

Kompetensutvecklande aktivitet 3 – Etik, möjligheter och risker

Mål:

- Reflektera kring etiska aspekter
- Värdera potentiella möjligheter och risker med användandet av AI

Nyckelbegrepp:

Algoritmisk diskriminering, bias (vinkling), övervakning, datalagring, GDPR, integritet, makt, monopol, fusk

Information om modulen:

Denna modul är tänkt att genomföras i två steg. Deltagarna får först förbereda sig individuellt (1 h) för att sedan delta vid en gemensam fysisk träff med de andra deltagarna (2 h). Material och frågor för förberedelsen finns tillgängligt för deltagarna via någon plattform. Den gemensamma träffen leds av en handledare. Deltagarna behöver vid det tillfället ha tillgång till någon digital enhet (dator, mobil, e. dyl.) Total tidsåtgång för både förberedelse och genomförande är tänkt att vara 3 h.

Individuell förberedelse inför träff (1h).

Upplevelse/Inspiration:

Titta på:

How I'm fighting bias in algorithms | Joy Buolamwini (9 min)

https://www.youtube.com/watch?v=UG_X_7g63rY

AI, ain't I a Woman? - Joy Buolamwini (3 min)

<https://youtu.be/QxuyfWoVV98>

Läs:

Bhargava, R. (2018). The Algorithms Aren't Biased, We Are. MIT Media Lab.

<https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2> (5 min läsning)

HAI Issue Brief: The European Commission's Artificial Intelligence Act

https://hai.stanford.edu/sites/default/files/2021-06/HAI_Issue-Brief_The-European-Commissions-Artificial-Intelligence-Act.pdf (6 sidor)

What to expect from artificial intelligence av Ajay Agrawal, Joshua S. Gans och Avi Goldfarb (2017)

<https://static1.squarespace.com/static/578cf5ace58c62ac649ec9ce/t/589a5bfe20099e9d2837a707/1486511104226/What+to+Expect+From+Artificial+Intelligence.pdf>

(6 sidor)

Individuell förberedande reflektion:

Reflektera över innehållet ovan du tagit del av. Är det någon etisk aspekt som du tycker är extra viktig? Vad vill vi använda AI till? När lämpar det sig väl att använda AI?

Gemensam fysisk träff alla deltagare + handledare (2 h).

Praktiskt arbete (60 min):

Genomförs förslagsvis i mindre grupper om 3–4 deltagare. Handledaren går runt mellan grupperna och ser till att alla kommer i gång.

1. Hitta på och beskriv någon tillämpning av AI som enligt er mening skulle vara olämplig ur ett etiskt perspektiv! Utgå ifrån en eller flera av nyckelbegreppen till denna modul. Tillåt er att vara lite dystopiska och beskriv så detaljerat som möjligt. Motivera även tydligt varför er tillämpning är olämplig.
2. Generativ AI använder maskininlärning där en modell tränas på sådan data som man vill lära modellen att skapa. På senare tid har vi sett hur flera tjänster poppat upp som kan generera både text, bilder och musik. Träningsdatan är nästan alltid insamlad via internet av stora företag som till exempel Google eller Microsoft där miljontals användare utan ersättning bidragit med data. Tjänster baserade på dessa modeller säljs sedan vidare på prenumerationsbasis till de som har råd. Föreställ er att det vore annorlunda. Beskriv ett scenario där generativ AI kan utvecklas i en annan riktning och användas på ett etiskt försvarbart sätt för att hjälpa alla människor lösa en del av de globala problem och utmaningar vi står inför!
3. Sök bland nyhetsartiklar via internet efter minst tre exempel på hur AI används för övervakning av folket i mer eller mindre auktoritära stater. Hjälp varandra i gruppen att vara källkritiska och dubbelkolla gärna den information ni hittar mot flera källor!

Rast (15 min)

Gemensam reflektion/diskussion (45 min):

Genomförs förslagsvis i lite större grupper där flera av de mindre grupperna slagits ihop alternativt med samtliga deltagare. Handledaren styr samtalet.

Del I: Samhällsnyttan med AI

- Lyft kortfattat varje grupps olämpliga tillämpningar av AI.
- Vad vill vi använda AI till? När lämpar det sig väl att använda AI?

Del II: AI inom högre utbildning

En av uppgifterna ni arbetade med innan rasten handlar om generativ AI. Ett exempel på generativ AI är tjänsten ChatGPT som lanserades i slutet av 2022 av forskningsbolaget OpenAI och som delvis finansieras av Microsoft. Nu har Microsoft redan börjat integrera generativ AI i en del av sina tjänster och planerar att fortsätta göra så även framöver.

- Var går gränsen till fusk när generativ AI används inom studier och forskning?
- När lämpar det sig väl att använda AI inom högre utbildning?