.



But if I then used a remote-control car to finish my running plan...



IDEA

1. Why are students using AI?



Then it is VERY unlikely
that I will achieve my
goal of running a
marathon!

IDEA



1. Why are students using AI?

If you want your students to use tools efficiently...

- Make sure that they understand:
 - What the learning goal IS
 - How course assessments/exam helps them progress TOWARDS the goal.

The task isn't to complete the running app workouts; it's to cross the finish line on race day.



IDEA



2. The Effects of AI on Exams

Quick poll: do you think that students who use AI during exams are more likely to do better or worse than their peers?

Anecdotal experiences tell us: **WORSE!**

But why?

- Which students are using AI for exams?
- WHY are they using it?



2. The Effects of AI on Exams

Student Use	Potential effect on exam performance
Explaining topic, syllabus, concept, or terminology (70%)	Misinformation or decontextualized content
Discussion partner (40%)	Misinformation or decontextualized content
Translation services (30%)	Out-of-context mistranslations
Generating text (20%)	Student fails to display relevant connection to course material and experiences

IDEA



Why GenAI is often more harmful than helpful for students:

What's the issue with this response?

2. The Effects of AI on Exams



What is the most cited economics paper of all time



The most cited economics paper of all time is "A Theory of Economic History" by Douglass North and Robert Thomas, which was published in The Journal of Economic History in 1969. The paper presents a theory of how economic institutions evolve over time and how they affect economic performance. It has been cited over 30,000 times according to Google Scholar, making it the most highly cited economics paper of all time. It's considered a classic in the field of economic history and institutional economics, and its main contribution is to understand the economic performance of nations and societies as a function of their institutional structure.

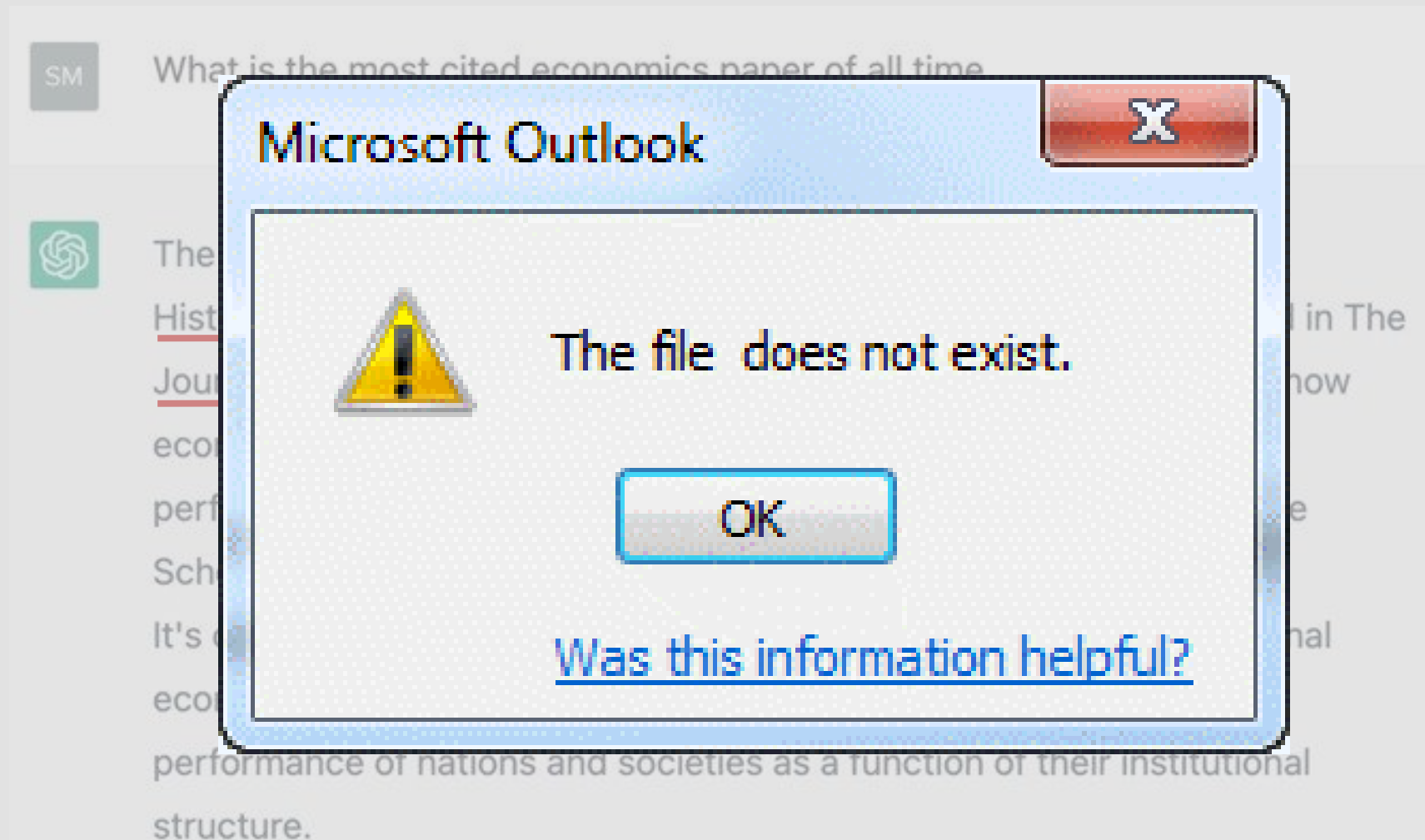
IDEA



Why GenAI is often more harmful than helpful for students:

What's the issue with this response?

2. The Effects of AI on Exams



IDEA

2. The Effects of AI on Exams

How do I measure human intelligence?

Human intelligence is contextualized!

Artificial vs. Human Intelligence

AI: Intelligence is based on statistical likelihood

HI: Intelligence is contextualized in real world experiences

IDEA



(Online version: presentation notes)

- We can start to think of GenAI not as this technological interference to be defeated in order to prevent the downfall of academic integrity, but rather as a tool that demands we consider and continue to build our pedagogical practices in response to a developing digital landscape.
- I am not suggesting that we all become AI experts and learn to immediately transform our learning environments for the digital sphere. Some of us might not have any desire for themselves or for their students to use GenAI in our classrooms at all.
- But what do these technological developments demand of us as educators? How do we continue to meet our goal of building effective learning environments, or, more topical for the day, to develop and perform effective exams that accurately measure our students' progress?



3. Addressing Misuse of AI: Human Intelligence

One of the best ways we can discourage students' use of GenAI on exams is by being intentional with our exam design.

- Clarify learning goals
 - «Skill development/content mastery» + «resource used» + «goal approached/achieved»

I can complete a running training regime using Nike Run Club in order to successfully run the Oslo Marathon.



3. Addressing Misuse of AI: Human Intelligence

CONTEXT	LEARNING GOAL	ASSESSMENT FORM
Students learning via lecture and assigned readings in order to build foundational subject knowledge.	<i>Students can accurately recall and explain course content using lecture notes and reading materials in order to develop a holistic understand of x.</i>	Traditional exam
Students participating in praxis experiences in order to practically apply learned theories	<i>Students can apply theoretical pedagogical knowledge using the classroom context in order to better understand their role as a classroom teacher.</i>	Reflection paper

IDEA

3. Addressing Misuse of AI: AI Resistance

AI-RESISTANT ASSIGNMENTS

Some faculty are already employing strategies to subvert the impact of artificial intelligence on their assessment of students, including:



Oral exams



In-class writing



Oral presentations



Reflection papers



Group discussions and debates



Hands-on experiments or labs



Peer teaching



Field trips and reports



Simulations and role plays



3. Addressing Misuse of AI: Future Integration?

UiO is still in process of developing and publishing standardized procedures for academic use of AI.

However, we can consider potential future applications where GenAI is a part of the exam.

- Home exams with submission of GenAI chat logs
- Academic text with cited and explained use of GenAI (with chat logs)
 - *Requires additional critical reflection*

If these become approved practices, we open the potential to help students use the tools at their disposal to their advantage!

IDEA

4. UiO Guidelines on AI

UNIVERSITETET
I OSLO

NO EN

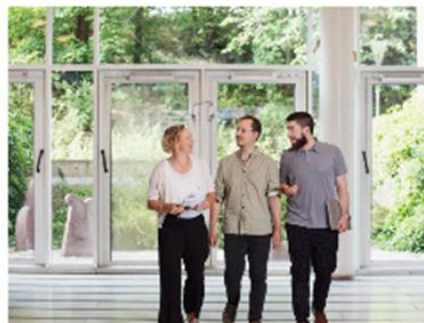
Meny

← Tjenester og verktøy

Kunstig intelligens (KI) ved UiO

Utviklingen innen kunstig intelligens vil påvirke vår læring, kunnskapsutvikling og arbeidsmåte. På denne siden formidler vi perspektiver, prinsipper og innsikt i bruk av KI i høyere utdanning.

Godkjente KI-verktøy på UiO



→ Hvordan bruke KI som underviser



IDEA

4. UiO Guidelines on AI



UNIVERSITY OF OSLO

← Services and tools ← IT services

Artificial intelligence and GPT

About AI and GPT at UiO

Artificial intelligence (AI) is becoming increasingly available and increasingly used. AI is used in a number of services, and here at UiO the technology is available to all IT users at UiO in the services below.

See also:
[The UiO web site about good use of AI](#)

Services

Show what is allowed to store

GPT UiO IT in research, IT in education	● Yellow and green data*
Autotekst – speech to text with Whisper from OpenAI Audio and video recording, IT in education, Web publishing and search	● ● ● Up to red data*
Microsoft Copilot (Bing Chat Enterprise) IT in research, IT in education, Administrative IT services	● Green data

GPT UiO

Ny samtale

UNIVERSITY OF OSLO

Autotekst

UiO : Universitetet i Oslo

COPILOT

IDEA



RECAP:

Key Points and Takeaways

1. Understand why students use AI/GenAI in assessments and exams
2. Examine the effect of GenAI on examination
3. Consider ways to address misuse of AI by students during exams
4. Review guidelines and regulations



IDEA

Workshop



**Tilrettelagt
diskusjon**

Tenk, par, del





Finn diskusjonsguiden i e-postbekreftelsen

TENK

Bruk fem minutter til å reflektere selvstendig over diskusjonsspørsmålene.

Skriv refleksjonene dine inn i diskusjonsguiden.

PAR

Bruk fem minutter til å diskutere refleksjonene dine med en partner/i en liten gruppe.

Velg én person som vil være villig til å dele med hele gruppen!

DEL

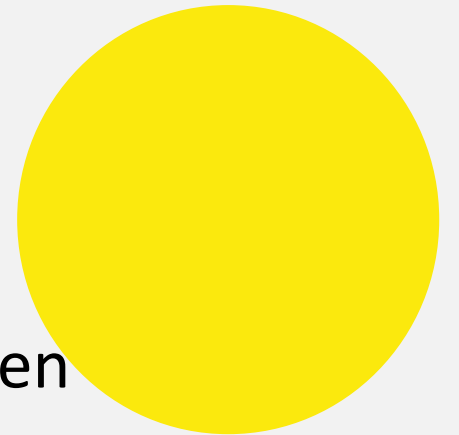
Vi vil bruke fem minutter på å diskutere alle våre refleksjoner som gruppe.

Vi kan bruke disse ideene i vårt kommende gruppearbeid.



TENK, PAR, DEL: DISKUSJONSSPØRSMÅL

- Hvor mange av studentene dine bruker AI? Hvordan og hvorfor bruker de det på denne måten?
 - Hvis du ikke vet svaret, hvordan kan du gå frem for å finne ut svaret?
- Vurder den potensielle effekten av GenKI på eksamen(er) innen ditt fagfelt.
 - Hvordan kan GenKI påvirke eksisterende læringsmål?





IDEA

Workshop



Gruppearbeid

Konseptualisering av
eksamensdesign



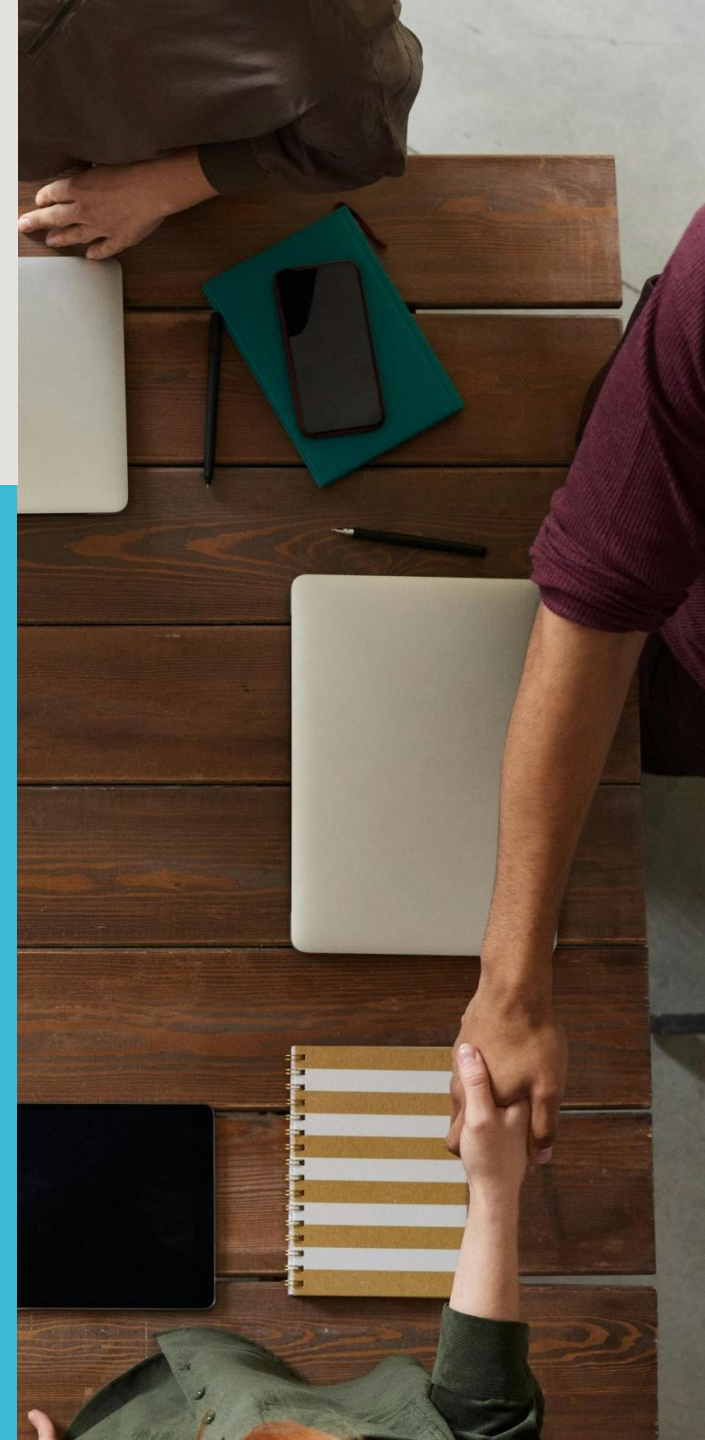
Gruppearbeid

*Premiss: minst halvparten av studentene i et emne skal bruke UiO GPT i arbeidet med hjemmeksamen. Spørsmål: Hvordan kan vi utforme eksamener som vil stille krav til **studentenes** kompetanse?*

1. Innhold og kompetanser:
2. Identifiser hvilket faglig innhold som skal vurderes
3. Bestem læringsmål, vurderingsform(er), og eventuelt kriterier

2. Design for eksamen (skisse, stikkord).

3. Hvordan inkluderer/ekskluderer denne oppgaven AI-integrasjon? Konsekvenser?





IDEA

Workshop



Q&A

Elizabeth Beech and Morten Andreas
Høglend, *IDEA StudentTeam*

Bjørn Sverre Gulheim
UiO IT





IDEA

Workshop

Tilbakemelding



Sluttrefleksjon

