# *Læring i feltet mellem hjælpemiddel og plagiat*

## GAI – sikkerhed og GDPR

# *Læring i feltet mellem hjælpemiddel og plagiat*
## GAI – sikkerhed og GDPR Udgangspunkt 2023-24

*Optimale løsning:*

- Databehandler aftale med OpenAI

*Alternativer*

- Dummy Mails og telefonnumre til alle involverede?
- Add On : 3. parts leverandør til lagring af data på lukket disk?
- MS BING-CHAT?

*Gangbar løsning (bygger på vurdering fra Bech-Bruun + forlæg fra Skanderborg Gymnasium).*

- Retningslinjer for elever og læreres brug.

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

## Kodeks 2023-24 I/II

Disse retningslinjer gælder for elever og lærere i forbindelse med brug i undervisningen af AI-tjenester som f.eks. ChatGPT.  De skal sikre en forsvarlig datasikkerhed.

*Oprettelse af konto*

- Hvis du ikke allerede har en konto på den pågældende tjeneste, så brug din skolemail til at oprette en. Vælg en stærk adgangskode, som du ikke deler med andre.

- Hvis du har mistanke om, at din konto bliver misbrugt eller andre uregelmæssigheder, skal du rapportere det til din lærer og/eller Christian på IT-kontoret (chm@herning-gym.dk).

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# Kodeks 2023-24 II/II

*Beskyttelse af personlige oplysninger*

Ved brugen af tjenesten skal du være opmærksom på og bevidst om en række datasikkerhedsaspekter. Det er vigtigt, at du overholder anvisningerne.

- Din konto er personlig. Dine anvendelser af tjenesten vil kunne henføres til dig. Skolen kræver ikke, du opretter adgang til ChatGPT eller andre tilsvarende tjenester.
- Alt det, du indtaster, vil blive gemt og logget og bruges f.eks. til at udvikle tjenesten og i nogle tilfælde til profilering og videresalg til eksterne.
- Tjenesterne giver ikke garanti for, at dine oplysninger behandles fortroligt og opbevares sikkert.
- *Du skal undgå at indtaste personlige oplysninger (navn, adresse, mobilnummer, alder m.v.) og i særdeleshed personfølsomme oplysninger (billeder, sygdomme, seksualitet, religiøs overbevisning m.v.), når du bruger tjenesten.*
- *Du må heller ikke dele dine kammeraters eller andres personlige oplysninger ved brug.*

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

## Persondata – Datatilsynet

### Er AI-modellen personoplysninger i sig selv?

Datatilsynet lægger til grund, at en AI-model som et klart udgangspunkt ikke i sig selv udgør personoplysninger, men alene er resultatet af behandlingen af personoplysninger. Det svarer til, at en statistisk rapport ligeledes ikke vil anses som personoplysninger, hvis rapporten alene indeholder konklusioner og aggregerede data, der er resultaterne af den statistiske analyse.

Visse maskinlæringsmodeller kan dog angribes på forskellige måder (såkaldte *model inversion attacks* og *membership inference attacks*), der gør det muligt at re-identificere de borgere, hvis oplysninger har indgået i modellens træningsdata. Et vellykket angreb, som resulterer i re-identifikation af borgernes oplysninger i træningsdata, kan være et brud på persondatasikkerheden og skal håndteres derefter.

Risikoen for, at en ondsindet aktør genidentificerer borgere ved bevidst at gennemføre et angreb for at udlede data, der har indgået i træningsdata, indebærer efter Datatilsynets opfattelse således ikke, at modellen skal anses som personoplysninger i sig selv.

# Retnings-linjer

## Valg af Gen-AI

**Gemmer systemet data?**

fx login, navn, e-mail, telefonnummer osv.

**Nej**

Lav retningslinjer til elever og ansatte der beskytter personoplysninger

**Ja, i EU**

- Analyser datastrømme
- Lav konsekvensanalyse
- Indgå databehandleraftale

Lav retningslinjer til elever og ansatte og oplys om databehandlingen

**Ja, uden for EU**

Dette er grundlæggende problematisk i forhold til GDPR

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# CHATGPT

| | |
|---|---|
| Firma: OpenAI | https://openai.com/chatgpt |
| Er databehandleraftale mulig: NEJ | Vurdering lavet 01.02.2024 af Rune Gråbæk rune@zbc.dk |

OpenAI tilbyder tre former for adgang til ChatGPT. En gratis konto (GPT3.5), en plus konto (GPT4.0 20$) og en team konto(GPT4.0 30$). Ved alle tre konti afgives personhenførbare data til OpenAI i form af Navn, mail og telefonnummer. Dermed bliver OpenAI en databehandler for skolen og skal leve op til GDPR.

CIU har fået vist dokumentation fra OpenAI af en vurderingsbeskrivelse af eventuelle og mulige konsekvenser ved overførsel af data til USA og i øvrigt til andre lande på stort set hele kloden i forbindelse med brug af Open AI (ChatGPT). Vurderingen fra Rune Gråbæk, IT- og digitaliseringschef ved ZBC er: *"Jeg mener umiddelbart dokumentet er indholdsmæssigt værdiløst og ikke viser en tilstrækkelig beskyttelse af personoplysninger."*




Herning Gymnasium

# COPILOT (BING)

| | |
|---|---|
| Firma: Microsoft | https://copilot.microsoft.com |
| Er databehandleraftale mulig: (JA) | Vurdering lavet 01.02.2024 af Andreas Kubsch andreas@ciuud.dk |

CoPilot er også navnet for Microsofts Generative AI som bruges via browser. Den kan bruges på stort set samme vilkår som en hvilken som helst anden online tjeneste uden et login (fx google, amazone, reddit). Her vurderes det ikke nødvendigt med en databehandleraftale. Til gengæld har den en begrænsning på 2000 tegn i en prompt og 5 prompt i en tråd.

Med en skolelicens (A3 eller A5) kan en underviser bruge Copilot som chatrobot i et beskyttet miljø i browseren bing (den normale browser på en Windows computer). Her kan man bruge den som chatrobot, søge på nettet og arbejde med billeder. Dog gemmer den ikke samtalerne og prompts må max være 4000 tegn med en chattråde med 20 beskeder (dog op til 18.000 tegn i deres "notesbog"). Denne adgang er ikke tilgængelig for elever under 18 år, men Microsoft undersøger om det kan frigives til elever ned til 13 år.

For elever vil den online tjeneste (uden login) give en funktionalitet der muliggør at eleverne kan lære færdigheder i at bruge en chatbot.

# COPILOT (BING)

Fra sommer 24:

- A3 giver elever adgang til CoPilot via deres MS-login.
- Dækket af Databehandleraftale med MS.

# COPILOT (BING)

## Technical breakthroughs

1. Next-generation OpenAI LLM model
2. Microsoft Prometheus model
3. Applying AI to core search algorithm
4. New user experience

# BARD/ GEMINI

| | |
|---|---|
| Firma: Google | https://gemini.google.com |
| Er databehandleraftale mulig: NEJ | Vurdering 7. marts 2024 af Peter Bruus |
| Google skriver i deres privatlivspolitik at de indsamler og behandler personlige oplysninger. Det er på nuværende tidspunkt ikke muligt at fravælge eller lave aftaler omkring databehandling og derfor ikke lovligt i en sammenhæng hvor en skole bærer et dataansvar. | |

Adobe ☰ Adobe

# Brugerretningslinjer for Adobes generative AI

## 1. Ingen AI-/ML-træning

Når du bruger vores generative AI-funktioner, accepterer du, at du kun vil bruge dem til dit kreative og produktivitetsfremmende arbejdsprodukt og ikke til at træne modeller inden for kunstig intelligens eller maskinlæring.

Det betyder, at du ikke må og ikke må instruere tredjeparter i eller tillade tredjeparter at bruge noget indhold, data, output eller anden information, der er modtaget eller afledt fra generative AI-funktioner, herunder eventuelt Firefly-output med henblik på direkte eller indirekte at oprette, træne, teste eller på anden måde forbedre maskinlæringsalgoritmer eller kunstige intelligenssystemer, herunder arkitekturer, modeller eller vægtninger.

# Brugerretningslinjer for Adobes generative AI

## 2. Vær respektfuld og sikker

Brug ikke Adobes generative AI-funktioner til at forsøge at oprette, uploade eller dele stødende eller ulovligt indhold eller indhold, der krænker andres rettigheder. Denne omfatter, men er ikke begrænset til, følgende:

- Pornografisk materiale eller eksplicit nøgenhed
- Hadefuldt eller stærkt stødende indhold, der angriber eller umenneskeliggør en gruppe baseret på race, etnicitet, national oprindelse, religion, alvorlig sygdom eller handicap, køn, alder eller seksuel[illegible]ing
- Eksplicit vold eller lemlæstelse
- Fremmelse af, glorificering af eller trusler om vold

Dine prompter og de resultater, der er genereret af generative AI-funktioner i Creative Cloud-produkter, kan både gennemgås via automatiserede (f.eks. maskinlæring) og manuelle metoder til forebyggelse af misbrug og indholdsfiltrering.

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

## GAI – sikkerhed og GDPR

**VIDENSKABELIGT FORLAG SÆLGER ADGANG TIL FORSKERES ARTIKLER TIL MICROSOFTS AI-UDVIKLING – AU-PROFESSOR FØLER SIG TAGET VED NÆSEN**

Den akademiske forlagsvirksomhed Taylor & Francis sælger adgang til sine forskningsartikler til techvirksomheders AI-udvikling. Forskerne, der har udgivet hos forlagets tidsskrifter, har intet fået at vide om det. Professor Peter Dalsgaard føler sig taget ved næsen og efterlyser en ny tilgang til samarbejder med større forlag.

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

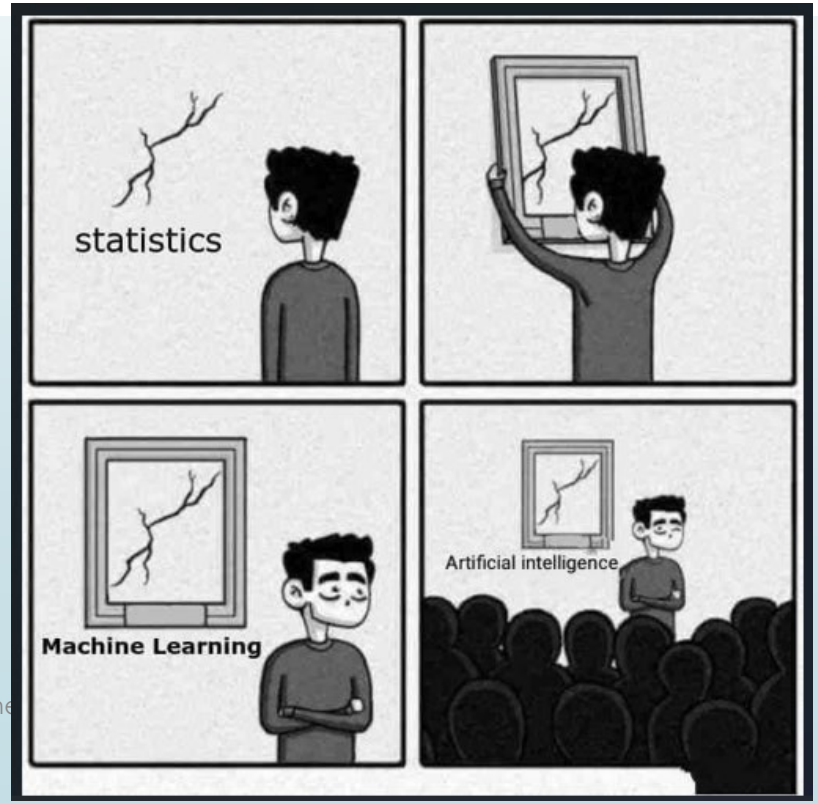
Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

## GAI – sikkerhed og GDPR

### OPHAVSRET

Undervisningsmaterialer, vejledninger, beskrivelser, billeder og andet indhold fra internettet er ofte underlagt ophavsret. Det betyder i praksis at man ikke uden godkendelse kan uploade det til en online Gen-AI model uden et samtykke. En underviser må derfor ikke tage eksterne dokumenter og bearbejde dem med CoPilot eller ChatGPT. For yderligere information kontakt Peter Leth på peterleth@ciuud.dk

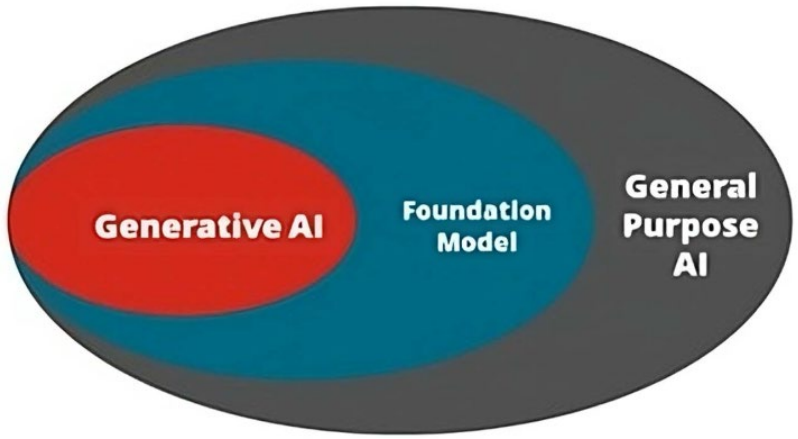





Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

# EU's AI-forordnina




*First, the AI Act requires providers of general-purpose AI models, such as those used for natural language processing or computer vision, to put in place* ***a policy to respect EU copyright law****, in particular to identify and respect the reservations of rights expressed by right holders. This obligation applies regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of these models take place.*

*Second, the AI Act requires providers of general-purpose AI models to draw up and make publicly available a sufficiently* ***detailed summary about the content used for training of the model****, such as the main data collections or sets and a narrative explanation about other data sources used. This obligation aims to enhance the transparency and accountability of AI systems and to facilitate the enforcement of rights and remedies by right holders and other affected parties.*

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# *Læring i feltet mellem hjælpemiddel og plagiat*

## GAI – sikkerhed og GDPR

### Grading Foundation Model Providers' Compliance with the Draft EU AI Act

| Draft AI Act Requirements | GPT-4 | Cohere Command | Stable Diffusion v2 | Claude | PaLM 2 | BLOOM | LLaMA | Jurassic-2 | Luminous | GPT-NeoX | Totals |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Data sources | ● ○ ○ ○ | ● ● ● ○ | ● ● ● ● | ○ ○ ○ ○ | ● ● ○ ○ | ● ● ● ● | ● ● ● ● | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | 22 |
| Data governance | ● ● ○ ○ | ● ● ● ○ | ● ● ○ ○ | ○ ○ ○ ○ | ● ● ● ○ | ● ● ● ● | ● ● ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ○ | 19 |
| Copyrighted data | ○ ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | 7 |
| Compute | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | ● ● ● ● | ○ ○ ○ ○ | ● ○ ○ ○ | ● ● ● ● | 17 |
| Energy | ○ ○ ○ ○ | ● ○ ○ ○ | ● ● ● ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | ● ● ● ● | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ● | 16 |
| Capabilities & limitations | ● ● ● ● | ● ● ● ○ | ● ● ● ● | ● ○ ○ ○ | ● ● ● ● | ● ● ● ○ | ● ● ○ ○ | ● ● ○ ○ | ● ○ ○ ○ | ● ● ● ○ | 27 |
| Risks & mitigations | ● ● ● ○ | ● ● ○ ○ | ● ○ ○ ○ | ● ○ ○ ○ | ● ● ● ○ | ● ● ○ ○ | ● ○ ○ ○ | ● ● ○ ○ | ○ ○ ○ ○ | ● ○ ○ ○ | 16 |
| Evaluations | ● ● ● ● | ● ● ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ○ ○ | ● ● ● ○ | ● ● ○ ○ | ○ ○ ○ ○ | ● ○ ○ ○ | ● ○ ○ ○ | 15 |
| Testing | ● ● ● ○ | ● ● ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ○ ○ | ● ● ○ ○ | ○ ○ ○ ○ | ● ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | 10 |
| Machine-generated content | ● ● ● ○ | ● ● ● ○ | ○ ○ ○ ○ | ● ● ● ○ | ● ● ● ○ | ● ● ● ○ | ○ ○ ○ ○ | ● ● ● ○ | ● ○ ○ ○ | ● ● ○ ○ | 21 |
| Member states | ● ● ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ○ ○ | ● ● ● ● | ○ ○ ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ○ ○ ○ | ○ ○ ○ ○ | 9 |
| Downstream documentation | ● ● ● ○ | ● ● ● ● | ● ● ● ● | ○ ○ ○ ○ | ● ● ● ● | ● ● ● ● | ● ● ○ ○ | ○ ○ ○ ○ | ○ ○ ○ ○ | ● ● ● ○ | 24 |
| Totals | 25 / 48 | 23 / 48 | 22 / 48 | 7 / 48 | 27 / 48 | 36 / 48 | 21 / 48 | 8 / 48 | 5 / 48 | 29 / 48 | |

Claus Gregersen, uddannelseschef, Herning Gymnasium, cg@herning-gym.dk

Herning Gymnasium

# Diskussion - Spørgsmål