

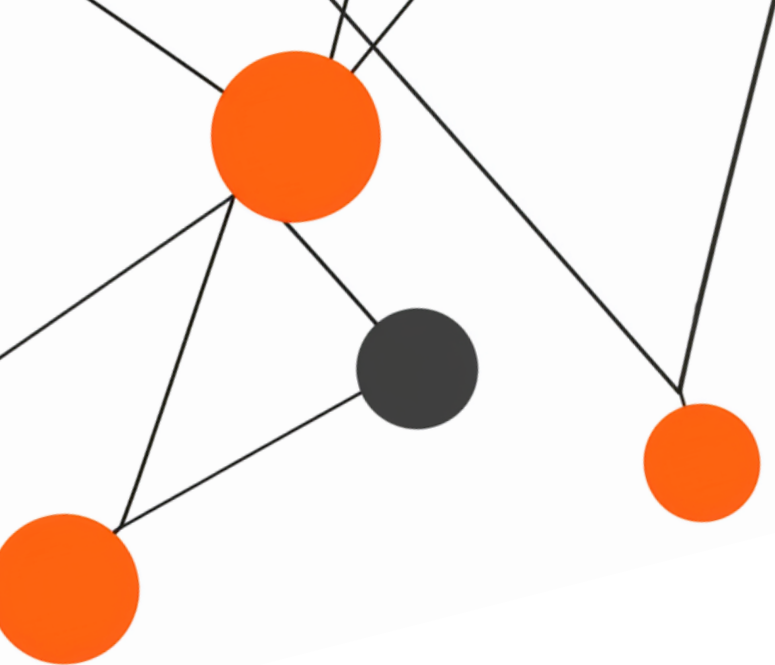
DIGITAL
INFORMATIONSLITTERACITET

AI GUIDE för lärare



FaktaBaari **EDU**

MINNA ASLAMA HOROWITZ,
ARI HAASIO, PIPSA HAVULA,
KARI KIVINEN, JARI LARU,
MIIKA MIININ, HARTO PÖNKÄ,
PATI RUIZ, AKI SAARIAHO,
PASI SILANDER OCH
TARMO TOIKKANEN.



Digital informationslitteracitet AI GUIDE för lärare

ISBN 978-952-65508-5-5 (mjuka pärmar)

ISBN 978-952-65508-6-2 (PDF)

2025 Avoin yhteiskunta ry / Faktabaari EDU

Översättningen av guiden har finansierats
av Svenska folkskolans vänner rf.

Ansvarig redaktör: Kari Kivinen

Författare: Minna Aslama Horowitz, Ari Haasio,
Pipsa Havula, Kari Kivinen, Jari Laru, Miika Miinin,
Tiina Härkönen, Harto Pönkä, Pati Ruiz, Aki Saariaho,
Pasi Silander och Tarmo Toikkanen.

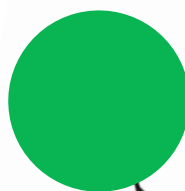
Citat: Kivinen, K., Aslama Horowitz, M., Haasio,
A., Havula, P., Laru, J., Miinin, M., Pönkä, H., Ruiz,
P., Saariaho, A., Silander, P. & Toikkanen, T. (2025)
Tekoälyopas opettajille. Faktabaari, Helsinki

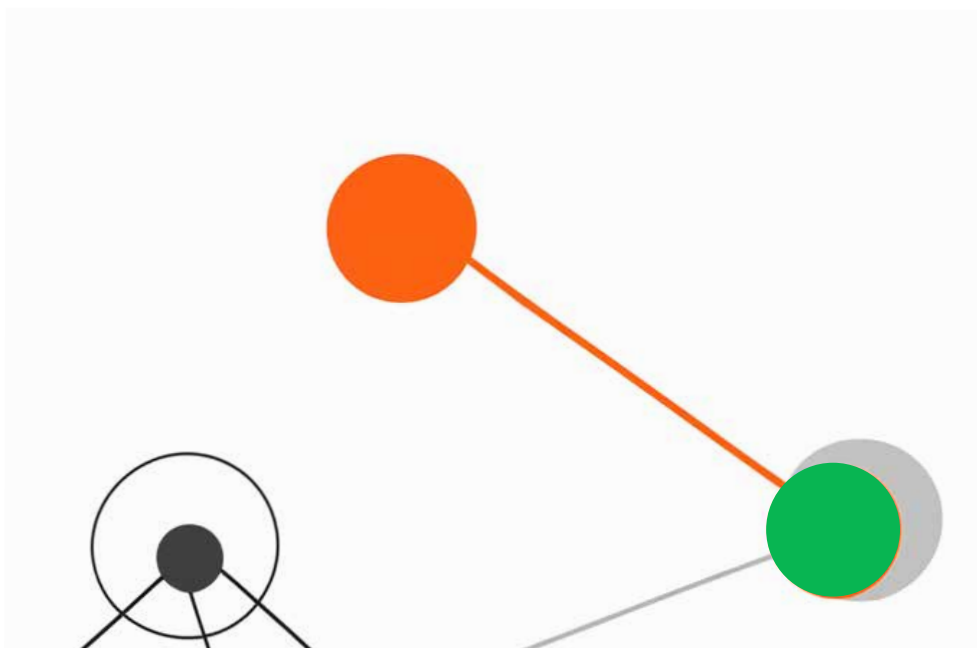
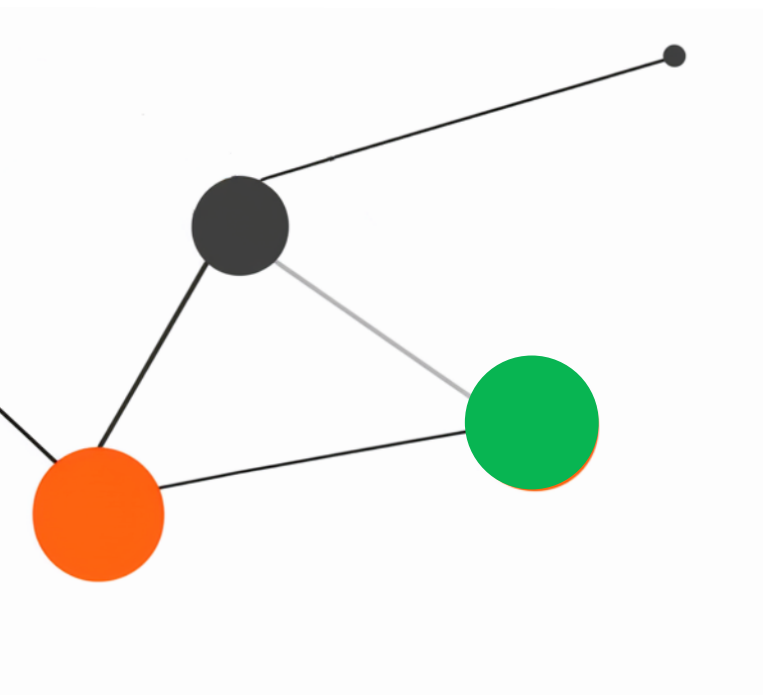
Omslag och layout: Elina Salminen

Innehåll

1. Inledning	5
2. SNABBGUIDE: Ta i bruk GenAI-verktyg nu!	8
3. AI som epistemisk förändring i undervisning och lärande	10
4. Hur fungerar generativ artificiell intelligens?	13
5. Digital kompetens, informationsresiliens & lärare	19
6. Etiska utmaningar för utbildningsanordnare och lärare vid användning	25
7. Referensramar för AI-kompetens för lärare	30
8. Bedömningsskala för användning av AI	34
9. Vad bör en lärare veta om EU:s AI-förordning?	36
10. Undervisning om AI-färdigheter för elever med hjälp av Generation AI-verktyg	39
11. Generation AI-projektet	44
12. AI, det gör ont – AI förändrar pedagogiken i Villmanstrand	46
13. Förnekelse vs teknooptimism: AI ur lärarens synvinkel	49
14. Deepfake härmar verkligheten	52
15. AI-baserade rekommendationsalgoritmer på sociala medier	55
16. AI – hot eller möjlighet för demokratin?	65
17. AI och kreativitet – risk eller möjlighet?	67
18. Amerikanska AI-rekommendationer för lärare?	70
19. GenAI, upphovsrätten och lärare	73
20. AI-verktyg och dataskydd	75
21. Bilaga: Faktabaaris anvisningar till elever	82

Denna och andra guider från
Faktabaari (på finska) finns på:
faktabaari.fi/dil





1. Inledning

KARI KIVINEN, FAKTABAARI

AI-GUIDEN ÄR en fortsättning på AI-guiden om digital informationskompetens, som Faktabaari gav ut 2024. Syftet med publikationen är att ge utbildningsanordnare, lärare och också elever grundläggande information om hur artificiell intelligens (AI) fungerar, vilka möjligheter och begränsningar och etiska utmaningar AI har, hur AI kan vara till nytta eller missbrukas och vad som avses med AI-pedagogik och AI-kompetens.

Syftet är att belysa följande frågor ur ett utbildningsperspektiv:

- **VAD** är artificiell intelligens och hur fungerar den (AI som lärandemål)?
- **HUR** kan AI användas ansvarsfullt och etiskt så att den stöder undervisningen och främjar elevernas lärande (AI som verktyg för läraren)?
- **HUR** kommer AI att förändra samhället (AI:s inverkan på samhället inom områden som arbetsmarknad, informationshantering, miljö, kommunikation, framtidens yrken osv.)?
- **VARFÖR** är det viktigt med ett sunt kritiskt förhållningssätt gentemot det som genererats av AI (Utmaningar med AI)?

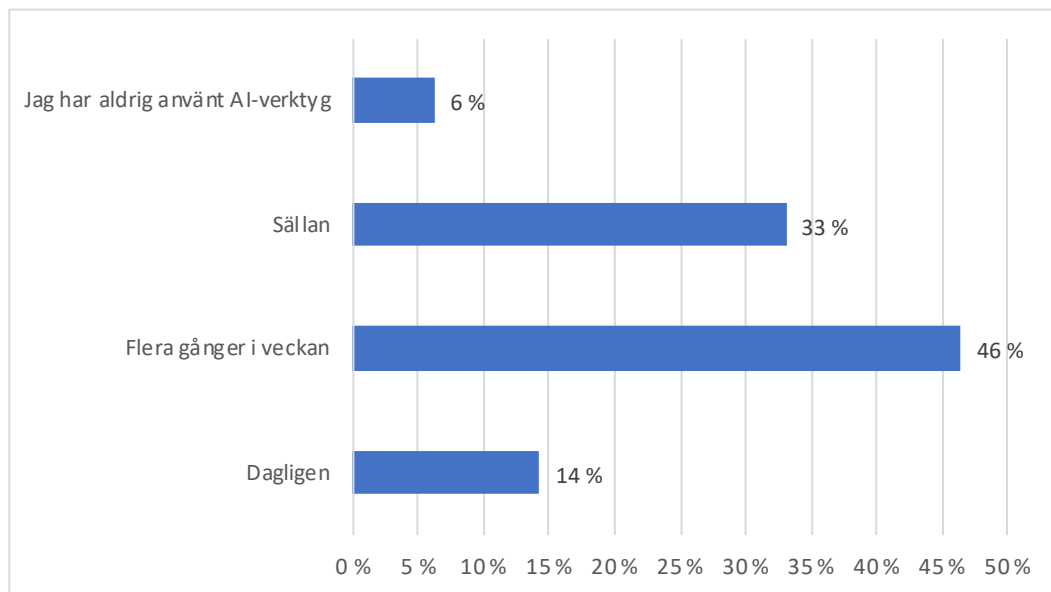
Målet har varit att skapa en så tydlig guide som möjligt, så att den är lätt att läsa och förstå. Med hjälp av källhänvisningar kan man söka efter ytterligare information.

Eftersom AI-tekniken fortfarande är i omvandling och utvecklas oerhört snabbt har vi i guiden fokuserat på kunskap som består på lång sikt.

Artikelsamlingen bygger på flera internationella rekommendationer för användning av AI i utbildning, som getts ut bland annat av EU^{1 2 3}, Unesco^{4 5 6}, OECD⁷, Unicef⁸ och andra instanser.

AI-tekniken har utvecklats så snabbt under de senaste åren att de som ansvarar för utbildning har haft svårt att hålla jämna steg med utvecklingen och att erbjuda lärare uppdaterade anvisningar och relevant fortbildning. Lärare har därför haft en rätt besvärlig sits. Hur lär man eleverna trygg och ansvarsfull användning av AI-teknik, när skolorna inte har tydliga anvisningar om vilken programvara och vilka tjänster som är säkra att använda, och hur lär man ut nya färdigheter utgående från en delvis föråldrad läroplan?

Hur ofta använder du AI-verktyg, som ChatGPT, i din vardag eller i ditt arbete? (N=112)



Förhållandet mellan utbildning och AI håller på att förändras snabbt. Utvecklingen av AI och nya, lättanvända AI-tjänster, särskilt generativ AI, har redan en betydande inverkan på undervisning och lärande. I och med att maskiner blir mer intelligenta och arbetsmarknader och ekonomier förändras måste utbildningssystemen anpassas för att utrusta studenterna med nödvändiga kunskaper och färdigheter inför framtidens arbetsuppgifter.

Den här utvecklingen förutsätter att utbildningsanordnarna uppdaterar läroplanerna

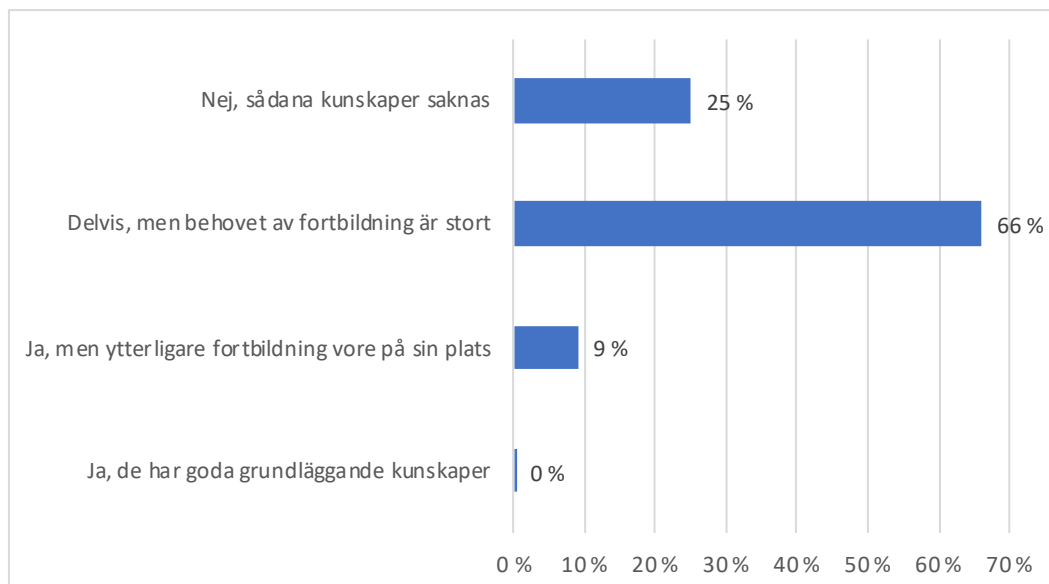
och att lärare fortbildar sig för att kunna dra nytta av den nya tekniken i undervisningen.

Finländska lärares användning av AI-verktyg

I samband med en utbildning, som ordnades av regionförvaltningsverket, fick de deltagande lärarna (N=112, oktober 2024) besvara frågan hur ofta de använder AI-verktyg i vardagen och i sitt arbete. Nästan 40 % av de deltagande lärarna uppgav att de sällan eller aldrig använder dem.

Resultatet kan bero på att många lärare

Har lärare tillräckliga grundläggande kunskaper om artificiell intelligens? (N=100)



upplever att de inte har de nödvändiga grundkunskaperna för att börja använda AI. Ingen av deltagarna upplevde att de hade goda kunskaper, och så många som 91 % (N=100) uppgav att de var i behov av fortbildning.

Mänsklig intelligens + artificiell intelligens = förstärkt intelligens

Förstärkt intelligens syftar på användning av teknik på ett sätt som kombinerar styrkorna hos människo- och maskinintelligens. Det betyder att AI inte ska ersätta mänsklig intelligens utan snarare utvidga och stödja den.

Syftet med förstärkt intelligens är att hjälpa individer och organisationer att fatta bättre beslut med hjälp av datorkraft och dataanalys⁹. Även i utbildningen uppnås de bästa resultaten när mänskligt och förstärkt intelligens får samarbeta.

AI-yrsel

Sitran utsåg nyordet "AI-yrsel" (fi. 'tekoälyhuimaus') till ett av de heta temana år 2025. Termen beskriver känslan som uppstår på grund av att artificiell intelligens utvecklas så snabbt att det är svårt att hänga med¹⁰.

AI-guiden från Faktabaari är motmedicinen mot AI-yrsel för lärare. I guiden har vi samlat bakgrundsfakta om hur AI fungerar, sammanfattningar av internationella ramverk och rekommendationer, praktiska erfarenheter från lärare som använt sig av AI och experters aktuella tankar kring AI ur olika synvinklar.

¹ Vuorikari, R., Kluzer, S. & Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens – With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48883-5, <https://doi.org/10.2760/490274>

² European Commission, Directorate-General for Education, Youth, Sport and Culture, Final report of the Commission expert group on artificial intelligence and data in education and training – A executive summary, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/65087>

³ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

⁴ Miao, Fengchun, Cukurova & Mutlu (2024): AI competency framework for teachers. Unesco <https://www.unesco.org/en/articles/ai-competency-framework-teachers>

⁵ Unesco (2023): Guidance for generative AI in education and research. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

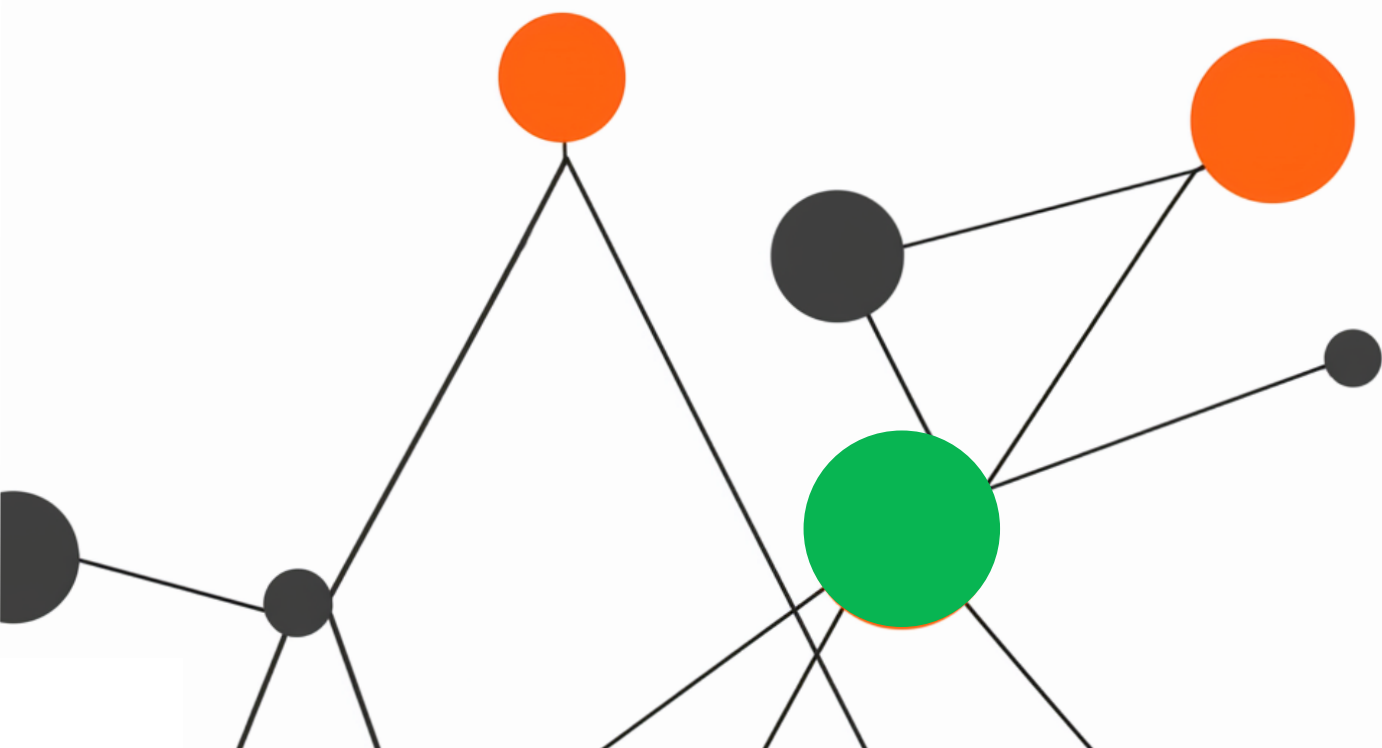
⁶ Fengchun, M. (2023): Guidance for generative AI in education and research. Unesco. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

⁷ OECD (2024) https://comite-etica.upc.edu/ca/actualitat/media/8-sti-explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_final_v3.pdf

⁸ Unicef (Hämtad: 19.12.2023): Policy guidance on AI for children. <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>

⁹ Jfr Lari Numminen (Hämtad: 31.12.2024): Mitä on tukiäly? <https://www.finnishup.com/mika-on-tukialy/>

¹⁰ Vuoden 2025 puheenaiheet (hämtad: 31.12.2024): Sitra <https://www.sitra.fi/artikkelit/vuoden-2025-puheenaiheet-pida-naita-silmalla/>



2. SNABBGUIDE: Ta i bruk GenAI-verktyg nu!

DET ÄR BÅDE LÄTT och roligt att använda AI-verktyg! Allt du behöver är en internetanslutning och en webbläsare. De flesta AI-verktyg kan användas direkt i webbläsaren.

Välj först verktyget du vill använda.

Här är en lista över sex AI-verktyg, som för närvarande är gratis:

AI-verktyg	URL	Beskrivning
OpenAi ChatGPT	https://chatgpt.com/	Mest känd, mångsidig, pionjär.
DuckDuckGo	https://duckduckgo.com/aichat	DuckDuckGo AI Chat är en gratis betaversion, som låter en chatta anonymt med AI-modeller från tredje part.
Google Gemini	https://gemini.google.com	Snabb, vidareutvecklas hela tiden
Microsoft Copilot	https://copilot.cloud.microsoft	Ingår också i sökmotorn Bing. Man kan öva att skapa bilder tex. här: https://www.bing.com/images/create .
Perplexity	https://perplexity.ai	Klarar av att föreslå källor, saklig.
Pi	https://pi.ai	Artig, användarvänlig, söt.

Varje AI-verktyg har sina egna styrkor. Du kan testa olika alternativ och hitta det som passar dig bäst. Av alla AI-verktyg finns också betalversioner, som är mångsidigare och effektivare. Det finns också gott om AI-verktyg för till exempel musikskapande, röstkloning, podcast-inspelning och videoredigering, som vi inte kommer att ta upp här.

Öppna verktyget och logga in.

I de flesta fall måste du skapa ett gratis användarkonto.

Se efter inmatningsrutan eller promptrutan:

Det är där du ska skriva det du vill fråga om eller det du vill AI-verktyget ska göra.

Skriv prompten:

Exempel på textpromptar:

- Skriv en tankadikt där följande ord ingår: vinter, snö, is, sol och vår.
- Förklara varför tokens är viktiga för hur generativ AI fungerar.
- Skriv en lärorik saga för fem år gamla barn om en räv, en igelkott och en kråka.

- **Sammanställ en lektionsplan för en lektion i årskurs 5, där temat är hur man kontrollerar källorna i en text som genererats av ett AI-verktyg.**

Du kan låta AI-verktyget precisera och förbättra det genererade svaret genom att ge preciserande promptar, tills du är nöjd med slutresultatet. Ju mer detaljerat du beskriver vad du vill AI-verktyget ska göra, desto sannolikare är det att du kan vara nöjd med svaret du får.

De flesta AI-verktyg förstår svenska och producerar grammatikaliskt korrekt text. Det är emellertid bra att komma ihåg att AI-verktygens "modersmål" i huvudsak är engelska.

Kontrollera alltid svaret som AI-verktyget genererat!

Det är viktigt att alltid kontrollera om svaret som AI genererat går att använda. Ibland kan svaren innehålla fel eller vara påhittade. Det är inte så farligt, om det är fråga om en dikt eller en saga, men med faktatexter måste man vara noggrannare.

AI:s egna instruktioner

Låt oss då fråga Google Gemini om vilka tips den ger en AI-nybörjare:

AI:s egna instruktioner

Låt oss då fråga Google Gemini om vilka tips den ger en AI-nybörjare:

Skapa en lista med de fem viktigaste tipsen för dem som inte tidigare har använt generativa AI-verktyg.

Absolut! Här är de fem viktigaste tipsen för dem som inte tidigare använt generativa AI-verktyg:

1. Ställ detaljerade och tydliga frågor:

- **Ju mer detaljerade frågor du ställer, desto precisare blir också svaret.**
- **Använd nyckelord** som beskriver den information du är ute efter så bra som möjligt.
- **Var så exakt som möjligt:** Istället för att skriva "Kan du berätta för mig om AI?" ska du fråga till exempel: "Vilka är de största utmaningarna med AI just nu?"

2. Fundera på ditt ordval:

- **Samma fråga kan ställas på många olika sätt.**
- **Prova synonymer** och olika meningsstrukturer.
- **Om du inte får ett önskat svar kan du omformulera frågan.**

3. Granska tillförlitligheten i svaret du får:

- **Lita inte automatiskt på allt:** AI kan ibland producera felaktig information.
- **Kontrollera informationen i andra källor, speciellt i viktiga beslutssituationer.**
- **Förmågan att tänka kritiskt är viktig när man använder AI.**

4. Dra nytta av styrkorna hos AI:

- **AI är ett ypperligt verktyg** för att söka och kombinera information.
- **AI kan hjälpa dig att skriva och översätta texter** och att skapa olika typer av innehåll.
- **Var inte rädd för att prova olika funktioner** och att hitta sådana sätt att använda AI som bäst passar dig.

5. Var medveten om AI:s begränsningar:

- **AI är inte en människa:** Den förstår inte kontexter lika bra som en människa, och den kan göra misstag.
- **Lita inte blint på de svar AI genererar.**
- **Använd AI som ett verktyg, som du kan ta hjälp av, men låt den inte fatta beslut åt dig.**

Kom ihåg: AI är en teknik som ständigt utvecklas, och dess användningsområden ändras också hela tiden. Var öppen inför att prova nya saker och att lära dig mer om möjligheterna med AI.

Lär dig mera: Den finska webbläraren Esa Riutta har publicerat pedagogiska videor om hur man kan använda olika AI-verktyg. Dem hittar du på hans Youtube-kanal: <https://www.youtube.com/c/webopettaja>

3. AI som epistemisk förändring i undervisning och lärande

PASI SILANDER

UTVECKLINGEN AV ARTIFICIELL INTELLIGENS har en betydande inverkan på det kunnande och de arbetslivsstrukturer som behövs på olika samhällsområden, och därmed också på skola och utbildning. Ny teknik har genom historien fått samhällen och näringsstrukturer att förändras, vilket uppfinningar som tryckpress, ångmaskin, el och halvledarteknik är goda exempel på. De möjliggjorde övergången från ett jordbrukssamhälle till ett industrisamhälle och senare till ett informationssamhälle och

"Kan vi alls längre säga vad som är verkligt och vad som skapats av artificiell intelligens? Har vi metoder för hur vi kan skilja sanningen från det osanna?"

digitaliseringens tidevarv. Dessa tekniker har haft en stor inverkan på tillverkningen av olika produkter och logistik, liksom på kommunikation och ekonomi, men ingen av dem har verkligen utmanat vår förståelse av information och kunskap – även om såväl tryckpress, internet som digitalisering visserligen avsevärt har ökat mängden information samt tillgången till information.

Utvecklingen av artificiell intelligens utmanar oss och vår uppfattning om

information och kunskap på ett helt nytt sätt. Generativ AI bygger och framställer information i hög grad på samma sätt som en människa, även om den bygger på stora språkmodeller (large language models) och alltså i praktiken på beräkningar av sannolikheter. Det är dock häpnadsväckande hur väl språkmodeller kan producera information, både i skrift och i bilder. En central fråga i sammanhanget är därför vilken typ av information eller kunskap eleverna kommer att behöva i samhället eller i arbetslivet i framtiden. På vilket sätt bör skolundervisningen förändras, så att eleverna ska kunna tillägna sig det kunnande som är nödvändigt i denna AI-era?

Epistemisk förändring av hur vi uppfattar information och kunskap

Vår förståelse av kunskap ger en inblick i vår uppfattning om vad kunskapen har för ursprung, natur, varaktighet och föränderlighet och hur kunskap uppstår eller produceras. Vi är vana vid att uppfatta

	Ytligt lärande	Djupt lärande	Framtidens lärande
Mål:	Komma ihåg information	Förstå	Skapa nya lösningar
Resultat:	Förmågan att tillämpa kunskap i enskilda situationer	Förmågan att använda och tillämpa kunskap i olika situationer	Förmågan att skapa nya lösningar för nya situationer
Metoder:	Kunskapsutbyte, upprepning av kunskap	Kollaborativt kunskapsbygge	Gemensamma innovationer, samskapande
Fokus:	Faktakunskaper	Kunnande	Mångfasetterade strategier och färdigheter för tänkande

Tabell 1. Lärandet samt målen och metoderna för lärandet måste betraktas i ett nytt ljus i och med utvecklingen av AI.²

information och kunskap som något människan har skapat och som ofta också avspeglar verkligheten och utgör en välgrundad, sann föreställning. Den rådande konstruktivistiska synen på kunskap betonar kunskapens kumulativa natur, antingen som något en individ producerat eller något som blivit till (konstruerats) som ett resultat av flera individers gemensamma kunskapsbygge. I kärnan av kunskapsproduktionen finns sociokulturella artefakter som språk eller matematik. Kunskap har tidigare ansetts kunna uppstå som ett resultat av människans kognitiva processer, dvs. tänkande.

Artificiell intelligens bearbetar, processar och producerar emellertid information, även om den inte har människans förmåga att förstå eller varsebli. Till skillnad från människan kan AI inte skapa betydelser eller tolka information, även om speciellt generativ AI blivit duktig på att ge sken av människoliknande beteende i de svar den genererar. Den information som producerats av AI baserar sig dock alltid på maskininlärningsmodeller, algoritmer och underliggande data. AI har inte människans förmåga till etiska eller moraliska reflektioner – och inte heller samvete eller förmåga att uppleva känslor eller empati. Ansvaret för kunskapen – för hur den AI-generade informationen

används och tolkas – ligger därför hos människan som använder sig av AI.

Förändring av pedagogik

Förändringen av vår uppfattning om kunskap leder oundvikligen

också till en förändring av pedagogik, eftersom man bör ta hänsyn till ett nytt slags behov av kunnande, kunskaper och färdigheter. Den konstruktivistiska synen på kunskap har starkt präglat den rådande pedagogiken ute i skolorna – den konstruktivistiska synen på lärande –, där lärandet ses som ett resultat av antingen elevernas individuella eller elevernas gemensamma kunskapsbygge.

I en AI-era behövs emellertid en annan typ av epistemisk flexibilitet¹, som kan definieras som förmågan att flexibelt kombinera olika former av kunnande, kognitiva strategier och olika former av kunskap (såsom den som producerats av AI) utgående från situationen i fråga. En central del av epistemisk flexibilitet är metakognitiva färdigheter, dvs. förmågan att reglera och styra eget tänkande.

På grund av AI kommer behovet av att lära sig till exempel att läsa och räkna eller att tillägna sig centrala begrepp i olika läroämnen inte att försvinna inom en snar framtid, men AI kommer att innebära en förändring av vilken typ av färdigheter att konstruera kunskap och tänka eleverna behöver. Eleverna behöver färdigheter till exempel i algoritmiskt tänkande (computational thinking³) och datahantering, som är nödvändiga för att de ska kunna förstå vad AI är och hur den fungerar. Sådana färdigheter behövs för att vi ska kunna förstå användningen av automatiserat beslutsfattande och maskininläring

"Uppfostrar vi elever som styrs av AI eller elever som styr AI?"

"AI är i sig en innovation, men framför allt är AI en plattform för nya innovationer."

till exempel på sociala medier, i hybridpåverkan, i nättjänster för informationssökning eller i banktjänster. Algoritmiskt tänkande är starkt förankrat i förmågan att skapa abstraktioner, i språklig problemlösning och i matematisk-logiskt tänkande.

"Färdigheterna kring AI är ofta helt nya för oss, och i bästa fall innebär de en gemensam och parallell inlärningskurva för lärare och elever."

samt problemlösning, källkritik och kreativitet, när det gäller användningen av AI. Dessa färdigheter är en förutsättning för ett aktivt deltagande samt för yttrandefrihet och stärkt demokrati. Dessutom är det viktigt att skolan fokuserar på etiska frågor kring AI och en etisk användning av tekniken.

Användningen av AI i utbildningen förutsätter att användarna förstår

AI i skolan

Skolans förmåga att bemöta förändringen i det digitaliserade samhället är bland annat beroende av vad utbildningssystemet har för syn på teknik (synen på och förståelsen av artificiell intelligens), syn på lärande och undervisningspraxis och av att kunnandet på området har hängt med i utvecklingen. Genom att utveckla dessa element kan skolan och utbildningssystemet navigera i AI-eran och se till att eleverna har försetts med uppdaterade kunskaper, som samhället och arbetslivet förutsätter av dem.

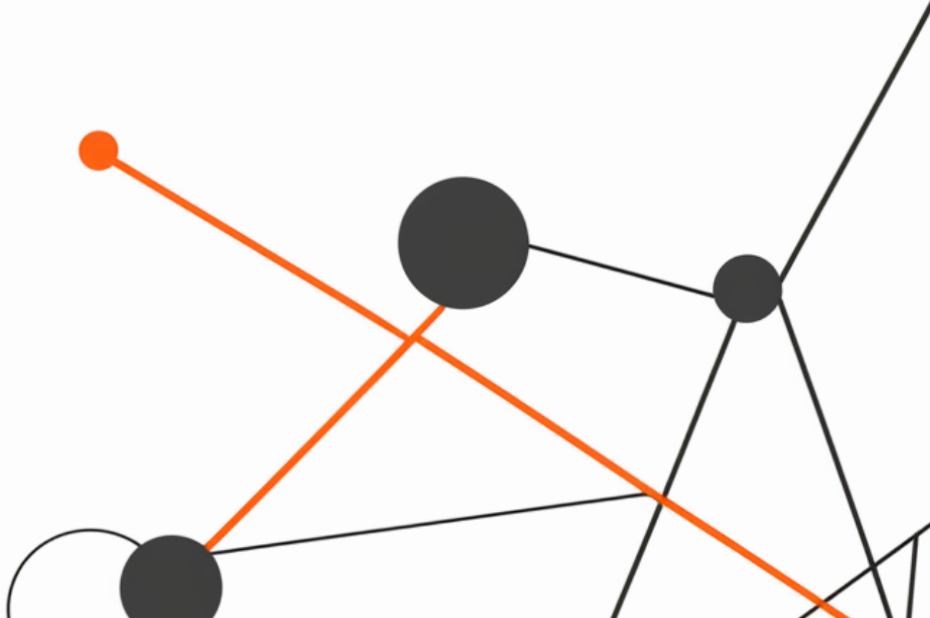
AI kan inte enbart behandlas som en ny teknik i skolan, utan användningen av AI är en central faktor som påverkar demokratin och yttrandefriheten i samhällen och utgör då ett verktyg för såväl automatisk censur, hybridpåverkan som kunskapsproduktion. En central utmaning i skolan är att förse eleverna med färdigheter kring kritiskt tänkande

vad AI är, vilken typ av felaktiga slutsatser AI kan generera och vilken typ av snedvridna resultat AI kan producera. Det är viktigt att beakta att AI-verktyg inte heller existerar utanför alla kulturella och historiska kontexter, värderingar och andra bakomliggande faktorer. AI-verktygen baserar sig för närvarande i hög grad på multinationella storbolags kommersiella utvecklingsarbete, data dessa bolag använder och kapaciteten hos de molnservrar som de tillhandahåller. I AI-verktygen har man till exempel sällan tagit hänsyn till särdragen hos små språk med få talare eller den kulturella kontext som de existerar i.

¹ Markauskaite, L. & Goodyear, P. (2017): Epistemic fluency: Innovation, knowledgeable action and actionable knowledge. Springer Dordrecht.

² Silander, P., Riikonen, S., Seitamaa-Hakkarainen, P. & Hakkarainen, K. (2022): Learning Computational Thinking in Phenomena-Based Co-creation Projects: Perspectives from Finland. I: Computational thinking education in K-12: Artificial intelligence literacy and physical computing (s. 103–119). MIT press.

³ Wing, J. M. (2008): Computational thinking and thinking about computing. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 366(1881), 3717.



4. Hur fungerar generativ artificiell intelligens?

KARI KIVINEN, FAKTABAARI

Generativ artificiell intelligens (GenAI) baserar sig på stora språkmodeller (large language models) och neurala nätverk. Det handlar om en typ av artificiell intelligens, som producerar nytt innehåll istället för att bara analysera eller kategorisera befintligt innehåll eller befintliga data.

Följande definition av generativ artificiell intelligens är en bearbetad version av definitionen som ingår i Utbildningsstyrelsens AI-rekommendationer, som sänts på remiss¹:

Med generativ artificiell intelligens (AI) avses AI-verktyg som kan producera nytt innehåll, till exempel texter, bilder och videor. Verktögen bygger ofta på maskininlärning och en stor mängd träningsdata. Utifrån träningsdata kan maskininlärningsbaserade verktyg lära sig nytt innehåll. En form av generativ AI utgörs av de så kallade stora språkmodellerna (large language models), som används för att producera och bearbeta naturligt språk som liknar människans språk.

Även om det är vanligt att interagera med GenAI på naturligt språk, är dess funktionalitet inte begränsad till det utan promptarna kan också bestå till exempel av kod eller bilder. GenAI producerar nytt innehåll med hjälp av en modell som skapats utifrån träningsdata. De stora språkmodellerna har byggts upp med antingen offentligt tillgängligt innehåll på internet eller med mer begränsade, slutna datamängder.

Utöver stora språkmodeller kan generativ AI också bygga på andra typer av modeller och tekniker, som är optimerade för specifika ändamål, såsom för att skapa ny musik, videor eller bilder. Till exempel har de så kallade diffusionsmodellerna fokus på att skapa och redigera bilder och ljud.

Stora språkmodeller

Stora språkmodeller tränas med stora mängder data. Dessa data kan bestå av böcker, artiklar, kod eller andra former av skriftlig kommunikation samt av bilder och grafik.

Datainsamling kan göras kontinuerligt på nätet tack vare så kallade datainsamlare. Nuförtiden kan de webbansvariga förhindra (eng. 'opt-out') datainsamling på sina webbplatser. Det är bra att komma

ihåg att träningsdata som samlas in också kan innehålla en stor mängd felaktig, diskriminerande och fördomsfull information.

Chatbotar och virtuella assistenter använder sig av språkmodeller för att producera svar utifrån användarnas promptar. Innehållsproducenter dra också nytta av dessa modeller för att skriva artiklar, bloggtexter och marknadsföringsmaterial. Inom utbildningssektorn hjälper modellerna lärare att skapa till exempel undervisningsmaterial och övningsuppgifter.

Språkmodellerna använder det insamlade materialet för att lära sig relationerna mellan ord och satser samt innebörden i olika typer av texter. På så sätt kan de utföra olika typer av uppgifter, såsom översätta språk, producera kreativt innehåll och svara på frågor. I takt med att AI-tekniken utvecklas kommer dessa modeller att utvecklas ytterligare och möjliggöra tillämpningar på ännu flera områden.

Språkmodellerna byggs upp med hjälp av neurala nätverk och speciellt den så kallade transformatorarkitekturen. Det är en metod baserad på djupinlärning som möjliggör ett effektivt sätt att bearbeta till exempel meningar, textavsnitt och kontexter. För att det ska vara möjligt måste texten konverteras till ett numeriskt format, vilket möjliggör komplexa matematiska beräkningar. Detta kallas för tokenisering.

Datapunkter och tokens dvs. teckensträngar

Tokenisering är en process, där text eller bilder omvandlas till ett numeriskt format, så att de kan hanteras av datorer. Under processen spjälkas en text upp i mindre delar, tokens, som kan bestå av ord, ordelar eller till och med enskilda tecken. Varje token ges en numerisk representation vid inbäddning (eng. 'embedding'), som återspeglar

dess betydelse i den aktuella kontexten. Metoden möjliggör matematikbaserade operationer mellan olika ordsekvenser (tokens), till exempel i form av jämförelser

av hur lika sekvenserna är. För att modellen ska kunna förstå innebörden i en mening eller ett textstycke är det viktigt med kontextualisering, som tar hänsyn till den textuella kontexten som ordsekvensen förekommer i.

En datapunkt utgörs av en enskild observation eller förekomst i en datamängd. En token är en mindre, odelbar enhet, som har identifierats i en datapunkt. En token kan också vara en orddel (t.ex. kan ordet "apparat" delas upp i följande tokens: "ap", "par", "at"). En token är en grundläggande byggsten som används för att framställa och bearbeta data i en AI-modell. Datapunkter och tokens är grundläggande dataenheter, som är centrala för träning och användning av AI-system.

En datapunkt kan ses till exempel som ett komplett recept med en lista över alla ingredienser, medan en token snarare motsvarar en enskild ingrediens i receptet.

AI-modellerna fungerar så att de går igenom flera datapunkter, som alla innehåller ett flertal tokens. AI-modellen analyserar relationerna mellan dessa tokens för att lära sig mallar och göra förutsägelser.

Tokens, datapunkter och mallar används för att producera ett svar utifrån en prompt. Man kan till exempel ställa en fråga till ett GenAI-verktyg i form av en textprompt, som i sin tur utgör indata för en statistisk modell. Den statistiska modellen omvandlar prompten till siffror. Därefter utför modellen beräkningar för att producera ett resultat utifrån statistiska samband.

Det mest kända generativa AI-verktyget heter ChatGPT, där förkortningen GPT står för "generative pre-trained transformer" ('generativ förtränad transformator'), vilket syftar på metoden som använts för att träna ChatGPT så att den kan hantera och förstå stora mängder textdata. Tack vare träningen kan ChatGPT ge svar på olika frågor och förfrågningar, som liknar människoproducerade svar. Modellen är som en "transformator" (eng. 'transformer'), eftersom den använder sig av en kraftfull arkitektur bestående av neurala nätverk.

**Språkmodellerna
skapas med
hjälp av neurala
nätverk.**

Exempel (skapat med Gemini 2.0 Flash Experimental)

Vid tokenisering spjälkas texten upp i mindre delar. Vid inbäddning (eng. 'embedding') omvandlas orden till siffror (vektorer). Kontextualiseringen ändrar vektorerna så att de återspeglar ordets betydelse i den aktuella kontexten. På så sätt kan datorn "förstå" att samma ord kan betyda olika saker i olika meningar. Med hjälp av vektorerna kan man också beräkna hur lika olika ord är: vektorerna som ligger nära varandra är också betydelsemässigt lika.

Ordet "bok".

1. **Betydelse 1:** Träd (lövtred)
2. **Betydelse 2:** Längre skrift

Meningar:

1. "Intill brunnen på gården stod en stor **bok**." (träd)
2. "Jag har läst en spännande **bok**." (längre skrift)

Tokenisering:

I båda meningarna har ordet "**bok**" tokeniserats.

Kontextualisering:

- Mening 1 ("Intill brunnen på gården stod en stor bok."): Kontexten "på gården" tyder på något som finns utomhus.
- Mening 2 ("Jag har läst en spännande bok."): Kontexten "läst" och "spännande" syftar på en skrift och en läsupplevelse.

Numerisk representation (inbäddning, embedding) med kontexterna:

Vid kontextualisering skapas olika vektorer för ordet "**bok**" i olika meningar:

- "bok" (träd, mening 1): [0.1, 0.9] (Den här vektorn ligger närmare vektorer för andra träd, tex. "tall", "björk".)
- "bok" (skrift, mening 2): [0.9, 0.1] (Den här vektorn ligger närmare vektorer för andra kulturupplevelser, tex. "film", "podcast".)

Vektorerna skiljer sig nu tydligt från varandra. Det här visar att ordet har olika betydelser beroende på kontext.

Vilket material används för att träna språkmodellerna?

Enligt OpenAI² har de nuvarande modellerna som används av ChatGPT utvecklats med hjälp av tre huvudsakliga datakällor:

1. **information som är allmänt tillgänglig på internet**
2. **information som OpenAI får tillgång till tack vare samarbetet med tredje part**

3. **information som användarna eller tränarna och forskarna tillhandahåller eller producerar.**

Tyvärr saknar dessa enorma datamängder ofta tydlig information om vad de innehåller och var de kommer ifrån. AI-företag berättar oftast inte vilka data de har använt för att träna sina modeller³. En orsak till det är att de vill skydda sin egen konkurrensfördel. En annan orsak är att de förmodligen inte heller själva vet var alla data

Inom ramen för det finländska forskningsprojektet Generation AI har det utvecklats en AI-applikation för skolor, som utformats särskilt med eleverna i den grundläggande utbildningen i åtanke. Syftet med projektet är att få elever, lärare och lärarstuderande att förstå säkerhetsfrågor, tekniker och modeller som berör AI eller maskininlärning med hjälp av experiment och uppfinningsrikedom. Allt utvecklingsarbete bygger på högkvalitativ internationell forskning.

Hundratals fjärde- och niondeklassare i Joensuu och Uleåborg har deltagit i skolexperiment inom ramen för utvecklingsarbetet kring maskinen som ska tränas i projektet. Applikationen har utvecklats vidare utifrån de erfarenheterna. I anslutning till projektet har det också utvecklats undervisningsmaterial kring applikationen, som finns tillgänglig på projektets webbplats. Du kan läsa mer om projektet lite senare i den här guiden.

kommer ifrån, eftersom datamängder slås ihop, paketeras och distribueras på ett komplext och ogenomskinligt sätt.

De data som används för att träna AI-modeller har också starkt fokus på den västerländska sfären: mer än 90 procent av de datamängder som analyserats av forskare härstammar från Europa och Nordamerika och mindre än 4 procent från Afrika⁴. Det engelska språkets dominans i träningsmaterialet förklaras delvis av att internet fortfarande till över 90 procent fungerar på engelska och att det fortfarande finns många platser i världen där internetanslutningar är mycket dåliga eller icke-existerande. Ett så ensidigt träningsmaterial har också en inverkan på det som AI kan genereras, och därför är det alltid förnuftigt att förhålla sig kritisk till det.

Holmes och Tuomi⁵ understryker att man inte får glömma människans roll i hur AI-system fungerar. Det är människor som samlar in eller sammanställer träningsdata (tex. bilder eller texter), som sedan används av AI, och det är människor som skriver algoritmerna, dvs. programkoden, bakom AI-tekniken och bestämmer vad de ska användas till.

Förträning och neurala nätverk

Under förträningen (eng. 'pre-training') går algoritmerna igenom en enorm mängd

datapunkter och tokens utgående från en omfattande datamassa.

Ett viktigt element i generativ AI utgörs av ett neuralt nätverk, som efterliknar de allmänna principerna för hur nervcellerna i en djurhjärna fungerar. Neurala nätverk består av sammankopplade noder ordnade i lager.

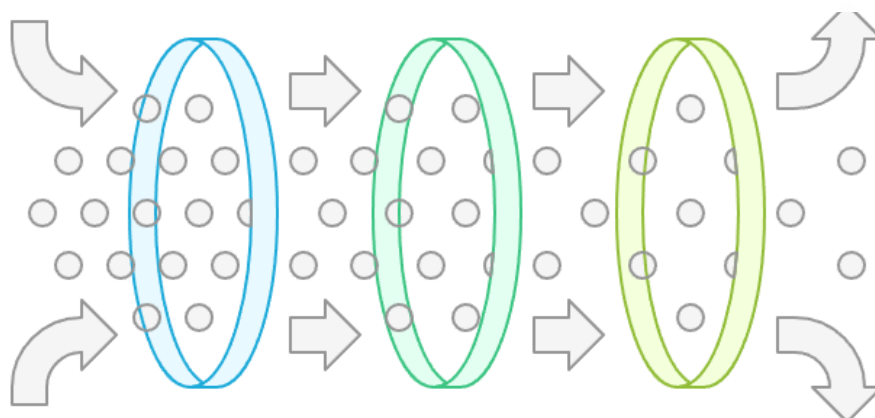
Inmatningslagret tar emot data, som därefter passerar genom ett antal dolda lager, där det utförs olika beräkningar. Det sista lagret, utmatningslagret, producerar det slutliga resultatet.

I samband med träningen lär sig AI-modellen de interna relationerna och formlerna i de aktuella data genom att justera styrkorna hos kopplingarna mellan noderna. Detta kallas för viktning. Det är en del av träningsprocessen att viktningarna ständigt justeras, så att det ska vara möjligt att minimera mängden fel och så att modellen allt bättre ska kunna göra förutsägelser och skapa innehåll. För justering av viktningar används i allmänhet optimeringsalgoritmer, varav den viktigaste för närvarande är "GTP", som ChatGPT också döpts efter.

När AI-modellen har tränats tillräckligt kommer den att kunna producera innehåll utgående

De data som används för att träna AI-modeller har starkt fokus på den västerländska sfären.

Behandling av data i ett neuralt nätverk



Inmatningslager

Tar emot och
bearbetar
originäldata

Dolda lager

Utför komplexa
beräkningar utifrån
data som matats in

Utmatningslager

Producerar det
slutliga resultatet
eller svaret

Gjord med Napkin

från prompten den får. Utifrån formlerna och kopplingarna modellen lärt sig tack vare träningsdata klarar den av att producera svar som är logiska och kontextuellt korrekta i ljuset av data den tränats med.

Generativ AI använder sig av modeller för djupinlärning, som är en typ av modeller för maskininlärning. Med maskininlärning avses att ett system på egen hand kan lära sig av de data som det matas med, utan att en människa bestämmer över alla dess funktioner. Med hjälp av djupinlärning kan AI-modellerna lära sig att identifiera mönster och kopplingar i enorma mängder data, som består av till exempel bilder eller texter. Därefter kan de producera nya, likartade data, som har samma egenskaper som originaldata. På så sätt skiljer de sig från traditionella regelbaserade AI-system, som följer vissa förutbestämda regler och som inte klarar av att lära sig eller bli bättre över tid.

För att träna AI-modeller behövs det enorma mängder data, teknik med hög prestanda och ekonomiska resurser. Endast mycket stora företag har möjlighet att utveckla sådana modeller. Som tur är kan dessa modeller användas för en mängd olika ändamål.

Även om generativa modeller är effektiva har de också begränsningar. De kan till exempel producera felaktiga slutsatser och snedvridna resultat (AI-bias).

Finjustering

Syftet med finjustering (eng. 'fine-tuning') är att anpassa grundmodellen (eng. 'foundational model') till en specifik uppgift. Modellen tränas då med en mindre, avgränsad datamängd, som berör det önskade användningsområdet. Finjusteringen förbättrar modellens prestanda på användningsområdet i fråga och gör modellen mer kostnadseffektiv.

Tack vare finjusteringen kan GenAI-verktyg utvecklas för ett brett spektrum av olika ändamål. På marknaden har det därför också dykt upp dussintals nya AI-verktyg, det ena märkligare än det andra.

Promptar eller instruktioner för AI

En prompt (input) är en order som användaren ger ett AI-verktyg. Det kan

bestå av text, ljud, bild eller programkod. Utifrån prompten producerar AI-verktyget det önskade resultatet. Ju mer exakt och detaljerad prompten är, desto bättre blir oftast också resultatet. År 2024 anges det vanligtvis i en prompt tydligt vilken roll AI

ska ha, vad man önskar för slutresultat och vilka instruktioner AI-verktyget behöver ha. Till exempel: "Sammanställ en lista över presidenterna under Finlands självständighet."

Svaret som generativ AI producerar är helt beroende av promptarna. Konsten att formulera effektiva promptar är därför en viktig färdighet som det är bra att öva på. Som tur är finns det gott om instruktioner och exempel på internet, och många AI-tjänster erbjuder också verktyg för det samt exempelpromptar. Det är också möjligt att direkt be AI-tjänster om instruktioner för hur man formulerar bättre promptar. Oftast gör man klokt i att inte lita på det första svaret ett AI-verktyg ger, utan det är bra att jobba vidare på det genom att be AI-verktyget att precisera och förbättra svaret.

För att träna AI-modeller behövs det enorma mängder data, teknik med hög prestanda och ekonomiska resurser.

Generativa AI-verktyg och sökmotorer är fundamentalt olika

Det är mycket viktigt att vara medveten om att "traditionella" sökmotorer och generativa AI-verktyg tillgodoser olika behov och tillhandahåller olika tjänster åt sina användare. Sökmotorerna söker efter befintlig information, medan GenAI-tjänster skapar nytt innehåll.

Sökmotorerna använder sig av algoritmer som analyserar innehållet i, antalet länkar på och andra egenskaper hos webbplatser, för att bestämma hur relevanta de är för och i vilken ordning de ska visas i sökresultaten. De söker

Även om de svar som GenAI-tjänster producerar ofta ger ett mycket tillförlitligt intryck kan de vara felaktiga.

efter och organiserar data på internet utifrån de sökord användaren ger. Söktträffarna är baserade på både användarens tidigare sökningar och rekommendationerna från sökmotoralgoritmen.

Den som gör sökningen

får sedan oftast tusentals eller till och med miljontals direkta länkar till material på internet. Det är på användarens ansvar att välja ut den söktträffen som bäst motsvarar det hen varit ute efter. Fördelen är att användaren själv direkt kan bedöma informationskällornas tillförlitlighet. Tyvärr är logiken bakom sökningarna i sökmotorerna, såsom Google, inte transparent, och de ordnar och censurerar söktträffar med hjälp av användarprofilering. De första söktträffarna brukar vara betalda eller sponsrade träffar, eftersom sökmotorerna drivs av kommersiella intressen.

Alldeles nyligen har GenAI-tjänsterna börjat föreslå en kombination av dessa två olika tillvägagångssätt. Till exempel erbjuder Google Gemini (version 1.5 Flash) möjligheten att "kontrollera" svaret GenAI producerat med Googles sökmotor. Man kommer åt funktionen i menyn bakom tre prickar under själva svaret ("Dubbelkolla svaret").

Om man söker information om till exempel Faktabaari med en "vanlig" sökmotor dyker det upp hundratals eller tusentals sinsemellan väldigt olika träffar, däribland kritik om Faktabaari. Frågar man ChatGPT:s webbsökningsfunktion vad Faktabaari är får man läsa en kort sammanfattande text om Faktabaari. Man kan fråga sig varför just dessa meningar har valts ut för sammanfattningen och huruvida de ger en realistisk och heltäckande bild av Faktabaari. Skulle det med tanke på medborgarnas tillgång till information kanske ändå vara bättre att få en mer nyanserad bild av Faktabaari?

Funktionsmekanismen för generativa AI-verktyg är dock en annan. De skapar nytt innehåll utgående från stora datamängder. Användaren får vanligtvis bara ett svar i taget, som baserar sig på sannolikhetsberäkningar i enlighet med processen som beskrivits ovan. Å andra sidan kan användaren be om och få till exempel hundratals svar och sedan jobba vidare på och precisera resultaten. Även om de svar som GenAI-tjänster producerar ofta ger ett mycket tillförlitligt intryck kan de vara felaktiga. GenAI-modellerna kan inte själva förstå den text de producerar, utan de baserar sig på statistiska modeller och träningsdata. Därför kan resultaten de producerar vara felaktiga eller vilseledande. Därför är det speciellt viktigt att granska resultaten GenAI producerar med källkritiska ögon. Därför står det också under promptfältet i Google Gemini så här: "Gemini kan göra fel, också när det gäller människor. Du bör alltså kontrollera alla svar du får."

Användbara länkar

- **How Chatbots and Large Language Models Works-video** <https://youtu.be/X-AWdfSFCHQ>
- **Santeri Kallio (2024): Mitä on generatiivinen tekoäly – GenAI-opas** <https://santerikallio.com/genai-opas/>
- **Ylen oppiminen: Älyä tekoälyä. Lärarens verktygslåda för undervisning (på finska)** – <https://yle.fi/a/74-20069158>

¹ Artificiell intelligens inom småbarnspedagogik och utbildning – lagstiftning och rekommendationer (hämtad 17.10.2024, lausuntopalvelu.fi) <https://www.lausuntopalvelu.fi/SV/Proposal/Participation?proposalId=a0d6af03-67e1-4ec7-9269-fab75bb05807>

² OpenAI (hämtad 13.12.2024) <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed>

³ MIT Technology Review (hämtad 18.12.2024) <https://www.technologyreview.com/2024/12/18/1108796/this-is-where-the-data-to-build-ai-comes-from/>

⁴ MIT Technology Review (hämtad 18.12.2024) <https://www.technologyreview.com/2024/12/18/1108796/this-is-where-the-data-to-build-ai-comes-from/>

⁵ Holmes, Wayne och Tuomi, Ilkka (2022): State of the art and practice in AI in education. Special Issue: Futures of artificial intelligence in education. European Journal of Education. Volume 57, Issue 4. <https://onlinelibrary.wiley.com/doi/10.1111/ejed.12533>

5. Digital kompetens, informationsresiliens & lärare

KARI KIVINEN, FAKTABAARI

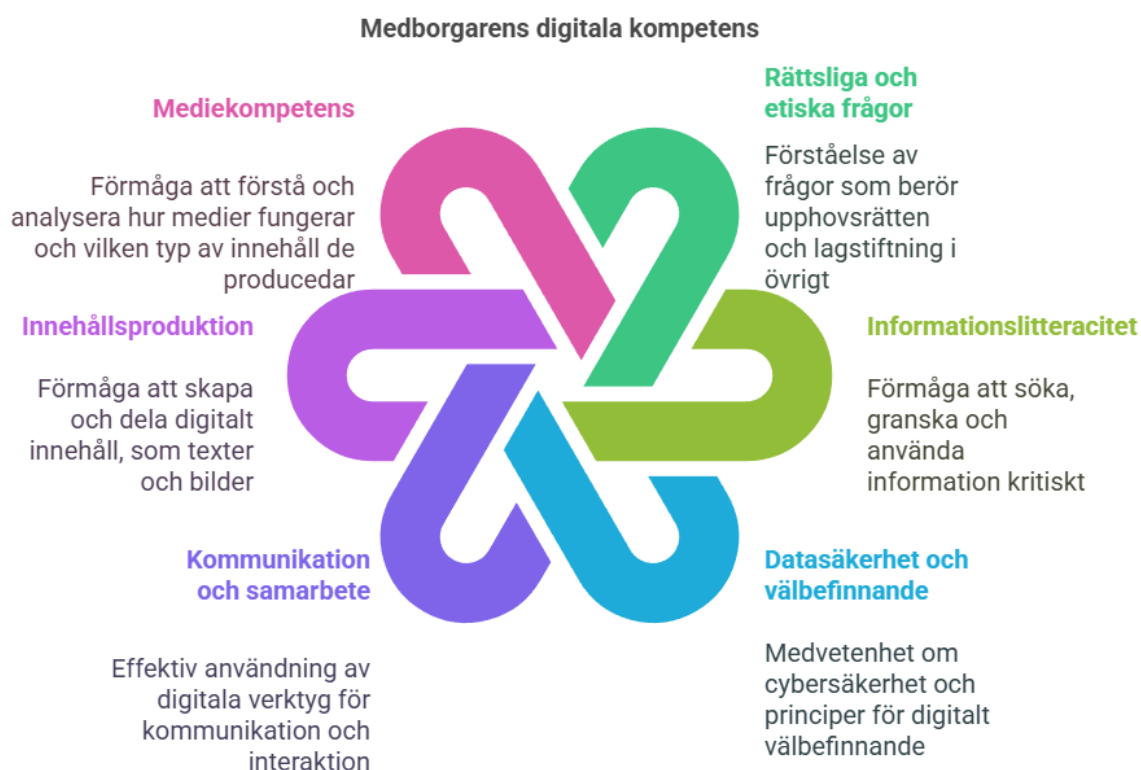
Digitaliseringen har blivit en del av vår vardag, och den har en inverkan också på utbildningssektorn. Men vad är då digital kompetens?

För att vi med tillförsikt samt kritiskt och tryggt ska kunna använda oss av nya, framväxande tekniker och artificiell intelligens bör vi ha goda grundläggande kunskaper om och förståelse av digitala tekniker och AI.

Kunskapsbeskrivningarna i Europeiska ramverket för digital kompetens DigComp 2.2¹ ger konkreta exempel på vilka kunskaper, färdigheter och attityder medborgarna behöver för att till exempel kunna utvärdera digitalt innehåll:

KUNSKAPER

- Medborgaren ska vara medveten om att onlinemiljöer innehåller alla typer av information och innehåll inklusive felaktig information och desinformation.



Digital kompetens är inte begränsad till användning av dator eller smarttelefon, utan kan bland annat omfatta följande områden.

Gjord med Napkin

- Medborgaren ska veta vikten av att identifiera vem som ligger bakom informationen som finns på internet och att verifiera den med hjälp av flera källor.

FÄRDIGHETER

- Medborgaren ska veta hur man skiljer sponsrat innehåll från annat innehåll på internet.
- Medborgaren ska veta hur man analyserar och kritiskt granskar sökresultat och aktivitetsflöden på sociala medier, identifierar deras ursprung och ser skillnad på faktarapportering och åsikter.
- Medborgaren ska veta hur man hittar författaren eller källan till informationen, för att verifiera om den är trovärdig (t.ex. en expert eller myndighet inom en relevant disciplin).

ATTITYDER

- Medborgaren ska vara beredd att faktagranska information och bedöma dess noggrannhet, tillförlitlighet och auktoritet, samtidigt som hen ska föredra primära informationskällor framför sekundära, i den mån det är möjligt.

Medborgarnas resiliens mot informationspåverkan

Under de senaste två åren har diskussionen allt starkare kretsat kring informationspåverkan, som sker på sociala medier, valpåverkan och behovet av att stärka medborgarnas resiliens.

Desinformation kan störa valprocesser och orsaka oroligheter och konfrontationer.

I 2024 års upplaga av Världsekonomiska forumets Global risks-rapport² är desinformation det allvarligaste hotet på kort sikt (2 år) och det femte allvarligaste hotet på lång sikt (10 år). Enligt rapporten kan desinformation störa valprocesser och orsaka oroligheter och konfrontationer. Det leder i sin tur till växande misstro mot medier och regeringskällor

och fördjupar polariserade synsätt även i stabila samhällen.

Det finns många varnande exempel på detta, varav det senaste och allvarligaste gäller Rumänien. Rumänien blev den 6 december 2024 första EU-landet som tvingades ogiltigförklara ett valresultat på grund av misstänkt utländsk inblandning. Europeiska kommissionen misstänker att Tiktok försummat sina skyldigheter att bekämpa yttre påverkan i det rumänska presidentvalet³.

Desinformation sprider sig nästan obehindrat på sociala medieplattformar oavsett om landsgränserna har stängts eller inte. Finland är inte heller skyddat från hotet om desinformation.

Ledartexten i tidningen Ilta-Sanomat⁴ (4.8.2024) ger en utmärkt sammanfattning av vad som bör göras: "Förutom tillförlitliga medier förutsätter kampen mot desinformation bland annat statlig beredskap för att bekämpa hotet – och en högklassig undervisning i mediekunskap. I främsta linjen i den kampen behövs vaksamma medborgare, som kan identifiera påverkansförsök och som inte blir provocerade, när de provoceras."

Resiliens uppstår dock inte av sig själv, och kunskaper och färdigheter måste uppdateras flitigt, särskilt i takt med att AI-tekniken utvecklas. Detta behövs det mer av:

- **reglering av plattformar samt spridning av information om bestämmelserna och övervakning av att de följs (t.ex. EU:s förordningar om digital kommunikation och AI utgör bra rättesnören)**
- **forskning (t.ex. Critical-gruppens internetforskning⁵)**
- **internationellt samarbete kring bekämpning av desinformation (t.ex. Faktabaari, Nordis⁶, EDMO⁷)**
- **tillförlitlig journalistik och självreglering av medier (t.ex. Opinionsnämnden för massmedier)**
- **utbildning i digital kompetens och AI (på skolor, bibliotek och arbetsplatser, som en del av militärtjänstgöringen och i vuxenutbildningen osv.)**

- **ytterligare resurser för bekämpningen av informationspåverkan.**

Digital informationslitteracitet

Digitaliseringen handlar alltså inte bara om teknisk infrastruktur, teknik, applikationer, teknisk utrustning eller internetanslutningar. I sin verksamhet har Faktabaari valt att särskilt fokusera på att i samarbete med lärare främja ungas förmåga

- **att kritiskt granska innehåll på nätet och innehåll som producerats av generativ AI**
- **att förstå hur algoritmerna påverkar sökresultat och rekommendationer**
- **att bli medvetna om att det hela tiden samlas in information om oss, som sedan används för informationspåverkan,**
- **att beakta integriteten och säkerheten i digitala miljöer.**

I sin DIL-guide⁸ har Faktabaari gett en definition av digital informationslitteracitet. Den har senare kompletterats så att AI-kompetensen har inkluderats i den digitala kompetensen:

DIL-guiden⁹ tar också upp multilitteraciteten i de finländska läroplansgrunderna samt olika internationella definitioner av litteracitet och olika former av internetkompetens och -strategier¹⁰.

Under de två år som gått sedan guiden kom ut har det visat sig vara förnuftigt att fokusera speciellt på tre frågor, när det gäller olika typer av internetinnehåll.

När man tar del av innehåll på internet, som man inte känner till är det bra att försöka hitta svar på tre enkla nyckelfrågor, innan man lägger ner tid på att utforska innehållet närmare:

1. **Vem ligger bakom informationen? Från vilken källa härstammar informationen?**
2. **Vad baserar sig påståendet på? Finns det belägg som stöder påståendet?**
3. **Vad står det i andra källor?**

Att lära ut tekniker för den här typen av reflektioner har visat sig vara ett effektivt sätt att främja medborgarnas färdigheter i källkritik

och lateral läsning¹¹. Med lateral läsning avses att läsaren kontrollerar bakgrunden till informationen hen hittar på nätet (skribentens trovärdighet, fakta, statistik, källor osv.) på olika webbplatser och i olika källor, innan hen sätter sig närmare in i texten.

För att stötta lärare har Faktabaari gett ut undervisningsvideor¹² och -material¹³ till exempel om rekommendationsalgoritmer, de tre frågorna ovan, lateral läsning, omvända bildsökningar och om identifiering av fel, lögnar och skvaller.

Enligt rekommendationerna som tagits fram i det finländska forskningsprojektet Critical¹⁴ behöver varje ung person ha kritisk läsförmåga, dvs. en förmåga att söka information i olika medier och att bedöma olika texters tillförlitlighet och att jämföra informationen i dem. Enligt forskningsprojektet bör den kritiska och den grundläggande läsförmågan stödjas parallellt och barn med utmaningar med läsning bör erbjudas stöd så tidigt som möjligt. Critical-gruppen har också skapat verktyg för undervisningen om kritisk läsförmåga på nätet¹⁵.

Läroplansgrunderna och nya typer av läskunnighet

Den nya och bredare definitionen av digital informationslitteracitet – eller -läskunnighet – fokuserar inte enbart på datatekniskt kunnande, utan den omfattar alla aspekter av mångsidig kompetens i de finländska läroplansgrunderna.

Kunskapsbeskrivningar för digital kompetens för småbarnspedagogik och undervisning

Undervisnings- och kulturministeriet gav ut Kunskapsbeskrivningarna för digital kompetens¹⁶ år 2022, precis innan generativa AI-tjänster slog igenom i hela världen.

Syftet med kunskapsbeskrivningarna var att förtydliga grunderna för planen för småbarnspedagogik och läroplansgrunderna för förskoleundervisningen och för den grundläggande utbildningen, när det gäller digital kompetens och informations- och kommunikationstekniskt kunnande. Målet var att anordna utbildning

och småbarnspedagogik skulle uppdatera sina digitala strategier och planer samt sina läroplaner, så att de skulle vara i linje med den nationella referensramen för goda digitala kunskaper.

Digital kompetens och informations- och kommunikationstekniskt kunnande består av fyra huvudsakliga delområden:

4. praktiska färdigheter och egen produktion

5. säkert och ansvarsfullt agerande

6. informationshantering samt undersökande och kreativt arbete

7. interaktion.

Det ingår konkreta och åskådliga exempel i kunskapsbeskrivningarna, som uppmuntrar lärare till mångsidig användning av digitala miljöer i undervisningen. Med digitala miljöer avses alla

digitala lösningar, tjänster och verktyg – numera också AI-verktyg – samt all digital utrustning som man har tillgång till i undervisningen och i småbarnspedagogiken. Kunskapsbeskrivningarna ger en fingervisning

Nya tekniker ger eleverna möjlighet att ha rollen som aktiv aktör, innehållsproducent och forskare.

om vilken typ av digital kompetens som ska tas upp i undervisningen på olika utbildningsstadier. Tack vare ny teknik får elever i olika åldrar möjlighet att också prova rollen som aktiv aktör, innehållsproducent eller forskare, istället för att bara vara passiva användare på sociala medier.

För att kunna erbjuda mångsidig undervisning i digital kompetens behöver utbildningsanordnarna systematiskt fortbilda lärare, införskaffa ändamålsenlig utrustning och programvara och ge personalen tydliga anvisningar.

Undervisningen i digital kompetens är i slutändan på lärarnas ansvar på alla stadier och i alla skolämnen. Därför är det nödvändigt att anordna av undervisning och småbarnspedagogik ansvarar för att den utrustning och de tjänster lärarna har tillgång till följer EU:s allmänna dataskyddsförordning

och den nationella lagstiftningen. Därutöver bör lärarna systematiskt erbjudas fortbildning och de bör ges tydliga anvisningar för hur de på ett säkert och ansvarsfullt sätt jobbar i digitala undervisningsmiljöer.

PISA, digitalisering och mobilförbud

Enligt PISA 2022-undersökningarna

har 15-åringarnas kunskaper i matematik och naturvetenskaper samt deras läskunnighet försvagats i majoriteten av OECD-länderna, liksom i Finland.

Flera teorier har framförts om orsakerna till de försvagade resultaten. I den offentliga debatten har bland annat digitalisering, elever med invandrarbakgrund, pandemi, fenomenbaserat lärande, inklusion och självreglering nämnts som möjliga orsaker.

PISA-undersökningen ger dock en rätt begränsad bild av sambandet mellan användningen av digital utrustning och inlärningsresultaten

. Eleverna har upplevt att användningen av digital utrustning i viss mån känns störande i inlärningssituationer, men å andra sidan ser en måttlig användning av digitala resurser ut att ha en positiv inverkan på inlärningsresultaten.

Debatten om att förbjuda mobilerna i skolan har gått het, och mobilanvändningen har redan varit begränsad och kommer att begränsas ytterligare.

En begränsning av mobilanvändningen innebär dock inte att skolorna inte har i uppgift att erbjuda alla barn och unga möjligheten att lära sig viktig digital kompetens, ansvarsfull användning av sociala medier och en kritisk läsförmåga!

I diskussionen borde också tas upp huruvida lärarna i Finland har tillräckliga grundläggande kunskaper, så att de effektivt kan använda sig av digitala miljöer och AI-verktyg som stöd i undervisningen och hur den nödvändiga lärarfortbildningen kring digital kompetens kan ordnas och finansieras.

Utveckling av elevers digitala kompetens

Gjord med Napkin



Prova nya verktyg



Undervisa om
kritiskt tänkande



Diskutera säkerhet



Ta upp etiska
frågor



Uppmuntra till
kreativitet

Hur utveckla digital kompetens?

Genom att lära sig digitala färdigheter och till exempel testa och leka med olika typer av AI-verktyg kan man få upp ögonen för ett helt nytt sätt att klara av olika vardagssysslor. Med hjälp av AI kan det vara roligt att göra exempelvis rese- eller semesterplaner, översätta texter enkelt från ett språk till ett annat och skapa fina bilder och videor. AI-verktyg kan användas som effektiva hjälpmedel när lärare väljer material och skapar idéer för och planerar sin undervisning. Med sin egen digitala kompetens kan lärare inspirera elever att använda sina egna smarttelefoner och surfplattor till annat än spel och sociala medier.

Läraren har en central roll i utvecklingen av elevernas digitala kompetens. Det är nödvändigt med digital kompetens i arbetslivet, och den hjälper eleverna också att aktivt delta i samhällsdebatten. Den nya tekniken kan dessutom göra lärandet både intressantare och interaktivare.

AI-kompetens

AI-kompetens är en viktig ny digital färdighet. Med AI-kompetens avses förståelsen av hur AI fungerar och tillämpas och vilka principer och begrepp som gäller för AI samt medvetenheten om begränsningar, effekter och etiska aspekter användningen av AI har.

Den senaste definitionen av AI-kompetens, som för närvarande är den mest heltäckande, grundar sig på forskaren Aran Bozkurts¹⁶ definition från hösten 2024.

Tack vare AI-kompetensen kan lärare och elever skickligt dra nytta av AI-baserade tekniker och vara medvetna AI-användare. En viktig del av det är också medvetenheten om hur undervisningen och uppgiftsbeskrivningarna kan ändras så att eleverna kan använda AI som ett verktyg som främjar lärandet och förmågan till kritiskt tänkande.

Amerikanska TechAI har gett ut en hel del material om AI för lärare. Tabellen intill ger en översikt över potentiella fördelar och risker med AI i utbildningen.

Det är nödvändigt med digital kompetens i arbetslivet, och den hjälper eleverna också att aktivt delta i samhällsdebatten.

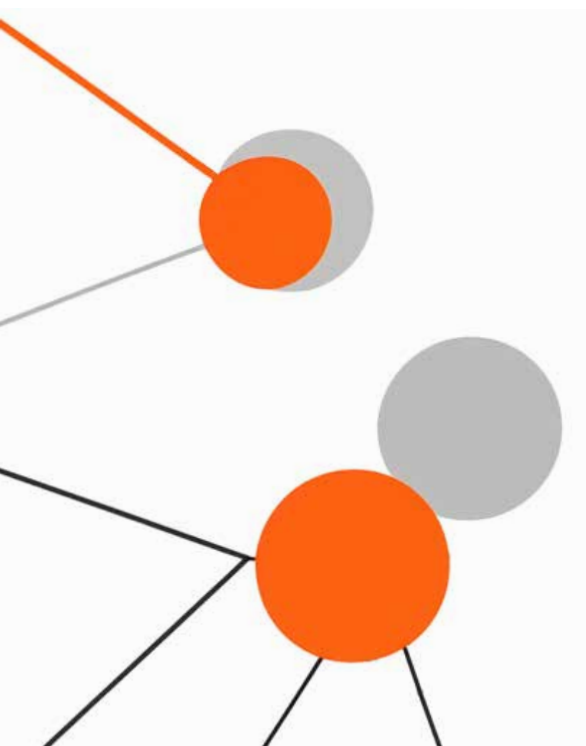
Några praktiska tips för utveckling av digital kompetens

1. **Prova nya verktyg:** Var nyfiken och prova olika verktyg och program i undervisningen. Det kan inspirera också eleverna.
2. **Undervisa om kritiskt tänkande:** Uppmuntra eleverna att kritiskt granska den information de tar del av. Diskussioner om var informationen härstammar ifrån hjälper eleverna att förstå informationens tillförlitlighet.

3. **Tala om säkerhetsfrågor:** Lär eleverna hur de kan skydda egna data på nätet. Det är en viktig del av digitalt välbefinnande.
4. **Diskutera etiska frågor:** Ta upp upphovsrätten och andra juridiska frågor, så att eleverna förstår vikten av att agera ansvarsfullt.
5. **Uppmuntra till kreativitet:** Låt eleverna skapa sina egna digitala projekt, till exempel bloggtexter eller videor.

-
- 1 DigComp 2.2. (2012): JRC <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>
 - 2 Global Risks 2024: Disinformation tops global risks (2024) World Economic Forum <https://www.weforum.org/press/2024/01/global-risks-report-2024-press-release/>
 - 3 European Parliament Think Tank (hämtad: 12.12.2024) Mis- and disinformation on social media and related risks to election integrity [https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA\(2024\)767150](https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2024)767150)

- 4 Ledare: Britannian mellakat ovat varoittava esimerkki sosiaalisessa mediassa levitettyjen valheiden vaikutuksista. (4.8.2024): Iltasanomat <https://www.is.fi/paakirjoitus/art-200010607694.html>
- 5 <https://educritical.fi/fi/>
- 6 Nordic Observatory for Digital Media and Information disorder NORDIS (hämtad: 12.12.2024) <https://www.nordishub.eu/>
- 7 European Digital Media Observatory EDMO (hämtad: 12.12.2024) <https://edmo.eu/>
- 8 Digitaalinen informaatiolukutaito-opas (2020): Faktabaari. <https://faktabaari.fi/dil/>
- 9 Digitaalinen informaatiolukutaito - mitä se on (2022): Faktabaari <https://faktabaari.fi/dil/1-digitaalinen-informaatiolukutaito-dil-mita-se-on/>
- 10 Verkkolukutaidot ja strategiat (2022): Faktabaari. <https://faktabaari.fi/dil/6-verkkolukutaidot-ja-strategiat/>
- 11 Osborne, J., Pimentel, D., Alberts, B., Allchin, D., Barzilai, S., Bergstrom, C., Coffey, J., Donovan, B., Kivinen, K., Kozyreva, A., & Wineburg, S. (2022): Science Education in an Age of Misinformation. Stanford University, <https://sciedandmisinfo.stanford.edu/>
- 12 Faktabaari (2024): Videomaterial <https://www.faktabaari.fi/dil/>
- 13 Faktabaari (2024): Undervisningsmaterial <https://www.faktabaari.fi/edu/opinimateriaalit/>
- 14 Critical (2023-2024): Utvalda artiklar och övrigt material: <https://educritical.fi/fi/tuotokset/selected-articles/>
- 15 Critical (2023): Työkaluja kriittisen nettilukutaidon opettamiseen <https://aoe.fi/#/kokoelma/294>
- 16 Bozkurt, A. (2024) 'Why Generative AI Literacy, Why Now and Why it Matters in the Educational Landscape? Kings, Queens and GenAI Dragons', Open Praxis, 16(3), s. 283-290. Tillgänglig på adressen: <https://doi.org/10.55982/openpraxis.16.3.739>



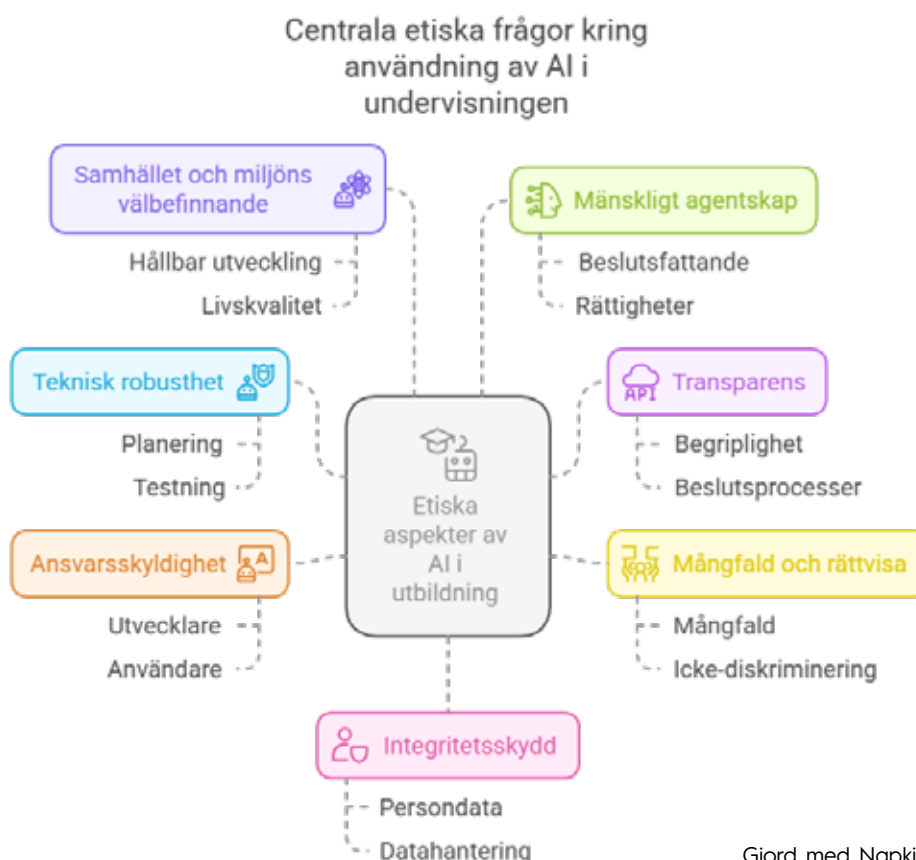
6. Etiska utmaningar för utbildningsanordnare och lärare vid användning av AI i utbildning

KARI KIVINEN, FAKTABAAARI

Den nya tekniken bör främja målen för undervisningen och inte vara en extra belastning.

ANVÄNDNINGEN AV AI-VERKTYG i undervisningen förutsätter noggrann planering, öppen kommunikation och kontinuerlig utveckling av läraren. På så sätt är det möjligt att se till att tekniken tjänar undervisningens syften på ett tryggt och pedagogiskt sätt. Undervisnings- och bedömningsplanerna bör omarbetas så att valet och användningen av AI-verktyg görs på ett ändamålsenligt sätt. Den nya tekniken bör främja målen för undervisningen och inte vara en extra belastning. När AI-baserade verktyg används rätt har de med sin snabba utveckling potential att effektivisera skolans vardagsrutiner, vara en källa till undervisningsmaterial och främja lärandet.

Det är på utbildningsanordnarens ansvar att bedöma hur användningen av AI-verktyg utformas i praktiken, så att informationssäkerheten, dataskyddet och upphovsrätten beaktas och



Gjord med Napkin

elevernas och personalens rätt till sina egna data och sin integritet respekteras. Samtidigt ska man se till att eleverna inte blir föremål för till exempel politisk eller ideologisk påverkan, kommersiella intressen eller marknadsföring och att de inte får ta del av skadligt eller stötande innehåll.

Utbildningsanordnaren ska därutöver utvärdera huruvida AI-verktygen uppfyller EU:s kriterier för AI-system och följer de tillhörande förpliktelseerna och villkoren.

Särskilt viktigt är det att beakta de etiska utmaningarna vid användning av AI-verktyg, stötta både lärare och elever och försäkra sig om verktygens tillgänglighet. Användningen av AI-verktyg bör stödja lärarens arbete, effektivisera det administrativa arbetet och främja alla elevers lärande i enlighet med målen i läroplanen.

Centrala etiska frågor kring användning av AI i undervisningen

EU-kommissionen¹ har gett ut en lista, som sammanställts av en expertgrupp, över etiska riktlinjer, som med tanke på användningen av AI är centrala för utbildningsanordnare och lärare. Riktlinjerna fokuserar på mänskligt agentskap, transparens, mångfald, samhällets välbefinnande, integritet, teknisk robusthet och ansvarsskyldighet.

Etiska krav på utbildningsanordnare

När AI-baserade system, verktyg och program införskaffas och används är det viktigt att de följer den lokala lagstiftningen och är i linje med läroplanerna och målen för lärandet. Lärarnas representanter får gärna erbjudas möjlighet att vara involverade i upphandlingen av systemen och beslutsprocessen kring den.

AI-system som införskaffas för undervisningsändamål bör vara etiska, tillförlitliga, säkra och människocentrerade. Med människocentrerade system avses bland annat att AI-systemen inte bör ges egen beslutanderätt utan mänsklig tillsyn till exempel vid elevantagning eller bedömning.

AI-system som införskaffas för undervisningsändamål bör vara etiska, tillförlitliga, säkra och människocentrerade.

Lagar och förordningar

Enligt EU:s AI-förordning bedöms användningen av AI i utbildningen ha hög risk, eftersom systemet kan utsätta användarens grundläggande rättigheter för risker som berör bland annat integritet, icke-diskriminering och yttrandefrihet. EU-förordningarna skyddar speciellt minderåriga barns digitala rättigheter.

Enligt artikel fyra i AI-förordningen ska leverantörer och tillhandahållare av AI-system vidta åtgärder "för att i största möjliga mån säkerställa att deras personal och andra personer som för deras räkning arbetar med drift och användning av AI-system har tillräcklig AI-kunnighet, med beaktande av deras tekniska kunskaper, erfarenhet och utbildning samt det sammanhang i vilket AI-systemen ska användas, och med hänsyn till de personer eller grupper av personer på vilka AI-systemen ska användas."

AI-systemen ska följa EU:s dataskyddsförordning. Tillgången till data ska vara förbehållen endast de personer som behöver komma åt dem, och användarna bör veta vad som händer med deras data och hur de används. Lärare och skolledare bör vara medvetna om hur de kan rapportera om problem som gäller integritets- eller dataskyddet.

Med tanke på eventuella personuppgiftsincidenter bör det finnas tillräckliga säkerhetsåtgärder som vid behov kan vidtas och övervakningsmekanismer för behandling av information (datainsamling, lagring, uppgiftsminimering). Personalen bör erbjudas ändamålsenlig utbildning så att de kan följa säkerhetsanvisningarna.

Potentiella risker med användningen av AI-verktyg, såsom farorna med ökat fokus på kommersiella intressen eller en ökad digital klyfta, bör övervakas och granskas kritiskt. På så sätt är det möjligt att säkerställa att AI-verktyg inte leder till ökad orättvisa eller diskriminering.

Tillgänglighet

AI-systemen bör vara tillgängliga för alla elever oavsett kön, färdighetsnivå eller andra personliga egenskaper. Särskild uppmärksamhet bör ägnas åt elever med särskilda behov.

Mänskligt agentskap	I utvecklingen och användningen av AI är det viktigt att beakta människans roll och påverkan. Det är viktigt att se till att människor förblir i centrum för beslutsfattandet och att deras rättigheter och behov beaktas.
Transparens	De operativa principerna och beslutsprocesserna för AI ska vara tydliga och lätta att förstå. Användarna och andra berörda grupper ska kunna förstå hur och varför AI fattar vissa beslut.
Mångfald, icke-diskriminering och rättvisa	I utvecklingen av AI är det viktigt att främja mångfald och förebygga diskriminering. AI-system ska vara rättvisa, och de får inte förstärka befintliga fördomar eller diskriminerande praxis.
Samhällets och miljöns välbefinnande	Användningen av AI ska främja samhällligt och ekologiskt välbefinnande, och AI-lösningar ska stödja hållbar utveckling och förbättra människors livskvalitet.
Integritet och datahantering	Integritetsskydd utgör en central fråga vid användning av AI. Det är viktigt att säkerställa att personuppgifter hanteras säkert och att användarna har tillgång till sina egna data.
Teknisk robusthet	AI-system ska fungera tillförlitligt och förutsägbart. Detta förutsätter noggrann planering och testning, så att risken för fel och misstag kan minimeras.
Ansvarsskyldighet	De som utvecklar och använder AI-system har ett ansvar för effekterna av dem. Det är viktigt att avgöra vem som bär ansvaret för de beslut som fattas av AI och konsekvenserna besluten kan ha.

Sammanfattning av de etiska riktlinjerna för användning av AI i undervisningen

Vid val av AI-verktyg är det viktigt att ta hänsyn till att de inte ska öka den digitala klyftan som redan existerar. Alla elever ska ha möjlighet att dra nytta av de möjligheter som AI erbjuder. Det är därför viktigt att se till att både lärare och elever har tillgång till samma AI-verktyg. Detta gör det också lättare att samarbeta, sprida information, ge vägledning och ordna lärarfortbildning.

Anvisningar

Både lärare och elever bör erbjudas tydliga anvisningar om hur de kan använda AI-verktyg.

Under våren 2025 ger Utbildningsstyrelsen och Undervisnings- och kulturministeriet ut rekommendationer och förordningar för användningen av AI i småbarnspedagogik, grundläggande utbildning, fritt bildningsarbete och andra stadiets utbildning. Syftet med stödmaterialet "AI i utbildning – lagstiftning och rekommendationer" är att stötta anordnare av småbarnspedagogik och utbildning i användningen av AI i undervisning och studier. Stödmaterialet kommer att ges ut på svenska, finska och engelska.

Etiska krav på lärare

AI innebär nya möjligheter för lärare, men samtidigt skapar den nya etiska utmaningar. Lärarna bör ha en förståelse av hur AI-tekniken

fungerar och vara medvetna om möjligheterna, begränsningarna och riskerna med AI-verktyg. Därutöver bör lärarna se till att AI används etiskt och på ett sådant sätt att den stöder lärandet.

Lärarnas representanter bör ha möjlighet att vara involverade när det fattas beslut om anskaffning och användning av AI-verktyg. Lärarna bör dessutom ha möjlighet att ge respons på verktygen och hur väl de fungerar. I undervisningen ska personalen använda endast sådana verktyg som godkänts av utbildningsanordnaren i enlighet med givna anvisningar.

Kännedom om och en ansvarsfull användning av AI-verktyg förutsätter fortbildning

För att AI-verktyg ska kunna tas i bruk är det viktigt att lärare får fortbildning och kan i viss mån sätta sig in i hur till exempel stora språkmodeller har tränats, hur de fungerar och varför användningen av verktygen alltid präglas av vissa osäkerhetsfaktorer, vilket gör att det som verktygen producerat alltid måste granskas.

Utbildningsanordnarna bör se till att lärarna har goda grundläggande kunskaper om verktygen och programvaran som används och att de ges tydliga anvisningar om hur verktygen fungerar, vad de kan användas till och vilka begränsningar de har. Utbildningsanordnarens

dataskyddsbud bör ge lärare information om vilka uppgifter systemen samlar in om eleverna, hur de insamlade uppgifterna behandlas, vad de används till och var de lagras. I slutändan är det lärarens ansvar att se till att verktygen används på ett ändamålsenligt sätt och så att de stöder undervisningen och lärandet.

Lärarna bör själva vara medvetna om vilka verktyg som baserar sig på AI-applikationer, hur de används och vilka potentiella risker som är förknippade med användningen av verktygen. Lärarna har också i uppgift att förklara för

eleverna hur AI-verktygen på skolan används och vägleda dem att använda verktygen på ett säkert sätt.

AI-systemen bör vara tillgängliga för alla elever oavsett kön, färdighetsnivå eller andra personliga egenskaper.

Pedagogik

AI-verktyg ska användas pedagogiskt utgående från barnens ålder och

i situationer där de tillför mervärde i lärandet. Lärarna bör vara medvetna om vilka verktyg som lämpar sig för olika åldersgrupper och hur de kan integreras i undervisningen så att de stöder lärandet.

Undervisningen ska planeras så att eleverna har lika möjlighet att använda AI-verktyg, och eventuella brister i tillgängligheten ska rapporteras till utbildningsanordnaren. På så sätt är det möjligt att se till att alla elever får tillgång till AI-verktyg oavsett bakgrund. AI har visat sig vara ett ytterst användbart verktyg i specialundervisningen, i och med att den erbjuder flera möjligheter till differentiering, personlig anpassning av undervisningsmaterial och användning av AI som mentor i lärandet.

I elevbedömningen ska man aldrig enbart och utan mänsklig kontroll förlita sig på AI-verktyg. Läraren bör se till att AI inte ersätter lärarens roll utan att AI fungerar som ett hjälpmedel som stöder undervisningen.

Det är viktigt att läraren informerar eleverna om hur och varför AI-verktygen används. Öppenheten

hjälp eleverna att förstå vilken roll AI har i undervisningen, vilket i sin tur motarbetar farhågor, rädslor och oärlig användning av AI.

Övervakning

AI-verktyg och -system utvecklas väldigt snabbt. Deras funktionalitet är något som bör utvärderas kontinuerligt, och personalen bör ha möjlighet att ge respons utifrån sina erfarenheter till utbildningsanordnaren. Med hjälp av responsen är det möjligt att se till att AI-verktygen utvecklas och förbättras kontinuerligt.

Lärare som aktivt deltar i användningen och utvecklingen av AI-verktyg kan skapa nya möjligheter till lärande och förbättra inlärningsresultaten. Det förutsätter i sin tur kontinuerlig dialog, fortbildning och övervakning, så att man kan säkerställa att AI främjar målen för utbildningen och inte utgör ett hot mot dem.

Lärare bör förse med tillräckliga grundläggande kunskaper så att de kan identifiera och rapportera eventuella personuppgiftsincidenter. Lärare bör också vara medvetna om att AI-verktyg kan missbrukas, till exempel för akademiskt fusk, och de bör kunna identifiera och minimera potentiella risker för missbruk.

Observera särskilt

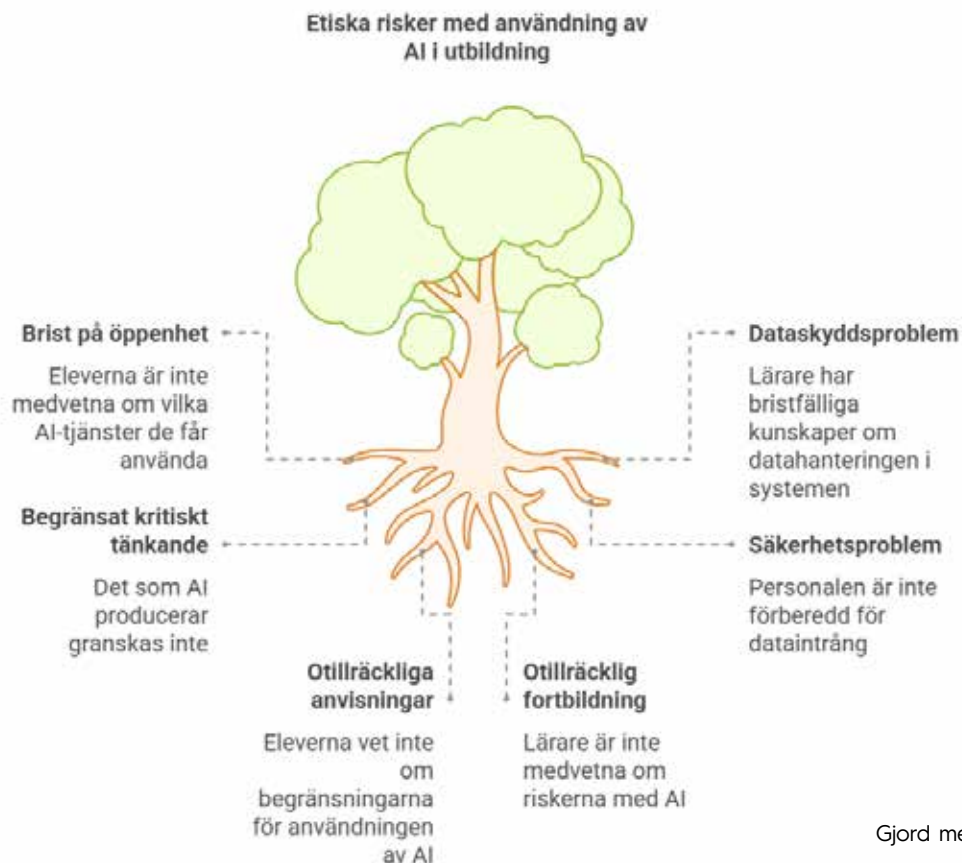
AI håller på att förändra vårt utbildningssystem genom att erbjuda mer personfokuserat lärande och individualiserad mentorhjälp. Samtidigt för AI med sig stora utmaningar.

Något som speciellt ger upphov till bekymmer är **datamissbruk**. Det är viktigt att se till att elevernas och lärarnas personuppgifter hanteras på ett ansvarsfullt sätt och inte används som inmatningsdata eller promptar i GenAI-verktyg.

En annan betydande risk har att göra med **informationens tillförlitlighet**. Informationen som genereras av

AI är inte alltid helt tillförlitlig, och eleverna måste lära sig att kritiskt granska det som

AI har visat sig vara ett mycket användbart verktyg i specialundervisningen.



AI producerat, till exempel genom att verifiera källornas tillförlitlighet.

Dessutom finns det en risk för att **ansvaret för lärandet läggs på AI**. Om eleverna förlitar sig för mycket på AI-baserade lösningar utvecklar de inte längre sin egen problemlösningsförmåga eller fördjupar sig i olika ämnen på egen hand. Det kan i sin tur leda till att lärandet i själva verket är bara ytligt.

Kommersialiseringen av utbildning är också en oroande trend. De som utvecklar AI-verktyg är för intresserade av ekonomiska fördelar, vilket kan leda till att innehållet i utbildningen är mer anpassat efter kommersiella än pedagogiska intressen.

Hanteringen av etiska risker förutsätter att lärare är vaksamma. Lärare, skolor och utbildningsanordnare bör tillsammans fastställa tydliga spelregler för användningen av AI i utbildningen. Det förutsätter en kontinuerlig diskussion, fortbildning och övervakning, så att användningen av AI tjänar målen för utbildningen. Ett mycket användbart verktyg för det är listan som utarbetats av en expertgrupp vid Europeiska kommissionen och som tar upp centrala etiska

och övriga frågor, som är viktiga med tanke på användningen av AI. Bilagan till AI-guiden² som Faktabaari gav ut 2024 ger en sammanfattning av de mest centrala frågorna.

Tarmo Toikkanen från Sitra har på ett briljant sätt sammanfattat varför det som produceras av AI behöver granskas och övervakas av en människa³.

Artificiell intelligens som producerar text ...

- **förstår inte vad ord betyder**
- **har inga åsikter, värderar inte**
- **tänker inte, ger inga förutsägelser**
- **känner inte, bryr sig inte**
- **tar inget ansvar**
- **hallucinerar.**

¹ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

² AI-guiden från Faktabaari (2024) <https://www.faktabaari.fi/dil/ai/>

³ Tarmo Toikkanen (hämtad: 31.12.2024) Tekoäly opetuksessa <https://drive.google.com/file/d/1dFhZjEWXRJn3OYP4iC-zTQJuPd5Qz-u7/view>

7. Referensramar för AI-kompetens för lärare

KARI KIVINEN, FAKTABAAARI

UNDER DET GÅNGNA ÅRET har det getts ut flera referensramar för AI-kompetens både för lärare och för elever.

I den här artikeln presenteras i korthet två olika referensramar för lärare. Unesco gav ut sin egen referensram för både lärare och elever hösten 2024. En forskargrupp (Filo m.fl., 2024) publicerade en referensram i november 2024, som gruppen utarbetat i samarbete med lärare och elever och som ser ut att vara mycket användbar.

I Finland håller Undervisnings- och kulturministeriet och Utbildningsstyrelsen på att utarbeta rekommendationer för användning av AI i småbarnspedagogik och utbildning. De ges ut under våren 2025.

Unescos referensram för AI för lärare

Unesco presenterade sin referensram för lärares AI-kompetens¹ i september 2024. I den är det fokus på lärarnas livslånga professionella utveckling, och den innehåller fem centrala områden.

Syftet med referensramen är att säkerställa att lärare har förutsättningar att använda AI på ett ansvarsfullt och effektivt sätt, samtidigt som potentiella risker för eleverna och samhället minimeras.

- **Det människocentrerade angreppssättet beskriver de värderingar och kritiska attityder gentemot interaktionen mellan människor och AI, som lärare behöver ha. Lärare uppmuntras att alltid se till att mänskliga rättigheter och behov relaterade till människors välbefinnande är i fokus vid användning av AI i**

Centrala områden	Kompetensnivå		
	Grundnivå	Fördjupad nivå	Avancerad nivå
Människocentrerat angreppssätt	Mänskligt agentskap	Människans ansvar	Socialt ansvar
Etik och AI	Etiska principer	Säker och ansvarsfull användning	Utarbetande av etiska riktlinjer tillsammans
Grunderna för och tillämpningar av AI	Grundläggande AI-tekniker och -tillämpningar	Tillämpnings-färdigheter	Skapande processer med stöd av AI
AI-pedagogik	AI-assisterad undervisning	Integrering av AI-pedagogik	AI-baserad pedagogisk förändring
AI och professionell utveckling	Livslång professionell utveckling med hjälp av AI	Effektivare lärande inom organisationen med hjälp av AI	AI som stöd i professionell utveckling

utbildningen. Lärare uppmuntras att utveckla olika sätt att kritiskt utvärdera fördelarna och riskerna med AI, att säkerställa människans eget engagemang och ansvar vid användning av AI och att förstå på vilket sätt AI påverkar samhället och hur det är att vara medborgare i AI-eran.

- **AI-etik:** Lärare är medvetna om de grundläggande etiska principerna för AI och förstår deras betydelse. De är införstådda med AI:s människodrivna natur och människans centrala roll i hur utvecklingen av AI utformas och utifrån vilka premisser. AI-etiken innefattar alla centrala etiska värderingar, principer, förordningar, institutionella lagar och praktiska etiska regler, som lärare bör förstå och som baserar sig på den snabbt växande kunskapen om AI-etik och dess inverkan på utbildning.
- **Grunderna för och tillämpningar av AI** definierar de grundläggande kunskaper och färdigheter som lärare behöver för att kunna ge stöd när det gäller vilka AI-verktyg som kan användas, hur de tillämpas och hur de kreativt kan anpassas för att skapa elevcentrerade, AI-assisterade undervisnings- och lärmiljöer. Lärare förväntas förstå hur AI definieras och ha grundläggande kunskaper om hur AI fungerar och vilka huvudkategorierna av AI-teknik är. Utöver det bör de ha nödvändiga färdigheter för att kunna bedöma hur AI-verktyg kan tillämpas och vilka begränsningar de har.
- **AI-pedagogik:** Lärare ska kunna identifiera och utvärdera de pedagogiska fördelarna med AI-verktyg och välja ändamålsenliga verktyg, som kan stödja undervisningsplaneringen, pedagogiken och bedömningen. Lärare måste utveckla sin förmåga att kritiskt bedöma när och på vilket sätt AI kan användas i undervisningen och lärandet på ett etiskt och människocentrerat sätt samt utveckla

och tillämpa engagerande AI-assisterad praxis för undervisning och lärande.

- **AI och professionell utveckling:** Lärare ska vara medvetna om på vilket sätt AI kan stödja deras kontinuerliga professionella utveckling och vara motiverade att använda AI i sitt livslånga professionella lärande. På grund av den snabba utvecklingen av AI behöver lärare vägledning i hur de kan fortsätta sin professionella utveckling i utbildningsmiljöer, där interaktionen mellan människor och AI ökar.

Delområdena i referensramen är sammanflätade och kompletterar varandra. De är inte oberoende av varandra. En framgångsrik undervisning (med eller utan AI) förutsätter ett heltäckande synsätt, som kombinerar olika kompetensområden.

Till exempel påverkas en lärares förmåga att tillämpa AI-pedagogik av hens förståelse av grunderna för AI, hens medvetenhet om anvisningarna som gäller för AI och hens engagemang för kontinuerlig professionell utveckling. På motsvarande sätt påverkas lärarens förmåga att hantera de etiska problemen med AI av hens förståelse av grunderna för AI och av hens erfarenheter av att använda AI i undervisningen.

Kompetens på ett område kan utveckla kompetensen på ett annat. En djupare förståelse av grunderna för AI kan förbättra lärarens förmåga att tillämpa de pedagogiska och etiska principerna kring AI, och kontinuerlig professionell utveckling ökar lärarens förståelse av alla dessa aspekter.

Referensramen betonar att AI-verktyg ska komplettera, inte ersätta, lärarnas avgörande roller och ansvarsområden inom utbildning.

Referensram för AI-kompetens för lärare och studerande

I en annan, relativt ny referensram för AI-kompetens för lärare och studerande, som också utvecklats i samarbete med lärare och studerande, identifieras fyra nyckelkompetenser²:

1. **kännedom om vad AI-mekanismer är och hur de fungerar**
2. **effektiv och kompetent användning av AI**
3. **hantering av AI på ett förutseende sätt, som också skapar mervärde**
4. **etisk användning av AI.**

Referensramen beskriver också de värderingar, attityder och kunskaper som behövs för att hantera AI i utbildningen, så att lärare och studerande kan förberedas för en värld där AI figurerar närmast överallt.

Enligt den föreslagna referensramen är kännedom om AI-mekanismer en ytterst viktig färdighet för både lärare och studerande, eftersom den lägger grunden för hur man navigerar bland och kritiskt hanterar AI-verktyg i olika kontexter.

Genom att förstå principerna för hur AI fungerar är det möjligt att bättre bedöma vad slags potential, begränsningar och etiska konsekvenser AI-verktygen har i vår vardag.

En effektiv och kunskapsbaserad användning av AI förutsätter också en förståelse av med vilken typ av promptar man kan få fram önskade resultat och en förmåga att välja ut rätt AI-verktyg för en specifik uppgift. Den tredje kompetensen går ut på att kritiskt granska de resultat som producerats av AI-verktyg, att bedöma deras exakthet, att identifiera eventuella felaktigheter och att förstå

vilka etiska konsekvenser användningen av AI har i vår vardag.

Med hantering av AI avses en etisk och rättvis användning av AI-verktyg så att de erbjuder människor stöd utan att ersätta deras insats. Ett sådant angreppssätt förutsätter en djup förståelse av möjligheterna med och begränsningarna av AI. Hanteringen handlar också om att AI används aktivt och på ett kreativt sätt, så att den blir en kraftkälla för människan och inte bara används passivt. Det här förutsätter en förmåga att kreativt tillämpa AI vid problemlösning och att skapa mervärde, som i dagens värld är en stor fördel. På så sätt kan människor använda sig av AI för att ha en positiv påverkan på sitt liv och sin omgivning. Samtidigt aktualiseras behovet av att förstå samarbetet mellan människor och maskiner och att kritiskt granska det som AI producerar.

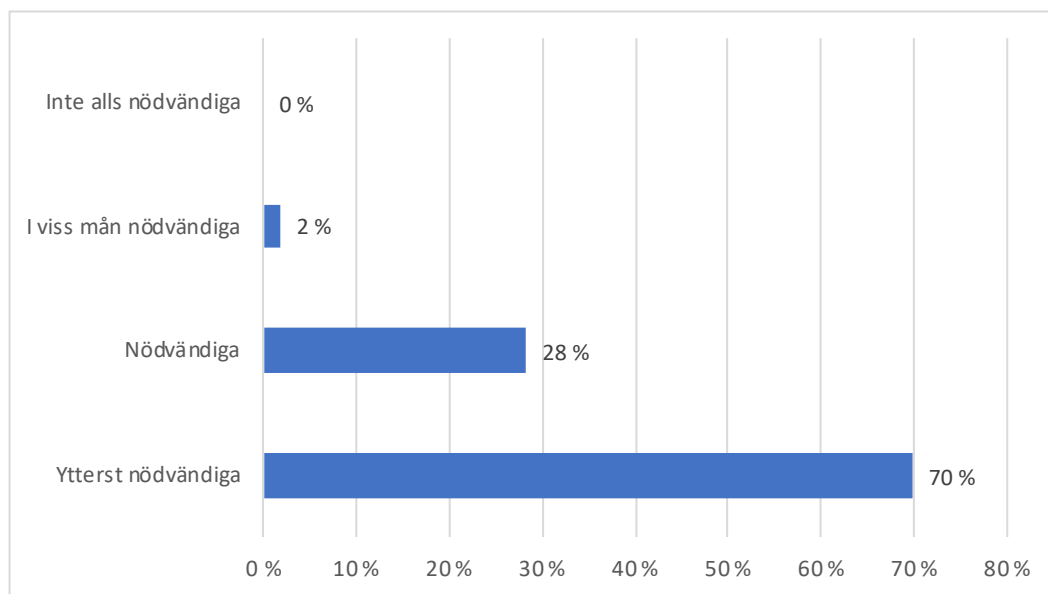
Syftet med en etisk användning av AI är att hjälpa elever och studerande att göra goda val. Det är viktigt att poängtera för elever och studerande att användningen och utvecklingen av AI-verktyg bör präglas av rättvisa, integritet, moral och ansvarsfullt beslutsfattande. En etisk användning av AI innebär att AI-systemens fördomar och missuppfattningar identifieras och korrigeras, att ojämlikheten minskas och att viktiga mänskliga värderingar och samhällets välbefinnande kan främjas. En etisk användning förutsätter en noggrann utvärdering av de

Referensram för AI-kompetens



Gjord med Napkin

Hur nödvändiga anser du att Utbildningsstyrelsens rekommendationer för användning av artificiell intelligens är? (N=106)



långsiktiga effekterna av AI, såsom försämring av människans egna färdigheter, diskriminering eller manipulering av människans egna uppfattningar. Samtidigt betonas individens ansvar och aktiva roll för att förebygga dessa utmaningar.

Finländska AI-rekommendationer

I oktober 2024 sände Undervisnings- och kulturministeriet³ och Utbildningsstyrelsen första delen av AI-rekommendationerna på remiss. Följande centrala delar ingår i rekommendationerna:

- En förutsättning för ansvarsfull och trygg användning av AI är att förstå hur AI fungerar och vilka effekter den kan ha.
- Det förutsätter i sin tur en kritisk utvärdering av möjliga användningsområden för och begränsningar av AI samt djupgående reflektioner över för- och nackdelarna med AI.
- Innan AI börjar användas är det viktigt att se till att användarna har en tillräcklig förståelse av AI och en förmåga att tolka och kritiskt utvärdera den information AI genererar.
- En central utmaning i fostran och utbildning är att förse elever och studerande med

färdigheter i bland annat kritiskt tänkande och problemlösning, källkritik och kreativitet, när det gäller användningen av AI.

- I en digitaliserad och globaliserad värld är dessa färdigheter en förutsättning för ett aktivt samhällsengagemang samt för yttrandefrihet och demokratins fortlevnad.
- Centrala teman inom fostran och utbildning blir då AI-etik och etisk användning av teknik.

Den slutliga versionen av rekommendationerna och stödmaterialet publiceras och kan användas som stöd i småbarnspedagogik, undervisning, utbildning och lärande år 2025.

I samband med en AI-utbildning i regi av regionförvaltningsverket fick de deltagande lärarna svara på frågan hur nödvändiga de anser att de nya rekommendationerna är. Svaret var tydligt. Enligt 98 % av lärarna är rekommendationerna nödvändiga eller ytterst nödvändiga.

¹ UNESCO AI Competency Framework for Teachers (2024): <https://unesdoc.unesco.org/ark:/48223/pf0000391104>

² Filo, Y., Robin, E. & Mor, Y. An Artificial Intelligence Competency Framework for Teachers and Students: Co-created With Teachers. European Journal of Open, Distance and E-Learning, 2024, Sciend, vol. 26 no. 1, s. 93-106. <https://doi.org/10.2478/eurodl-2024-0012>

³ Tekoäly varhaiskasvatuksessa ja koulutuksessa – lainsäädäntö ja suosituksset (hämtad: 17.10.2024 lausuntopalvelu.fi) <https://www.lausuntopalvelu.fi/FI/Proposal/Participation?proposalId=a0d6af03-67e1-4ec7-9269-fab75bb05807>

8. Bedömningsskala för användning av AI

KARI KIVINEN, FAKTABAARI

MÅNGA LÄRARE HAR sagt sig vara oroliga för att elever eventuellt använder sig av GenAI-verktyg för sina läxor, vilket kan ses som fusk, plagiering eller till och med akademiskt bedrägeri. Därför bör eleverna uppmuntras till att öppet tala om i hur hög grad de använt sig av AI i sina uppgifter.

Om något som producerats av AI inkluderas i en undersökning, en artikel eller i undervisningen är det bra att hänvisa till och citera AI på samma sätt som övriga källor. Jyväskylän universitet¹ ger till exempel följande instruktioner för dokumentation:

”Om du använder AI-verktyg i dina studier eller i din forskning ska du dokumentera användningen noggrant. Dokumentationen är en förutsättning för att du kan rapportera användningen av AI sanningsenligt och i enlighet med god vetenskaplig praxis.

Enligt riktlinjerna vid Jyväskylän universitet ska studerandena rapportera om användningen av generativ AI i kursuppgifter och i lärdomsprov.”

Generativa AI-verktyg är gratis tillgängliga för alla elever och studerande. Därför bör lärare fundera över vilken typ av uppgifter de ger eleverna och vad de vill bedöma. Det är inte längre ändamålsenligt att endast jobba med uppgifter som går ut på att eleverna söker upp information eller skriver essäer. Med hjälp av sökmotorer och GenAI kan eleverna lätt leta fram vilken som helst typ av information och producera text om vilket som helst tema.

Lärarna bör i förväg diskutera med eleverna hur artificiell intelligens kan användas och i vilka övningsuppgifter, som läraren sedan bedömer. Syftet med AI-bedömningsskalan, som tagits fram av Furzen & AI (AI Assessment Scale, AIAS)² är att göra kommunikationen mellan lärare och elever eller studerande tydligare. Bedömningsskalan kan användas för att klargöra för både elever eller studerande och personal vad det är som räknas som ändamålsenlig eller olämplig användning av AI i en uppgift.

Användningen av AI kan förbjudas till exempel i förberedande prov inför studentskrivningar. Å andra sidan finns det många uppgifter där AI kan användas som ett användbart idé- eller stödverktyg. AI kan med fördel användas i uppgifter som förutsätter problemlösnings- och innovationsförmåga.

Olika nivåer för användning av AI

Om användningen av AI begränsas och övervakas ska begränsningarna vara tydligt formulerade och alla bör ha tillgång till tydliga anvisningar. Furzes tabell har också kompletterats med tydliga elevanvisningar för olika situationer.

Anvisningar om användning av AI till elever

Om det är viktigt att eleverna visar prov på kunskaper och färdigheter helt utan AI-hjälp, måste de göra det under någon form av övervakning. Enligt Furze beror det på två saker³. "För det första finns det ingen garanti för att en elev som har tillgång till någon form av teknisk utrustning (mobil enhet, bärbar dator, surfplatta, Meta Ray Bans e.d.) inte använder sig av AI-hjälp. Om vi, för det andra, försöker använda detekteringsverktyg eller ta till andra tekniska lösningar finns det en risk för likabehandlingsproblem, eftersom en del elever kan ha tillgång till mer avancerade, avgiftsbelagda AI-produkter (och som kommer undan tack vare dem), medan andra har bara begränsad tillgång till gratis verktyg (och som därför sannolikt ökar fast)."

Många utbildningsinstitutioner har tagit i bruk AI-detekteringsverktyg som ska övervaka användningen av AI. Tyvärr har de visat sig vara

opålitliga och ofta ge falska positiva resultat⁴.

Risken med att förlita sig på sådana verktyg är att oskyldiga elever orättvist blir anklagade, medan verktygen eventuellt inte alls upptäcker sådana elevers skickliga oärliga användning av AI-verktyg, som avsiktligt undviker att bli upptäckta.

Om eleverna ges tillstånd att använda AI är det bäst att också be dem dokumentera sin arbetsprocess och till exempel presentera vilka promptar eller inmatningsdata de använt sig av, vilka resultat de fått och hur de har jobbat vidare på och redigerat det AI-genererade materialet. En sådan metod inte bara hindrar oärlig användning av AI utan den hjälper också att utveckla elevernas AI-färdigheter. En inblick i arbetsprocessen ger läraren också en uppfattning om på vilket sätt olika AI-verktyg eventuellt kan användas i uppgifterna som de ger eleverna⁵.

¹ Jyväskylä universitet (hämtad 15.12.2024) <https://www.jyu.fi/fi/opiskelijalle/kandi-ja-maisteriopiskelijan-ohjeet/tiedonhankinta-ja-aineistonhallinta/tekoaly-tiedonhankinnassa>

² The updated AI assessment scale (hämtad 12.12.2024) <https://leonfurze.com/2024/08/28/updating-the-ai-assessment-scale/>

³ Leon Furze (hämtad 29.12.2024) <https://leonfurze.com/2024/08/09/can-the-ai-assessment-scale-stop-students-cheating-with-ai/>

⁴ Jason M. Lodge (hämtad 29.12.2024) The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action <https://www.teqsa.gov.au/sites/default/files/2024-08/evolving-risk-to-academic-integrity-posed-by-generative-artificial-intelligence.pdf>

⁵ Jämför: Jason M. Lodge (hämtad 29.12.2024) The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action <https://www.teqsa.gov.au/sites/default/files/2024-08/evolving-risk-to-academic-integrity-posed-by-generative-artificial-intelligence.pdf>

1	INGEN AI	Bedömningen utförs i sin helhet utan AI-hjälp i en övervakad miljö, vilket säkerställer att eleverna eller studerandena kan stödja sig på sina existerande kunskaper och färdigheter och sitt kunnande.
2	PLANERING MED HJÄLP AV AI	AI kan användas för förberedelser inför en uppgift för att skapa idéer, göra skisser eller genomföra preliminära undersökningar. På den här nivån är det fokus på effektiv användning av AI vid planering, syntes och idéskapande, men i bedömningen ska det istället vara fokus på förmågan att utveckla och förädla idéerna självständigt.
3	AI SOM VERKTYG	AI kan användas som hjälpmedel i en uppgift för att till exempel skapa idéer, skissa, få feedback och göra preciseringar. Eleven ska kritiskt granska och göra ändringar i de svar som AI genererar och visa att hen förstår dem.
4	FRI ANVÄNDNING AV AI	AI kan användas i alla delar av en uppgift, och eleverna kan använda sig av AI för att uppnå målen för uppgiften. På den här nivån kan det också vara ett krav att eleverna ska använda sig av AI för att uppnå målen och lösa problemen i uppgiften.
5	KREATIV ANVÄNDNING AV AI	AI används på ett kreativt sätt för att främja problemlösningsförmågan, skapa nya insikter eller lösa problem på ett innovativt sätt. Eleverna och lärarna planerar tillsammans bedömningsgrunderna, som fokuserar på kartläggningen av specifika AI-verktyg som kan användas inom utbildningsbranschen.

Leon Furze m.fl. (2024), uppdaterad AIAS-tabell för lärare.

1	INGEN AI	Du får inte använda dig av AI i någon del av uppgiften. Du ska visa prov på relevanta kunskaper och färdigheter.
2	PLANERING MED HJÄLP AV AI	Du kan använda AI för att planera, skapa idéer och genomföra undersökningar. Det ska framgå av slutresultatet på vilket sätt du har vidareutvecklat och -förädlat idéerna.
3	AI SOM VERKTYG	Du kan använda dig av AI i vissa uppgifter, när du till exempel skriver en text, färdigslipar den och utvärderar den. Du ska kritiskt utvärdera och göra ändringar i det AI-producerade material du använder dig av.
4	FRI ANVÄNDNING AV AI	Du kan fritt använda dig av AI i hela uppgiften antingen som du vill eller enligt instruktionerna i uppgiften. Fokusera på att använda AI utgående från de uppsatta målen och på att samtidigt visa prov på kritiskt tänkande.
5	KREATIV ANVÄNDNING AV AI	Du bör använda dig av AI på ett kreativt sätt för att lösa uppgiften och eventuellt utforska nya angreppssätt i samråd med din handledare.

Leon Furze m.fl. (2024), uppdaterad AIAS-tabell för elever.

9. Vad bör en lärare veta om EU:s AI-förordning?

TARMO TOIKKANEN, SITRA

”Införandet av AI-system inom utbildning är viktigt för att främja digital utbildning av hög kvalitet och göra det möjligt för alla studerande och lärare att förvärva och dela de digitala färdigheter och kompetenser som krävs, inbegripet mediekunnighet, och kritiskt tänkande, för att aktivt delta i ekonomin, samhället och i demokratiska processer.”

- ur förordet till EU:s AI-förordning

OLIKA TYPER AV AI-BASERADE lösningar kan vara till nytta i undervisningen, men de bör utvecklas och tas i bruk på ett ansvarsfullt sätt. Om vi låter bara första bästa AI-chattbot bedöma elevernas prestationer och antecknar resultaten som sådana i studieregistret är det inget ansvarsfullt med det vi gör. För att alla ska vara på det klara med vad som avses med ansvarsfull

användning beskriver EU:s AI-förordning situationer som kräver särskild försiktighet.

Ansvaret för AI-lösningarna i undervisningen ligger hos utbildningsinstitutionen, inte läraren. Å andra sidan är alla företag skyldiga att säkerställa sina anställdas AI-kompetens. Därför måste utbildningsinstitutionen utbilda sina lärare i att arbeta med AI-system på ett ansvarsfullt

Utbildningsinstitutionen måste erbjuda lärarna tillräckliga anvisningar om hur olika AI-system ska hanteras.

sätt. I den här korta artikeln kommer jag att beskriva de allmänna principerna i AI-förordningen, som hjälper läraren att få en översikt. Det viktigaste budskapet är dock att utbildningsinstitutionen måste erbjuda lärarna tillräckliga anvisningar om hur de hantera olika AI-system.

Målet är tillförlitliga AI-verktyg

Nästan vad som helst kan föras med eller utföras med hjälp av AI: bilar, leksaker, rekryteringsprocesser, brottsbekämpning eller till exempel elevbedömning. Det är rätt uppenbart att ogenomtänkt tillämpning av AI till exempel i en självkörande bil eller vid bedömning av elevers uppsatser, innebär en hel del risker.

Riskerna kan delas in i två grupper. De kan

1. äventyra människors grundläggande rättigheter eller

2. de kan ge upphov till farliga produkter.

De grundläggande rättigheterna omfattar till exempel rätten till utbildning, arbete, yttrandefrihet, religionsfrihet och jämlikhet.

Farliga produkter regleras genom produktreglering. EU har lagar som reglerar säkerheten i till exempel bilar, hissar, livsmedel eller elektrisk utrustning. CE-märkningen indikerar att tillverkaren lovar att produkten är säker att använda. Samma princip ska gälla också den

delen av produkten som styrs av AI.

Syftet med AI-förordningen är att säkerställa att det på EU-marknaden finns endast tillförlitliga AI-tjänster. Omvänt innebär det att om ett AI-verktyg kan ge

upphov till besvär (eller risker) måste man se till att det är tillförlitligt – det ska alltså utformas med eftertanke.

Det är en hel del som då ska tas i beaktande, men sammanfattat är det fråga om en sak: god datahantering (eng. *data governance*). När dataflöden identifieras, beskrivs och hanteras på rätt sätt är det också lättare att säkerställa ansvarsfull användning och tillförlitlighet. Tack vare datahantering är det lättare att tackla dataskyddsfrågor och säkerställa tillförlitligheten hos AI. I samma veva är det bra att också fokusera på kvalitetskontroll och riskhantering.

Riskenivåer

AI-systemen kan kategoriseras utifrån riskerna de har.

1. **För hög risk: Vissa extrema användningsområden, som inte överensstämmer med europeiska värderingar, är helt och hållet förbjudna.**

Sådana användningsområden är bl.a. social poängsättning av medborgare genom att kombinera olika databaser, biometrisk fjärridentifiering av medborgare i realtid eller kognitiv-behavioristisk manipulering av människor.

2. **Hög risk: AI är en del av en reglerad produkt eller den kan utgöra en betydande risk för människors grundläggande rättigheter. Det här är tillåtet, men AI-tjänsterna måste utformas och implementeras med eftertanke, dokumenteras och godkännas av myndigheterna.**
3. **Minimal risk: Till exempel chattbottar, bildgeneratorer och liknande verktyg, som producerar innehåll som människor kan ta del av. Det enda kravet är att användaren informeras om att det är fråga om ett AI-baserat verktyg.**
4. **Obefintlig risk: Vissa bakgrundsprocesser, som skräppostfilter eller AI-generad terräng i dataspel. Dessa är inte knutna till några skyldigheter.**

Inom utbildningssektorn har bland annat följande AI-system hög risk:

- studentantagning
- bedömning av inlärningsresultat
- bedömning av en persons framtida utbildningsbehov
- övervakning av studerande under provtillfällen.

Listan ovan ger bara vissa riktlinjer, och risknivån påverkas också av hur självständigt AI-systemet tillåts fungera. Om ett AI-verktyg till exempel bara assisterar i bedömningen av inlärningsresultaten, som läraren utför, är det inte nödvändigtvis fråga om högriskanvändning. Om AI istället självständigt får ge förslag på bedömningsresultat är risknivån högre.

När du ska börja använda ett AI-verktyg

Leverantören av ett AI-verktyg måste känna till vilken risknivå som gäller för användningen av verktyget. CE-märkningen i AI-verktyget indikerar att det har utformats i enlighet med EU:s kvalitetskrav.

AI-förordningen förutsätter vissa åtgärder med syfte att säkerställa AI-verktygens tillförlitlighet, när verktyg med hög risk ska tas i bruk. Åtgärderna får dock med fördel vidtas även om det inte är obligatoriskt.

Även om leverantören av AI-verktyget har tagit hand om sin andel och informerat alla involverade bör ni som tar i bruk verktyget också göra er andel.

- **Datahantering:** Hur ska systemet användas, vilka data kan det matas med och vilka personer ska ha tillgång till dem? Datasekretess, dataskyddsfrågor och etiska aspekter ska beaktas.
- **Kvalitetskontroll:** Hur övervakas det att systemet fungerar och hur upptäcks felsituationer? Det är av största vikt att kontrollen utförs av personer med expertis och att den dokumenteras.
- **Riskhantering:** Planen vidare efter felsituationer, datasäkerhetsproblem osv.

Så här utvecklar du AI

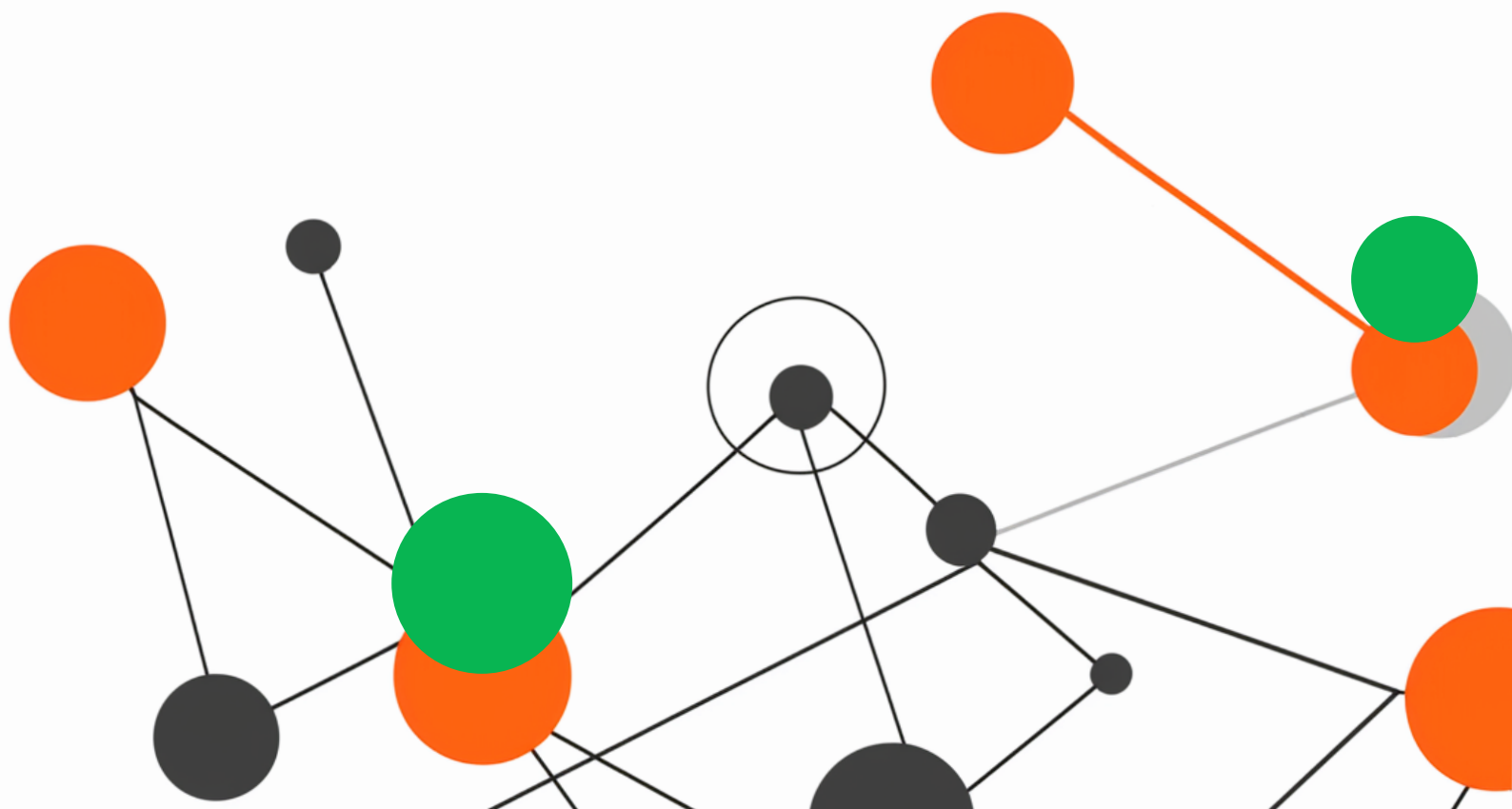
Som utvecklare av ett AI-system ska du identifiera vilka ändamål dina kunder har tänkt använda AI för och vad som är den högsta risknivån bland dem. Är det fråga om ett system med hög risk förutsätter AI-förordningen att vissa åtgärder vidtas, men de får gärna övervägas också annars.

- **Datahantering:** Datakällor, metadata, träningsalgoritmer, tillgång till data
- **Kvalitetskontroll:** Automatiserade test, kontroll av kvaliteten i träningen, kontroll vid användning
- **Riskhantering:** Systemiska risker, utvecklarens egna risker, risker som orsakas av kunders agerande. Riskhanteringsplan.

Läs mera:

I Finland utgör Traficom en samlande instans för tillsynsmyndigheterna. På Traficoms webbplats finns närmare anvisningar och myndigheternas kontaktuppgifter.

Sitra erbjuder en självstudiekurs om AI-förordningen, som kommer att ingå i kursserien Datatalouden ABC ('ABC i dataekonomi'). Kursen färdigställs och publiceras under våren 2025 på adressen <https://www.sitra.fi/hankkeet/datatalouden-abc/>



Undervisning om AI-färdigheter för elever med hjälp av Generation

10. AI-verktyg

JARI LARU, PED, UNIVERSITETSLEKTOR, FAKULTETEN FÖR PEDAGOGIK OCH PSYKOLOGI, ULEÅBORGS UNIVERSITET

GENERATION AI ÄR ett projekt som finansieras av Rådet för strategisk forskning (RSF) och som administreras av Finlands Akademi.

Det övergripande målet med Generation AI-projektet är att stödja barn och unga i att förstå tekniker som baserar sig på AI och maskininlärning. Projektet strävar också efter att hjälpa barn och unga att kritiskt och etiskt granska vilka bakgrundskrafter, användningsområden och effekter AI-baserade tekniker har. En viktig utgångspunkt är dessutom att främja barnens kreativa färdigheter och att hjälpa dem att växa till medvetna och ansvarsfulla människor som skapar vår gemensamma framtid. Målet är att eleverna ska få meningsfulla upplevelser av att de kan lyckas och känna att de kan påverka det de tycker är viktigt.

Det behövs applikationer som stöd för undervisningen om AI

Inom projektet utvecklar vi applikationer som är säkra, responsiva och tillgängliga. Vi samlar inte in data och vi spårar inte användarna. Rättsvetare inom projektet har säkerställt att våra applikationer uppfyller EU:s dataskyddskrav. Utöver det är våra applikationer gratis och de har lanserats under en MIT-licens.

Maskinen som lär sig

Olika tillämpningar av AI är starkt närvarande i vår vardag. Vi möter dem till exempel på sociala medier och i form av text- och bildgeneratorer samt exempelvis när vi låser upp telefonen med hjälp av ansiktigenkänning. Få av oss tänker ändå på hur AI faktiskt fungerar, hur AI-verktyg utvecklas och varför AI ibland gör fel.

Maskinen som lär sig (fi. 'opetettava kone') är ett verktyg som hjälper elever att lära sig hur AI fungerar genom att skapa applikationer som använder sig av maskinseende. Med hjälp av maskinen får eleverna sätta sig in i hela arbetsflödet för maskininlärning från att data samlas in till att den färdiga mobilapplikationen tas i bruk – helt utan några programmeringskunskaper! Ett sådant angreppssätt minskar behovet av tekniska färdigheter, vilket gör att

inlärningsprocessen blir mycket mer engagerande¹.

Applikationen lämpar sig för alla åldrar, och den har utvecklats särskilt med barn och unga i åtanke. Med hjälp av Maskinen som lär sig får eleverna på ett konkret sätt en inblick i vilka möjligheter det finns med AI och vilka begrepp som ligger i grunden för AI. Maskinen får eleverna också att fundera över vilka begränsningar AI har och vilka etiska frågor som användningen av AI aktualiserar.

Sociala mediemaskinen – en simulator för sociala medier

Sociala mediemaskinen är ett verktyg för förklarande artificiell intelligens (XAI), som riktar sig till eleverna i den grundläggande utbildningen. Maskinen har ett bekant, interaktivt användargränssnitt, som liknar populära plattformar på sociala medier, som eleverna redan känner till. Maskinen ger eleverna möjlighet att "titta in under huven" och avslöjar hur "Instagram fungerar". Syftet är att lära eleverna på ett lättillgängligt och intressant sätt hur centrala datadrivna och AI-baserade mekanismer på sociala medier fungerar och vilka centrala begrepp, såsom datainsamling, profilering och rekommendation, som förknippas med mekanismerna. Syftet är också att utveckla elevernas datahanteringsförmåga, AI-kompetens och förståelse av AI-etik².

**Allt material
från barnens
AI-projekt
publiceras under
våren 2025.**

Med hjälp av Sociala mediemaskinen får eleverna egna förstahandserfarenheter av hur även de minsta handlingarna, som att pausa en video, sparas

i användarens digitala fotavtryck och vidare i användarens personliga profil. I Sociala mediemaskinen används realtidsvisualiseringar, som hjälper användaren att se hur hens egna och andras agerande påverkar användarupplevelsen på sociala medier.³

Undervisningsmaterialet och närmare instruktioner för användningen av Sociala mediemaskinen publiceras i samband med Educa-mässan våren 2025, och efter det finns de tillgängliga på projektets webbplats.

Applikationer och material som skapas tillsammans

Generation AI ordnar årligen ett engagerande samplaneringsprojekt för alla projektets skolor i Joensuu och Uleåborg med praktiskt orienterade aktiviteter för barn i olika åldrar. I samband med projektet får lärare och elever tillsammans jobba med tvärvetenskapliga undersökningar.

Projektet har utformats utgående från att samma barn och unga varje år under tre på varandra följande år (se tabell 1) får delta i ett projekt, där de tillägnar sig grundläggande begrepp och färdigheter och lär sig om hur de samhälleliga effekterna syns i deras vardag. Samplaneringen går delvis ut på att utveckla undervisningsmaterial⁴ och verktyg⁵, som senare görs tillgängliga för alla lärare och elever i vårt land.

Det första skolprojektet genomfördes 2023. Temat för projektet var datadriven design. Inom ramen för projektet fick de deltagande skoleleverna planera maskininlärningsbaserade klassificeringsapplikationer med hjälp av Maskinen som lär sig, som utvecklats inom projektet.⁶

Maskinen som lär sig:
<https://tm.generation-ai-stn.fi>
Källkoden för maskinen:
<https://github.com/knicos/genai-tm>
Undervisningsmaterial:
<https://www.generation-ai-stn.fi/materiaalit>

Sociala mediemaskinen (fi. 'Somekone'):
<https://somekone.generation-ai-stn.fi>
Källkoden för maskinen:
<https://github.com/knicos/genai-somekone>
Materialet för Sociala mediemaskinen (på finska): publiceras under våren 2025 på adressen:
<https://www.generation-ai-stn.fi/materiaalit>.

Skolprojekt	Namn	Begrepp	Lektioner	Verktyg
2023	Datadriven design (Hur fungerar AI?)	Träningsdata, klassificerare, tillförlitlighet, svagheter, algoritmisk bias	Grundläggande AI-begrepp och -mekanismer Datadriven design Effekter av AI och AI-etik	Maskinen som lär sig + material
2024	Mekanismer på sociala medier	Datainsamling, profilering, rekommendation	Under utveckling, publiceras Q1/2025	Profilspele Sociala mediemaskinen + material, som publiceras Q1/2025
2025	Effekterna av sociala medier	Bubblor på sociala medier, oavsiktliga effekter, påverkan på beteendemönster och åsikter	Under utveckling	Under utveckling

Tabell 1. Skolprojekten under åren 2023–2025

Projekt	Lektion	Uppgiftsbeskrivning
Robotar (förskola)	1	• Barnen ritar robotar på papper och på en interaktiv tavla.
	2	• Barnen bygger robotar av återvunnet material.
	3	• Barnen använder sig av AI för att skapa en berättelse med lärarens hjälp. • Barnen presenterar robotarna som de har byggt.
Robotfabrik (grundläggande utbildning, åk. 1, specialskola)	1	• Barnen ritar robotar.
	2	• Läraren har skapat bilder med hjälp av AI, som utforskas tillsammans. • Barnen ritar, formar och bygger robotfabriker. • Läraren skapar bilder med hjälp av generativ AI utifrån barnens planer.
	3	• Barnen skapar berättelser om sina egna robotar i grupper på tre. • Läraren sammanställer en gemensam berättelse utgående från barnens egna berättelser med hjälp av ChatGPT.
Superhjältar (förskola)	1	• Barnen ritar superhjältar med färgpennor. • Barnen berättar vilka egenskaper superhjältarna har.
	2	• Läraren skapar bilder utgående från barnens ritningar med hjälp av generativ AI.
	3	• Läraren skapar en berättelse utifrån barnens beskrivningar med hjälp av generativ AI.

Tabell 2. Exempeluppgifter för förskoleundervisning och grundläggande utbildning

Steg	Rubrik	Beskrivning
1	Undersök profilerna som skapats med hjälp av Sociala mediemaskinen.	• Hur hänger bilderna och nyckelorden ihop? • Vilken typ av innehåll är viktigt eller intressant för dessa personer? • Vilken typ av visuellt innehåll tilltalar dessa personer och fångar deras uppmärksamhet?
2	Planera en reklamkampanj för den här målgruppen.	• Vad har kampanjen för bakgrundshistoria och budskap? • Vad vill du särskilt lyfta fram kring temat du har valt (det ursprungliga temat var Helsingfors)? • Vad har kampanjen för slogan?
3	Skapa reklammaterial för kampanjen med hjälp av generativ AI.	• Bilder • Faktablad, berättelser eller andra texter för reklamkampanjen • Videor

Tabell 3. Reklamkampanjen "Helsingfors – sommarstaden för unga" steg för steg

I praktiken deltog eleverna i tre workshoppar med fokus på grundläggande AI-begrepp och -mekanismer, datadriven design och AI:s effekter och etik⁷.

Ett annat skolprojekt år 2024 satte fokus på mekanismerna på sociala medier, som behandlades på två olika lektioner. På första lektionen fick eleverna med hjälp av ett profileringspel och övningsuppgifter utan maskininblandning lära sig om datainsamling, profilering och rekommendationer. På följande lektion fick eleverna använda simulatoren Sociala mediemaskinen⁸, som hjälpte dem att undersöka hur mekanismerna på sociala medier faktiskt fungerar. Allt material som ingår i undervisningshelheten publiceras under våren 2025, och efter det ges det också ut närmare information om helheten. Planeringen inför ett tredje skolprojekt ska också inledas snart, och närmare information om det ges också senare.

Hur använda AI på ett kreativt sätt?

Epistemisk osäkerhet ser kanske ut att vara en rätt svårbegriplig term, men det är den som utgör utgångspunkten för reflektioner över hur vi kan lära oss leva, existera, arbeta och studera i en värld fylld med AI-teknik. I och med uppkomsten av generativ AI är var och en av oss tvungen att i viss mån tvivla på tillförlitligheten i det som produceras av AI. En total misstro gör det dock omöjligt att hantera AI-tekniker, så det är fråga om ett slags balansgång⁹.

Förskole- och nybörjarundervisning

När det gäller förskole- och nybörjarundervisning har det hittills undersökts vad man kan göra med hjälp av AI och hur det kan knytas an till förskoleundervisningens läroplan. Forskningen har bland annat utmynnat i robot-, robotfabriks- och superhjälteprojekt, som alla har gått ut på att läroinnehållet i förskoleundervisningen behandlats med hjälp av

generativ AI. AI har använts till exempel till att berätta historier, skapa bilder och ställa frågor.

I projekten ovan har forskare kunnat identifiera olika metaforer som barn och deras lärare använder om AI och a) vad AI är: dator (program), människa (assistent), fantasi (trollkarl), b) hur AI fungerar: dator (kommandon), människa (intelligens, sätt att agera), jämförelser mellan människa och teknik, jämförelser mellan naturen och den digitala världen, jämförelser mellan fantasi och människans handlingar, c) AI:s existens: tvivel / tillit, att dämpa oron.

Grundläggande utbildning

När det gäller den grundläggande utbildningen har det tillsammans med eleverna genomförts olika projekt med syfte att dra nytta av generativ AI. Projekten bland annat tangerat teman som lyxkonst och sociala medier, medier och marknadsföring.

Tabellen nedan ger en beskrivning av en exempeluppgift, där sjätteklassare har i uppgift att skapa affischer för reklamkampanjen "Helsingfors – sommarstaden för unga».

I det här fiktiva projektet har sjätteklassarna använt sig av Sociala mediemaskinen för att skapa olika profiler för ungdomar, och därefter har de haft i uppgift att utarbeta marknadsföringsmaterial för unga med hjälp av generativ AI.

Allt material från AI-projekten inom den grundläggande utbildningen publiceras under våren 2025.

Beakta barnens digitala rättigheter

Förutom att utmynna i verktyg, metoder och undervisningsmaterial förser Generation AI-projektet finländska lärare och utbildningsanordnare med kunskap om nationell lagstiftning och EU-reglering. När det gäller EU är det bland annat fråga om dataskydds- och AI-förordning. Professor Susanna Lindroos-Hovinneimo har skrivit flera olika artiklar

om EU:s AI-förordning, varav en är en kort introduktion till EU:s AI-förordning, som publicerats i årsboken för medierätt ('Viestintäoikeuden vuosikirja') år 2023.¹⁰

Digitala miljöer är ofta inte utformade med barn eller barns rättigheter i åtanke, men trots det har de omfattande konsekvenser för barnens rättigheter. Det har också identifierats en tydlig spänning till exempel mellan behovet av att skydda barnen och barnens självständiga engagemang. Hur kan vi till exempel skydda barnen för olika risker och samtidigt se till att de får tillgång till digitala verktyg i sin skolvardag? Det här är också temat för en fritt tillgänglig forskningsartikel i den finska tidskriften *Lakimies*¹¹, som fått välförtjänt uppmärksamhet till exempel bland utbildningsanordnare.

Lärare och andra vuxna har en avgörande roll när det gäller att tillgodose barnens rättigheter i praktiken. Det räcker inte att rättigheterna endast diskuteras på en högre nivå. En särskilt viktig kontext med tanke på hur barnens digitala rättigheter kan tillgodoses utgörs av klassrummet och skolans vardag. Dataskyddsguiden från Generation AI-projektet (fi. 'Tietosuojapas')¹², som lagts upp på projektets webbplats, kan erbjuda stöd i det sammanhanget.

Sammanfattning

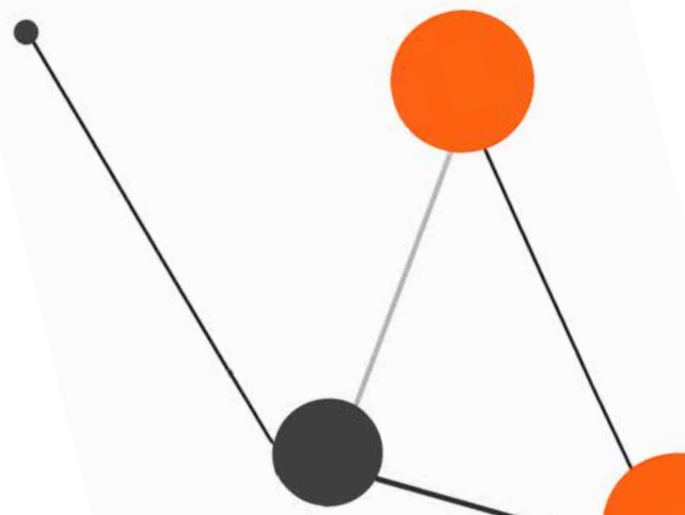
De applikationer, metoder, material och olika guider som utvecklats inom ramen för projektet Generation AI baserar sig på referentgranskad och högklassig internationell vetenskaplig forskning. Vi utvecklar grunden för AI- och säkerhetsfostran i vårt land utifrån det här kunskapsunderlaget.

Projektet ligger verkligen i tiden, eftersom det inte finns någon omfattande forskning om barnens rättigheter, när det gäller AI-kontexten, och eftersom det finns få forskningsbaserade pedagogiska modeller för undervisningen om AI och förmågan att vara aktiv i den digitala världen.

Vi utvecklar lösningar som också internationellt sett är unika. Du som lärare i Finland kan alltså känna dig privilegierad.

Följ vår resa och kom med på den!

- 1 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a): A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 2 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b): An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 3 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b): An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 4 Kahila, J., Vartiainen, H., Tedre, M., Laru, J., Arkko, E., Lin, A., Pope, N. & Jormanainen, I. (2024a): Tekoäly oppimisen kohteena ja luovan toiminnan lähteenä. Valkoinen kirja. January Collective. <https://www.generation-ai-stn.fi/kirjasto/opetettavakone/mika-on-opetettava-kone/opetettava-kone-ja-materiaalit-perustuva-tutkimukseen/> Kahila, J., Vartiainen, H., Tedre, M., Arkko, E., Lin, A., Pope, N., Jormanainen, I. & Valttonen, T. (2024b): Pedagogical framework for cultivating children's data agency and creative abilities in the age of AI. *Informatics in Education*. doi:10.15388/infedu.2024.15
- 5 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a): A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT). Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b): An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 6 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a): A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 7 Kahila, J., Vartiainen, H., Tedre, M., Laru, J., Arkko, E., Lin, A., Pope, N. & Jormanainen, I. (2024a): Tekoäly oppimisen kohteena ja luovan toiminnan lähteenä. Valkoinen kirja. January Collective. <https://www.generation-ai-stn.fi/kirjasto/opetettavakone/mika-on-opetettava-kone/opetettava-kone-ja-materiaalit-perustuva-tutkimukseen/> Kahila, J., Vartiainen, H., Tedre, M., Arkko, E., Lin, A., Pope, N., Jormanainen, I. & Valttonen, T. (2024b): Pedagogical framework for cultivating children's data agency and creative abilities in the age of AI. *Informatics in Education*. doi:10.15388/infedu.2024.15
- 8 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b): An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
- 9 Davies, S., Sormunen, K., & Kangas, K. (2024): Cultivating Epistemic Doubt: a Key Competence for Productive Participation in the Era of Artificial Intelligence. 17th International Conference on Computer-Supported Collaborative Learning (CSCL) 2024 - Buffalo, USA.
- 10 Lindroos-Hovinheimo, S. (2023): Lyhyt johdatus Unionin tekoälyasetukseen. I: Viestintäoikeuden vuosikirja 2023 (s. 67-76). Forum Iuris, Helsingfors universitet, Juridiska fakulteten. <https://www.edilex.fi/viestintaoikeuden-vuosikirja/1000880004>
- 11 Silvennoinen, E., Tedre, M., & Valttonen, T. (2024): Datafikoituva peruskoulu - tasapainoilua lapsen henkilötietojen suojan ja opetuksen tavoitteiden välillä. *Lakimies*, 122(5), 655-678.
- 12 Generation AI Tietosuojapas (hämtad 18.12.2024) <https://www.generation-ai-stn.fi/henkilotietojen-kasittely-alakoulussa/>



11. Generation AI-projektet

JARI LARU, PED, UNIVERSITETSLEKTOR, FAKULTETEN FÖR PEDAGOGIK
OCH PSYKOLOGI, ULEÅBORGS UNIVERSITET

GENERATION AI ÄR ett projekt inom programmet Säkerhet och tillit i algoritmernas tid (Shield). Shield är i sin tur en projekthelhet om AI i anslutning till Finlands Akademi, som finansieras av Rådet för strategisk forskning (RSF).

Shield-programmet fokuserar på säkerhet och tillit i en tid av informationsteknik och hybridpåverkan. Vid sidan av de mer traditionella säkerhetshoten påverkas människans handlingsmöjligheter på oförutsägbara sätt av utmaningar och möjligheter som digitaliseringen och den tekniska utvecklingen för med sig samt av olika typer av kriser från minskad biologisk mångfald till klimatförändringar, finanskriser och maktpolitiska konflikter.

Varför är Generation AI-projektet viktigt just nu?

Projektet Generation AI svarar på tre sammanlänkade samhällsutmaningar:

1. **Försämrade känsla av trygghet, kontroll och fri vilja i ett datadrivet samhälle**
2. **Försämrade tillit till myndigheterna, medierna, vetenskapen, civilsamhället och medborgarna**
3. **Snabbt ökande ojämlikhet, som är kopplad till förmågan att hantera och viljan att engagera sig i informations- och kommunikationstekniken, som är i ständig förändring.**

Inom ramen för projektet utvecklas interventioner i förskole- och skolundervisningen, som har som syfte att stärka barnens resiliens i en alltmer teknikcentrerad värld, att skapa förutsättningar för och främja deras förmåga att beakta cybersäkerheten och att förse lärare med insikter och pedagogiska verktyg, som behövs för att de ska kunna anpassa läroplanerna. Därtill främjar projektet forskningen om barnens rättigheter när det gäller artificiell intelligens.

Vad går projektet ut på?

Projektet skapar ett fundament för "AI-generationens" teknologifostran: det är fråga om teknologifostran som beaktar principerna för hur AI fungerar samt de möjligheter och dynamiska effekter som AI för med sig – utan att glömma bristerna hos och riskerna med AI eller

det faktum att AI-system kan öka motsättningar, diskriminering och känslan av otrygghet och urholka förtroendet i samhället.

Projektet främjar elevernas förståelse av säkerhet och tillit till AI- och maskininlärningsbaserade system. Projektet mäter hur detta påverkar elevernas datahanteringsförmåga och deras förmåga att arbeta med komplexa och icke-linjära processer och att göra dem mindre oförutsägbara och polariserande.

Projektet ger eleverna möjlighet att identifiera konsekvenserna av sina ofta omedvetna val på internet. På så sätt är det möjligt att visa hur fördomar smyger sig in i AI-system och skapar algoritmisk diskriminering och ojämlikhet. Det gör eleverna lyhörda inför frågor om kontroll och teknisk determinism och om vem som har makt över användningen av data och på vilket sätt användarnas tillit kan missbrukas.

Vilken är målgruppen för Generation AI-projektet?

Projektet för samman teknikutvecklare, skolor, myndigheter, företag och icke-statliga organisationer. Det stärker barn och ungas förmåga att hantera effekterna tekniken har på samhällen globalt och skapar pedagogiska verktyg för lärare kring lärandet som är i förändring.

En av de främsta utmaningarna på 2000-talet är att få barn och unga att komma till insikt om tekniken som framtidens samhälle i allt högre grad är beroende av. Inspirerat av den här utmaningen skapar Generation AI-projektet förutsättningar för och stärker den vetenskapliga förståelsen, som i sin tur möjliggör utbildningslösningar som beaktar AI-generationens kompetensbehov.

Projektets webbplats:

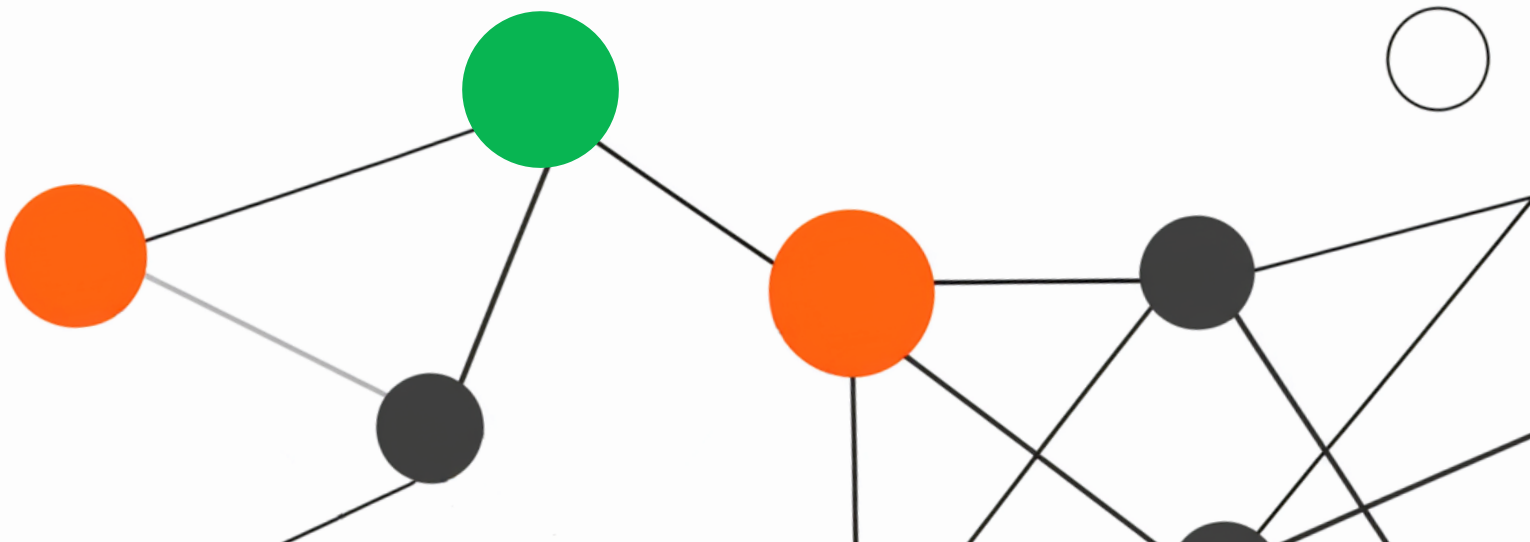
<https://www.generation-ai-stn.fi>

Undervisningsmaterial som utvecklats inom projektet: <https://www.generation-ai-stn.fi>

Applikationer som utvecklats inom projektet:

Maskinen som lär sig: <https://tm.generation-ai-stn.fi>

Sociala mediemaskinen ('Somekone'):
<https://somekone.generation-ai-stn.fi>



12. AI, det gör ont – AI förändrar pedagogiken i Villmanstrand

MIIKA MIININ, EDUKS

FÖR UNGEFÄR TVÅ ÅR sedan gjorde AI entré i klassrummen med buller och bång. Eleverna tog snabbt till sig det nya verktyget, och lärarna fick fundera över sin undervisning ur ett nytt perspektiv. Verket, som vid första anblicken inte verkade vara mycket mer än en teknisk nymodighet bland andra, visade sig skaka om hela skolvärlden.

I höstas avklarade AI studentskrivningarna galant och fick vitsordet laudatur i flera ämnen. Texterna som gavs det höga vitsordet hade skapats genom att mata in frågorna i ett AI-verktyg. Med andra ord förutsatte de AI-genererade svaren ingen specifik kompetens. Det här ger upphov till frågan om vad och hur man ska studera på gymnasiet.

På grund av den här nya utmaningen önskade lärarna i Villmanstrand sig först och främst hjälp med uppgiftsbeskrivningar, så att AI inte skulle kunna göra uppgifter åt studerandena. Vi kunde uppfylla lärarnas önskemål genom att ge ut en guide med tio konkreta tips för uppgiftsbeskrivningar i den nya AI-eran. Guiden blev till tack vare samarbetet med Eduks, som Utbildningsstyrelsen finansierade, Otavia, gymnasieväsendet i Kuopio och Mobie.¹

I guiden tar vi inte ställning till om man ska försöka begränsa användningen av AI eller om den ska göras till en del av inlärningsprocessen: guiden innehåller tips till alla lärare, oavsett hur de förhåller sig till AI.

Efter att den initiala chocken lagt sig började lärare se AI som en möjlighet istället för ett monster och vill veta mer om hur de kan effektivisera sitt eget arbete: Hur kan AI hjälpa mig att utföra mina arbetsuppgifter bättre? I följande guide fortsatte vi med konceptet med tio konkreta tips.² Den här gången var det fokus särskilt på de bakgrundsprocesser som gör att läraren inte fullt ut kan fokusera på det viktigaste, dvs. att bemöta studerandena. Guiden gav tips för hur man kan kommunicera, skapa idéer, producera nytt innehåll, hålla möten osv.

Det blir aktuellt med ansvarsfull användning av AI först efter att de mest akuta bränderna har släckts. För närvarande arbetar vi med en guide för lärare om etisk och datasäker användning av AI. En sådan

guide skulle inte ha varit till någon större nytta om lärarna inte först erbjudits önskad hjälp med de tidigare problemen.

Alla nämnda och också kommande

guider finns på adressen:

<https://tekoalyoppimisentukena.wordpress.com/tot-oppaat/>

Skomakare, bliv vid din läst

Många AI-tjänster tampas med två envisa utmaningar vid hantering av inmatningsdata: De hallucinerar, dvs. producerar text med felaktigheter förorsakade av felgissningar. Den andra utmaningen handlar om att AI kan avvika från källmaterialet: AI har tränats med en stor mängd olika texter, och i sina svar vandrar AI därför lätt i väg från det aktuella materialet till informationen i materialet som den tränats med.

ChatGPT, Copilot, Claude och andra generativa AI-verktyg har revolutionerat arbetet på många sätt. Ett AI-verktyg som fått mindre uppmärksamhet har emellertid blivit lite av en favorit inom Eduks speciellt i uppgifter som går ut på att hantera stora datamängder, söka information eller få en snabb översikt över ett nytt tema.

Googles Gemini-baserade Notebook LM sticker åtminstone för närvarande ut tack vare sin förmåga att hålla sig till källmaterialet. Notebook LM kan matas med enorma mängder material: dokument, webbplatser, Youtube-videor eller till exempel oredigerad text. När materialet har matats in är det möjligt att diskutera med Notebook LM, precis som med vilket som helst annat AI-verktyg. Notebook LM söker igenom det inmatade materialet (och endast det) för önskad information, sammanfattar innehåll, letar fram motstridigheter eller överlappningar och skapar vid behov ett samtal utgående från materialet, där maskinröster, som låter som människor, går igenom materialet som i ett vanligt podcastavsnitt om temat i fråga.

I praktiken har Notebook LM ett flertal andra intressanta användningsområden: om det inom utbildningsväsendet ska införas

ett nytt undervisningsprogram bör man på förhand göra en konsekvensbedömning och utarbeta ett eventuellt avtal om behandling av personuppgifter för det nya programmet. Den typen av bakgrundsarbete förutsätter ett stort antal utredningar, eftersom man måste gå igenom dataskyddsbeskrivningar, lagtexter, kommunens anvisningar, riktlinjer och så vidare. Arbetet kan utföras betydligt snabbare när allt nödvändigt material matas in i Notebook LM: genom att diskutera med AI-verktyget hittar man snabbt önskad information med tydliga källhänvisningar.

Å andra sidan kan skolor kan använda Notebook LM i sitt arbete kring läroplaner genom att mata in skolans gamla läroplan, läroplansgrunderna och närmare information om skolans fokusområden i verktyget. På så sätt är det möjligt att identifiera brister, utreda hur skolans egna fokusområden på ett effektivt sätt integreras i läroplanen och säkerställa att inget viktigt utelämnas eller formuleras på ett motsägelsefullt sätt.

AI på skärmen där, säg vem som en bättre president är!

Lektor Hanna Sydänmaanlakka tog vara på möjligheterna med AI på kursen om akademiskt skrivande på engelska. I Sydänmaanlakkas exempel förutsatte uppgifterna att studenterna använde sig av AI: studenterna skulle använda AI för att bearbeta och tillägna sig fakta.

Studenterna fick i uppgift att mata in tio så olika källor som möjligt om presidentkampen mellan Trump och Harris i ett valfritt AI-verktyg. Utifrån dessa källor skulle studenterna sedan bilda sig en egen uppfattning om vilken av kandidaterna som tilltalade dem själva mest samt vilken av dem som var bättre för Finland och varför. För att motivera sitt val skulle studenterna fundera över faktorer som påverkade deras val. De skulle också jämföra kandidaterna utifrån dessa faktorer samt utifrån olika politiska vägval och löften genom att diskutera med AI-verktyget.

Genom att testa olika AI-verktyg lärde sig studenterna vad som skilde kandidaterna åt: en

del verktyg vägrade helt och hållet diskutera politik, en del vandrade i väg utanför källmaterialet i sina svar och andra hallucinerade i största allmänhet. Studenterna skulle skriva en uppsats om sina slutsatser, ange källorna som använts i uppsatsen, räkna upp AI-verktygen de använt sig av och analysera hur diskussionen med AI lyckats. Tiden som gick åt till att göra uppgiften var två gånger 75 minuter.

Uppgiften vägledde studenterna att reflektera över olika infallsvinklar: för att peka ut kandidaten som var bäst för Finland fick man fundera över vilka teman som var viktigast ur Finlands synvinkel: ekonomi, säkerhet, tolerans eller någonting annat. Hur mycket tyngd ska varje tema ha i vågskålen och avviker den egna ståndpunkten från Finlands intressen – varför och på vilket sätt?

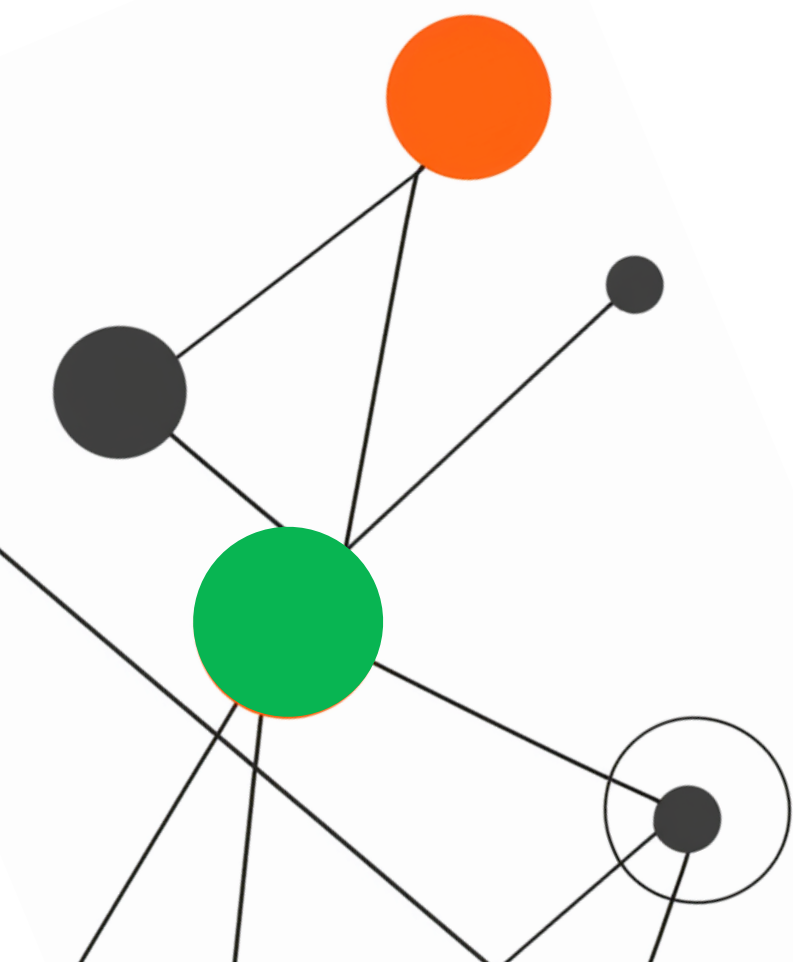
Studenterna utvecklade sin förmåga till kritisk läsning närmast obemärkt, då de kring samma tema lätt kunde identifiera ytterst polariserade synsätt med en stark agenda. Sydänmaanlakka blev glad över kvaliteten på studenternas

reflektioner. En student kommenterade uppgiften så här: *”Jag trodde uppgiften snabbt kunde göras med hjälp av AI, men uppgiften visade sig vara så intressant att jag lade ner mycket mer tid på den än planerat!”* Experimentet uppmuntrade studenterna också att jobba med AI. Sydänmaanlakka berättade om en student som tidigare känt en viss rädsla för AI, eftersom det är möjligt att diskutera med AI-verktyg på samma sätt som med intelligenta, riktiga människor. Tack vare uppgiften insåg studenten att AI inte är något man behöver vara rädd för utan att AI uttryckligen kan användas för att analysera material.

Med hjälp av AI kan människor mycket snabbare hantera betydligt större mängder källmaterial än med traditionella metoder. Med rätt användning av AI leder det oberoende av uppgift till bättre resultat: mera, bättre, med mindre ansträngning.

¹ <https://hankejulkaisut.mobiezine.fi/zine/14/cover>

² <https://hankejulkaisut.mobiezine.fi/zine/15/cover>



Förnekelse vs 13. teknooptimism: AI ur lärarens synvinkel

AKI SAARIAHO, OTANIEMEN LUKIO

VÅRT SAMHÄLLE HAR förändrats oåterkalleligt – eller har det? AI har smugit sig in i vår vardag och också skolvärlden dels lite i smyg, dels med buller och bång. Alltid när man talar om AI brukar man nämna att AI redan länge varit en del av tekniken och apparna vi använder¹. Å andra sidan var det en händelse som gjorde AI plötsligt aktuell

uttryckligen i undervisningen – lanseringen av ChatGPT i november-december 2022.

Generativ AI har inneburit förändringar i hur många lärare planerar sin undervisning, förhåller sig till lärande och lär sig nya saker tillsammans med eleverna. På olika håll har det skrivits och skrivs otaliga guider om och beskrivningar av vad AI är och hur den används^{2 3 4 5}. GenAI-verktygen är till hjälp vid lektionsplanering och innehållsskapande: elever och studerande använder sig gärna av en mirakelmaskin som

ser ut att producera svar som verkar trovärdiga. Just nu och mer än någonsin tidigare i mänsklighetens historia lär sig lärare att använda verktyg samtidigt som deras elever⁶.

När det gäller lärandet finns det tre huvudsakliga förhållningssätt gentemot AI-verktyg: förnekelse, försiktighet och teknooptimism.

Tre förhållningssätt

När det gäller lärandet finns det tre huvudsakliga förhållningssätt gentemot AI-verktyg: förnekelse, försiktighet och teknooptimism. Dessa kan säkert delas upp närmare och noggrannare på olika sätt än så, men dessa tre är förmodligen de tre huvudsakliga förhållningssätten.

Förnekarna har bestämt sig för att det finns så många risker, osäkerhetsfaktorer och rädslor förknippade med AI att de totalvägrar att ha någonting med AI-verktyg att göra. Sådana personer vill att det ska vara förbjudet att använda AI i undervisningen och när eleverna arbetar eller lär sig använda nya applikationer.

Argument som lyfts fram då är bland annat miljöaspekten, riskerna med informationshantering eller förlusten av olika färdigheter som människan med åren lärt sig, om vi förlitar oss på AI. Varje liter vatten som går åt till att hålla i gång ett AI-verktyg tolkas i det sammanhanget som ett tecken på att användningen av AI är ohållbar

från början och att den också alltid kommer att vara det.

Eftersom det inte går att bygga datacentraler på ett hållbart sätt eller att producera energi utan att påverka miljön är det ingen vits med att försöka göra AI-verktygen bättre. Läs- och skrivkunighet hotas av AI och kommer därför att dö ut, om vi i en större utsträckning börjar använda AI-verktyg, eftersom de i likhet med andra rusmedel leder till beroende och fördärv.

Teknooptimisten ser däremot AI som en möjlighet att ta människans potential till en högre nivå. AI-tekniken möjliggör ett skutt framåt för produktiviteten, i och med att vi istället för triviala uppgifter kan fokusera på att nå nya sfärer med vår kreativitet.

Teknooptimisten ser därför enbart möjligheter i AI-verktygen, litar på vår riskhanteringsförmåga och ser med stort intresse fram emot nästa utvecklingssteg. Hen vill inte vänta på myndigheternas gemensamma riktlinjer eller nationella anvisningar utan experimenterar modigt med de möjligheter AI erbjuder både i undervisningen och i privatlivet. För optimisten innebär den nya tekniken nya möjligheter, och den gör det möjligt att skapa tidigare otänkbara applikationer för vetenskaper och tekniker. Optimisten nämner läkarvetenskapen, transportbranschen, individuellt lärande och räddningsinsatser som exempel på något som tidigare varit utom hens räckhåll.

Den försiktiga håller sig noga inom bestämda ramar. Experiment med AI är möjliga, om det först fastställs vilka experiment det är tryggt att genomföra. Så länge vi inte vet var elevernas data – alla data, till exempel uppsatser, inspelade videor eller personuppgifter – lagras, vågar den försiktiga inte testa AI-verktyg i undervisningen. Risken för att personliga data hamnar i fel händer är alltför stor. Tillståndet till användningen av AI ska helst komma från chefsnivå.

Den försiktiga läser igenom diskussionsforum och -trådar på nätet innan hen kan tänka sig att testa AI-verktyg. Hen håller sig avvaktande, tills en pålitlig instans slagit fast att det är tryggt

att använda ett visst AI-verktyg. Den försiktiga vill att vi först har tagit i beaktande alla tänkbara avtal och förutsett alla fallgropar. Den försiktiga är rädd för att användningen av AI helt spårar ur⁷.

Orsaker till användning av AI och det man ska akta sig för

Det finns redan många olika AI-verktyg för olika användningsområden, och de blir fler hela tiden⁸. Just nu, vid årsskiftet 2024–2025, har vi med stor säkerhet bara sett början av vad AI-verktygen kan bli. Vi förväntar oss stora fördelar, som en del också redan varit med om, men en hel del andra förhåller sig fortfarande försiktiga och vågar inte ens testa AI-verktyg. Förnekarna blir kvar och slåss mot väderkvarnar – AI är här för att stanna, åtminstone i någon form.

Framför allt är det bra att se AI som en assistent som klarar av att hantera stora datamängder – så stora mängder data att människan eller människohjärnan helst inte ens ska försöka ta sig an dem. AI-verktygen är bland annat alldeles överlägsna när det gäller att hitta det väsentliga i olika slags källmaterial, jämfört med verktygen som användes före tiden med generativa AI-verktyg. Det är emellertid viktigt att kunna formulera fungerande promptar: bearbetningen av ett visst källmaterial kan göras effektivt och på ett ändamålsenligt sätt endast med sådana promptar som får AI-verktyget att leta fram önskade data.

Källmaterialet är förstås också bara så "bra" och "korrekt" som materialet det består av. Precis som det framgått i den här guiden kan källmaterialet som AI använder sig av vara korrupt eller ytterst snedvridet. Det kan i sin tur leda till att AI, så att säga, hallucinerar, vilket tas upp på andra ställen i guiden.

Det viktigaste vid användning av AI-verktyg är källkritik. GenAI-verktygen är inte renodlade

Framför allt är det bra att se AI som en assistent som klarar av att hantera stora datamängder.

sökmotorer – även om den gränsen nog också sakta suddas ut när sökmotorer integreras i verktyg som baserar sig på stora språkmodeller.

Med tanke på undervisning och lärande är den bästa utgångspunkten att eleven lärt sig vara källkritisk. Eleven ska kunna bilda sig en uppfattning om huruvida texten AI-verktyget genererat håller måttet, passar in på den aktuella kontexten och kan förbättras ytterligare, huruvida den information verktyget tagit fram ur källmaterialet är tillförlitlig och om den faktiskt svarar på frågan i prompten eller huruvida det bild-, ljud- eller videomaterial som skapats med hjälp av AI är ändamålsenligt och det man önskat sig och huruvida det sparar tid – och allt detta dessutom med hänsyn till etiska aspekter.

Slutkläm om det etiska

I skrivande stund finns det både bra och dåliga exempel på de etiska aspekterna av utvecklingen av AI-verktyg. De dåliga exemplen berör miljöfrågor och lagstiftning. Det är uppenbart att utvecklingen av AI och arbetet som utförs med hjälp av den kräver en enorm mängd både energi och naturresurser. I Finland har inte många skogar ännu huggits ned för att göra plats för datacentraler, men det kommer säkert att hända.

Energiproduktion kräver ökad kapacitet, om utvecklingen av AI ska fortsätta på samma nivå. Till exempel köpte några av de stora AI-utvecklarna, Google och Microsoft, år 2024 energi från olika kärnkraftverk i USA för sina

egna behov⁹. Ny teknik kräver nya lösningar och ger eventuellt upphov till ökad energiförbrukning. Samtidigt befinner vi oss i en tid som präglas av en ständigt ökande förlust av arter och biologisk mångfald, vilket gör att den typen av negativa konsekvenser uttryckligen borde undvikas.

Lagstiftningen släpar efter på många områden, när det gäller AI: det finns många åsikter om GDPR och de går ofta isär, upphovsrättsliga frågor är olösta¹⁰ och det finns olika ståndpunkter i olika länder. Integritetsskydd och MyData (egna data) måste skyddas effektivt.

På den positiva sidan kan nog nämnas både EU:s satsning på AI-förordningen (AI-akten) och GDPR:s många användningsområden samt det starka engagemanget för att hålla framme de etiska aspekterna¹¹, så att vi trots vårt AI-rus håller oss på körbanan i teknikens alla svängningar¹².

¹ Karmiva löytö TikTok-sovelluksesta: "Poista se viimeistään nyt", hämtad 10.12.2024

² Matteenan blogi: Generatiivinen tekoäly, hämtad 10.12.2024

³ Prompt Library – AI for Education, hämtad 10.12.2024

⁴ <https://tekoalyoppimisentukena.wordpress.com/tot-oppaat/>, hämtad 10.12.2024

⁵ Lagstiftning och rekommendationer för artificiell intelligens inom utbildning | Utbildningsstyrelsen, hämtad 10.12.2024

⁶ Change blindness – by Ethan Mollick – One Useful Thing, hämtad 10.12.2024

⁷ AI Reaches Human-Level Reasoning: Should We Be Worried? | by Fix Your Fin, hämtad 10.12.2024.

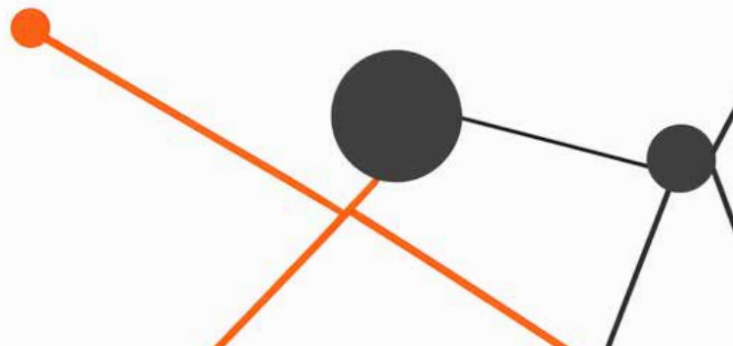
⁸ GenAI opetettava kone → Generation AI, hämtad 10.12.2024

⁹ Why Big Tech is turning to nuclear to power its energy-intensive AI ambitions, hämtad 10.12.2024.

¹⁰ The US Copyright Office says an AI can't copyright its art – The Verge, hämtad 10.12.2024

¹¹ <https://technologymagazine.com/articles/data-privacy-day-data-protection-in-the-age-of-gen-ai%5C>, hämtad 10.12.2024

¹² The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, hämtad 10.12.2024



14. Deepfake härmar verkligheten

PIPSA HAVULA, ANSVARIG REDAKTÖR PÅ FAKTABAARI

MED HJÄLP AV GENERATIV AI är det möjligt att mycket verklighetstroget ändra på en människas ansikte, kropp eller röst så att man kan få henne att göra och säga saker som hon i verkligheten aldrig har gjort eller sagt. Med hjälp av AI är det möjligt att skapa bilder och videor av närmast vad som helst och få dem att se ut som äkta foton eller videor. Den här typen av AI-teknik som härmar verkligheten kallas för deepfake (ibland också djupfejk).

Bakom deepfake ligger djupa neurala nätverk. De är modeller för maskininlärning som imiterar principerna för hur biologiska nervnät fungerar. Den tekniska utvecklingen har gjort stora framsteg under de senaste åren, delvis tack vare GAN-modellen, som presenterades av den amerikanska AI-forskaren Ian Goodfellow och hans team 2014.¹ I modellen har ett neuralt nätverk i uppgift att utvärdera ett annat och att försöka skilja åt människo- och maskinskapat material. På så sätt lär neurala nätverk sig själva.

Målet är att ett generativt nätverk lär sig producera så bra förfälskningar att det utvärderande nätverket inte längre kan skilja dem från äkta material. GAN-modellen har vidareutvecklats till allt kraftfullare deepfake-applikationer, och förfälskningar görs också med andra modeller.

Ofarlig, nyttig och skadlig användning av AI

Deepfake används numera till en uppsjö av olika ändamål från underhållningsindustrin till spridningen av desinformation. Med hjälp av deepfake-tekniken är det möjligt att skapa en digital tvilling av vem som helst, till exempel av en skådespelare som inte längre lever men som man vill ska återvända till vita duken. Tekniken används också i stor utsträckning till exempel på ljudboksmarknaden, där äkta människoröster ersätts av AI-röster som låter lika äkta. I reklambranschen kan det enkelt göras olika versioner av samma reklamfilm på till exempel nio olika språk.

I undervisnings- och inlärningssammanhang kan deepfake-tekniken användas till exempel till att skapa en mer lättsmält AI-podcastversion av en invecklad vetenskaplig artikel (tex. med Googles Notebook LM) eller till att skriva ett tal för Martin Luther King, som en AI-version av King läser upp.

Deepfake-tekniken har också gett upphov till mycket ont. Med hjälp av den kan det utan en persons eget samtycke skapas till exempel en sexvideo med personen, som ser autentisk ut, om det finns annat bild- eller videomaterial tillgängligt av hen. Tekniken används också runt om i världen till att skämma ut människor, speciellt kvinnor och flickor. Falsa videor som ser autentiska ut används till exempel vid mobbning och utpressning av journalister.²

Genom att kлона en människas röst har föräldrar också lurats på pengar, då de ringts upp av personer som låtit som deras egna barn.³ På så sätt har också en medarbetare på ett företag lurats till att överföra en stor summa pengar till ett konto, som pengarna definitivt inte borde ha överförts till,⁴ och på samma sätt har rektorn på en skola i USA kunnat svartmålas.⁵ Det finns gott om sådana exempel.

Deepfake-bilder, -videor och -ljudklipp sprids nuförtiden närmast utan undantag också under stora nyhetshändelser. Vid politiska val har deepfake använts till att smutskasta politiker och till att undergräva förtroendet för valsystemen. Inför presidentvalet i USA 2024 ringde en klonad AI-version av Joe Bidens röst upp människor och uppmuntrade dem att låta bli att rösta.⁶ Inför samma val försökte man måla upp en bild av vicepresidentkandidaten Tim Walz som pedofil bland annat genom att lägga ut en deepfake-video på nätet, där en person som felaktigt skulle framstå som Walz "tidigare elev" talade till kameran och berättade om sina fruktansvärda upplevelser.⁷ Det finns en lång lista också över sådana exempel, och den här typen av förfalskningar har förekommit också i Finland.⁸

För att deepfake ska kunna orsaka skada måste det vara möjligt att sprida det i meddelandetjänster och på sociala medier. Techjättarna, som Meta, Tiktok, Google och Microsoft, har uppmanats att ta ansvar för att begränsa spridningen av deepfake-material.

Identifiering av deepfake

Nuerna kan det vara närmast omöjligt att skilja skickligt gjorda deepfake-förfalskningar från äkta

material. Detsamma kan uttryckas också omvänt: det har blivit allt svårare att bevisa att en äkta video faktiskt också är äkta. Därför har det hänt i många situationer att äkta material har anklagats för att vara förfalskat. Det kan ha långtgående konsekvenser för spridning av korrekt information och för vad människor i slutändan väljer att tro på.

Fram till för ett tag sedan kunde deepfake-förfalskningar kännas igen på att AI gett människor till exempel för många eller för få fingrar eller att den haft svårt att skapa örsnibbar och glasögon som såg äkta ut. Småningom har dock den här typen av små fel kunnat åtgärdas, och förfalskningar som gjorts med bra AI-verktyg går inte längre att skilja från äkta material. Röstförfalskningar är också numera närmast felfria.

De generella instruktionerna nedan kan hjälpa en att utvärdera hur tillförlitlig informationen på nätet är – även i situationer där det är oklart huruvida materialet är deepfake eller inte:

- **Om en bild, en video eller ett ljudklipp på sociala medier ger upphov till en stark känslomässig reaktion är det första man bör göra att ta sig en funderare och därefter en till, innan man väljer att tro på materialet eller sprider det vidare.**
- **Tänk på tre frågor: Vem har lagt ut bilden, videon eller ljudklippet? Vilka bevis har lagts fram för att stödja innehållet? Vad säger andra källor om innehållet?**
- **Utöva lateral läsning: Handlar det om någonting som det också rapporterats om i tillförlitliga källor? Har någon ifrågasatt materialets äkthet i kommentarsfältet? Hittar du andra bilder, videor eller material om samma händelse?**
- **Gör en omvänd bildsökning med bilden eller videon (t.ex. med Google Lens) och undersök vilka sammanhang materialet tidigare har förekommit i.**
- **Var skeptisk speciellt om någon typ av skandalomsusade bilder, videor eller ljudklipp börjar spridas på sociala medier**

precis före ett val eller någon annan stor nyhetshändelse. Var vid behov i kontakt med myndigheter, faktagranskare eller journalistiska medier.

För att lösa problemet har man försökt använda sig av olika typer av identifieringstekniker, som kan hittas på nätet med sökorden "deepfake detector". Faktabaari rekommenderar dock inte att man förlitar sig enbart på dem.

Det finns en pågående kapplöpning mellan utvecklingen av deepfake och av motsvarande identifieringsteknik, där identifieringstekniken alltid är steget efter.⁹ Identifieringsverktygen kan felaktigt tro att äkta material är förfalskat eller att förfalskat material är äkta. Algoritmerna som försöker upptäcka förfalskningar är endast så bra eller dåliga som data som de har tränats med. Identifieringstekniken har också fördomar, och den kan vara dålig på att identifiera deepfake-förfalskningar med mörkhyade personer¹⁰ eller bli förvirrad om den äkta bilden eller videon minskats eller exempelvis suddats ut. Kvaliteten på identifieringsverktygen varierar mycket, och på samma sätt som alla andra AI-verktyg används de också i syfte att tjäna pengar.

Således kan identifieringsverktygen används som en del av en faktagranskningsprocess, men de ska hellre ses som assistenter än experter. Det är bra att försöka ta reda på vilken typ av källmaterial identifieringsverktyget har tränats med, vem det har utvecklats av och vilka begränsningar det har. Utgå aldrig från att identifieringsverktyget gör en sanningsenlig bedömning, utan kontrollera också vad det står om ämnet i andra källor.

Som med all annan teknisk utveckling gör man förnuftigt i att hålla sig uppdaterad om hur identifieringsverktygen utvecklas.

Faktabaari har som mål att informera sina läsare på sin webbplats och i sitt månadsbrev

om verktyg som vi använder i professionellt granskningsarbete.

I början av 2025 har Faktabaari uppdaterat sitt AI-to-spel med nya bildpar med hjälp av den senaste tekniken, för att stödja diskussionen som den allt vanligare deepfake-tekniken gett upphov till. Tanken är att erbjuda lärare och elever möjlighet att testa huruvida de själva kan känna igen AI-genererade bilder bland äkta bilder. Efter egna gissningar är det lättare att ta tag i temat tillsammans med dem som vuxit upp i digitaliseringens tidsålder. Testa spelet här (även på svenska): <https://peli.faktabaari.fi/fi>

Det finns en pågående kapplöpning mellan utvecklingen av deepfake och av motsvarande identifieringsteknik, där identifieringstekniken alltid är steget efter.

¹ se tex. MIT Technology Review (2018). The GANfather: The man who's given machines the gift of imagination, <https://www.technologyreview.com/2018/02/21/145289/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/>

² BBC (2024). Inside the deepfake porn crisis engulfing Korean schools <https://www.bbc.com/news/articles/cpdlpj9zn9go> ; BBC (2024). Girl, 12, victim of 'deepfake' bullying porn image <https://www.bbc.com/news/articles/ckvgez74kgo>; VOA (2024). Journalist 'haunted' by AI deepfake porn video, <https://www.voanews.com/a/journalist-haunted-by-ai-deepfake-porn-video/7624281.html>

³ Guardian (2023). US mother gets call from 'kidnapped daughter' – but it's really an AI scam, <https://www.theguardian.com/us-news/2023/jun/14/ai-kidnapping-scam-senate-hearing-jennifer-destefano>

⁴ Guardian (2024). Company worker in Hong Kong pays out £20m in deepfake video call scam, <https://www.theguardian.com/world/2024/feb/05/hong-kong-company-deepfake-video-conference-call-scam>

⁵ BBC (2024). The racist AI deepfake that fooled and divided a community, <https://www.bbc.com/news/articles/ckg9k5dv1zdo>

⁶ Reuters (2024). Consultant fined \$6 million for using AI to fake Biden's voice in robocalls, <https://www.reuters.com/world/us/fcc-finalizes-6-million-fine-over-ai-generated-biden-robocalls-2024-09-26/>

⁷ Wired (2024). Russian Propaganda Unit Appears to Be Behind Spread of False Tim Walz Sexual Abuse Claims, <https://www.wired.com/story/russian-propaganda-unit-storm-1516-false-tim-walz-sexual-abuse-claims/>

⁸ se tex. Cybersäkerhetscentrets veckoöversikt 03/2024, <https://www.kyberturvallisuuskeskus.fi/fi/ajankohtaista/kyberturvallisuuskeskuksen-ai-kohtaus-032024>

⁹ Binh, Le, Shahroz Tariq, Sharif Abuadbbba m.fl. (2023) Why Do Deepfake Detectors Fail?, https://www.researchgate.net/publication/368843207_Why_Do_Deepfake_Detectors_Fail

¹⁰ University at Buffalo. (2024) New deepfake detector designed to be less biased, <https://www.buffalo.edu/home/story-repository/host.html/content/shared/university/news/ub-reporter-articles/stories/2024/01/lyu-deepfake-bias.detail.html>

15. AI-baserade rekommendationsalgoritmer på sociala medier

HARTO PÖNKÄ, INNOWISE

I DISKUSSIONEN OM AI kommer många först att tänka på chattbottar som ChatGPT och Copilot och bildgeneratorer som baserar sig på AI-modeller. Det man eventuellt inte inser är att vi mycket oftare använder oss av AI i ett annat sammanhang än i form av AI-verktyg: nämligen på sociala medier.

På de populäraste sociala medieplattformarna har maskininlärning åtminstone sedan 2016 använts för att få algoritmer att ge rekommendationer. De tidigaste försöken med rekommendationssystem gjordes redan flera år tidigare.

Olika typer av rekommendationsalgoritmer

Rekommendationssystem används i många informationssystem och applikationer. På olika sociala medieplattformar kallas de ofta rekommendationsalgoritmer. Det är fråga om program som genom att gå igenom stora datamängder skapar rekommendationer för specifika användare. Oftast handlar det om att välja ut innehåll för användarens flöde på en viss plattform på sociala medier. I praktiken är det alltså rekommendationsalgoritmerna som bestämmer vilken typ av inlägg och i vilken ordning användaren får se dem på sociala medieplattformar och -applikationer.

Rekommendationssystemen kan delas in i tre huvudgrupper: innehållsbaserade och kollaborativa system samt hybridmodeller, som kombinerar flera metoder.^{1 2}

- **En innehållsbaserad algoritm skapar rekommendationer genom att jämföra den information som finns tillgänglig om både användaren och olika typer av innehåll.** En applikation kan exempelvis fråga användaren vilka teman hen intresserar sig för och anpassa innehållet utifrån ett antal utvalda kategorier.
- **En kollaborativ rekommendationsalgoritm skapar rekommendationer utifrån data som samlats in om användarens aktivitet.** Tanken är att innehållet filtreras utifrån information om flera användare istället för en enskild användare och att de som tidigare intresserat sig för vissa teman gör det också senare. Data som ligger till grund för filtreringen kan till

exempel bestå av användarnas reaktioner på inlägg eller deras sociala nätverk.

- **Hybridalgoritmerna kombinerar dessa två tillvägagångssätt och andra metoder. En algoritm kan till exempel rekommendera innehåll till nya användare utifrån var de befinner sig geografiskt eller vilka intressen de har och sedan gå över till kollaborativa rekommendationer allt eftersom det samlas in data om användaren. Dessa data kan i sin tur jämföras med data om andra användare och annat innehåll.**

Eftersom användarnas intressen, beteendemönster och aktivitet varierar har algoritmerna utvecklats så att de allt bättre kan anpassa sig efter användarnas vanor och skillnaderna mellan olika användare. Detta återspeglas också i hybridalgoritmernas ökade popularitet. Till exempel behövs det betydligt större antal och nyare rekommendationer för dem som är väldigt aktiva

på sociala medier än för dem som tittar in i apparna endast några gånger i veckan.

Dagens webbtjänster och sociala medieplattformar har som mål att samla in data som ger en så detaljerad bild av en användares aktivitet som

Det behövs nyare rekommendationer för dem som är aktiva på sociala medier än för dem som öppnar en app endast några gånger i veckan.

möjligt och som rekommendationsalgoritmerna kan använda sig av för att så bra som möjligt filtrera innehållet utifrån användarens intressen. Det har i sin tur inneburit att det samlas in allt mer olika typer av personuppgifter och data som beskriver användarnas aktivitet.

I takt med att mängden användardata växer och data präglas av större komplexitet har det för olika applikationer utvecklats också AI-baserade rekommendationsalgoritmer. Fördelen med AI-baserade algoritmer är att de effektivt kan dra nytta av ett stort användarspecifikt datalager och anpassa sig till varje användares beteende. Således kan olika faktorer betonas i

rekommendationerna beroende på vilken typ av data det finns om en användare. Det här är en tydlig fördel jämfört med icke-flexibla rekommendationsalgoritmer som behandlar alla användares data på samma sätt.

De nuvarande rekommendationsalgoritmerna använder sig ofta av AI-tekniker som maskininlärning, djupinlärning och naturlig språkbehandling, som används blanda annat i chattbotar. Å andra sidan kan användare trots användningen av AI uppleva att algoritmerna på sociala medier inte visar rätt sorts innehåll för dem.

Det är bra att vara medveten om att algoritmerna i webbtjänster och på sociala medier ändras ofta. Ändringarna beror på att algoritmerna utvecklas, eftersom sociala medieplattformar strävar efter att optimera algoritmerna så att de bättre tjänar sina syften. I takt med att AI utvecklas är målet till exempel att i allt högre grad utnyttja AI-funktioner i rekommendationsalgoritmerna.

Oavsett om AI används eller inte är det bra att komma ihåg två saker som inte ser ut att ändras: för det första är syftet med algoritmerna oftast att hålla användarna kvar på plattformar så länge som möjligt, och för det andra är algoritmerna lätta att manipulera. Till exempel annonsörer, inflytelserika personer inom politiken och de som sprider desinformation håller sig ständigt uppdaterade om utvecklingen av algoritmerna och försöker utnyttja dem för att ge sina egna inlägg så stor synlighet som möjligt.³

Facebook och filtrering av sociala nätverk

Pionjären bland AI-algoritmerna – på gott och ont – har varit Facebook. Facebook lanserade nyhetsflödet ('news feed') två år efter att själva plattformen lanserats år 2006. Till en början baserade sig flödet inte på en rekommendationsalgoritm, utan det visade allt ens Facebook-vänner hade haft för sig i kronologisk ordning.

Facebook lanserade sin första algoritm, som var avsedd för att styra nyhetsflödet, år 2009.

Händelsen kan så här i efterhand ses som startskottet för rekommendationsalgoritmerna på sociala nätverkstjänster. Tanken var att visa användarna det mest intressanta innehållet istället för deras vänners inlägg. Algoritmen baserade sig i huvudsak på tre variabler: hur nära relation det var mellan användaren och inlägget, hur viktigt innehållet ansågs vara och när inlägget lagts ut. Således var det alltså fråga om en kollaborativ rekommendationsalgoritm, som i hög grad beaktade det sociala nätverket mellan användarna. Senare har den här typen av algoritmer blivit vanliga uttryckligen på plattformar som är avsedda för nätverkande.^{4 5 6}

Den nya algoritmen innebar en betydande förändring för Facebooks dåvarande 300 miljoner användare. Eftersom användarna inte längre "behövde" läsa alla sina vänners inlägg blev det möjligt att inkludera allt fler personer i det egna nätverket. År 2009 var det genomsnittliga antalet Facebook-vänner 120, men 2011 hade siffran stigit till 190. Vid den tiden gjordes en annan stor ändring i algoritmen för nyhetsflödet, då en rekommendationsalgoritm som baserade sig på maskininlärning togs i bruk.⁷

År 2013 använde den nya AI-baserade rekommendationsalgoritmen sig av upptill 100 000 olika variabler. Detta innebar allt mer individanpassade rekommendationer: Användarna som var aktiva i Facebook-grupper såg fler inlägg från grupperna. De som ofta gillade bilder såg allt oftare bildinlägg. De som klickade på länkar fick i sin tur se allt fler länkar och så vidare.⁸

När AI gallrade fram de intressantaste inläggen i det sociala nätverket kunde användarna nätverka med allt fler människor, utan att nyhetsflödet fylldes med ointressant "brus". År 2016 hade användarna i USA i genomsnitt redan 350 Facebook-vänner och 18-24-åringar hade så många som 650 Facebook-vänner.

Under de första åren hade Facebook varit en plattform för kommunikation mellan bekanta, men ökningen av antalet vänner innebar en möjlighet att få mer och mer publicitet. Betydelsen av länkar som delats på Facebook blev större än

tidigare för kommersiella aktörer, eftersom organisk synlighet på Facebook kunde innebära tusentals och åter tusentals besökare på företagens webbplatser. Vid sidan av viljan att ta hand om sina vänskapsrelationer kunde Facebook också användas i försök att uppnå masspopularitet och att skapa politiskt inflytande samt att främja företagsverksamhet. Facebooks egna medarbetare beskrev förändringen som en kontextkollaps.⁹

År 2016 lanserade Facebook emoji-reaktionerna «älskar», «haha», «wow», «ledsen» och «arg». De tillät algoritmen att lista ut användarnas känsloreaktioner och rekommendera inlägg till användarna utifrån dem. Samtidigt blev den AI-baserade hybridalgoritmen så invecklad att den visade sig vara svår för Facebook att hantera.

Det fanns ett stort problem med hur rekommendationsalgoritmen påverkades av emoji-reaktionerna. Emoji-reaktionerna tilldelades mycket större tyngd i hur algoritmen beräknade rekommendationspoäng jämfört med vanliga gilla-markeringar. Det ledde till att inlägg som fick ett stort antal till exempel "arg"- och "haha"-reaktioner kunde få oproportionerligt stor synlighet, när algoritmen lyfte fram dem i användarnas flöden. Att väcka ilska och andra starka känslor blev på så sätt ett effektivt sätt att ge inlägg stor publicitet på Facebook. Detta gick inte obemärkt förbi hos dem som hade för avsikt att påverka människors attityder och åsikter.^{10 11}

Det tog Facebook flera år att erkänna och rätta till emoji-reaktionernas oproportionerligt stora påverkan på rekommendationsalgoritmen. Enligt dokument som läcktes av en tidigare Facebook-anställd 2021 hade Facebook försökt förhindra känsloreaktionernas alltför stora inverkan på rekommendationerna, men funktionen som utvecklats för ändamålet visade sig inte fungera. Det är talande att algoritmen för nyhetsflödet kunde ge inlägg med till exempel desinformation, hatretorik och klickrubriker stor synlighet, samtidigt som bolagets moderatorer efter bästa förmåga

Att väcka ilska och andra starka känslor blev ett effektivt sätt att ge inlägg stor publicitet på Facebook.

försökte bekämpa den typen av innehåll.

Forskare upptäckte att emoji-reaktionerna användes till att sprida felaktigt och skadligt innehåll på Facebook 2019. "Arg"-reaktionernas betydelse för rekommendationsalgoritmen minskades flera gånger, och till sist slopades metoden helt och hållet 2021. Detta gjordes efter att Facebook blivit föremål för hård offentlig kritik i efterdyningarna av presidentvalen i USA 2016 och 2020. Bolaget anklagades för att ha misslyckats med att bekämpa spridningen av desinformation och informationspåverkan inför valen. Känslolokaliserade Facebook-inlägg användes effektivt också i kampanjerna inför riksdagsvalet i Finland 2019.¹²

Sedan 2021 har Facebook fortsatt att utveckla sin AI-drivna rekommendationsalgoritm. I takt med att regleringen med fokus på integritetsskydd har ökat har Facebook regelbundet börjat ge ut information om vilken typ av inlägg algoritmen väljer ut för en genomsnittlig Facebook-användares flöde. Under de senaste åren har Facebook särskilt minskat synligheten för länkar i flödet och ökat antalet sådana inlägg som lagts ut av aktörer som inte är bland användarnas Facebook-vänner. Ändringarna innebär att Facebook försöker hålla användarna kvar på sin egen plattform, och bolaget går stegvis mot intressefokuserade rekommendationer som skapats med hjälp av profilering, vilket gör att användarnas sociala nätverk har allt mindre betydelse.¹³

Youtubes dilemma: Vilket av alla miljontals videor ska rekommenderas härnäst?

Youtube har ett uppenbart behov av en effektiv rekommendationsalgoritm, eftersom det år 2008, bara tre år efter lanseringen, hade laddats upp över 45 miljoner videor på plattformen. Det var möjligt att upptäcka videor enligt kanal samt med hjälp av en popularitetsbaserad lista och en textbaserad sökfunktion. Sökfunktionen fungerade dock inte särskilt bra, eftersom det enda giltiga sökkriteriet i de flesta videor var videons titel.¹⁴

Youtubes behov av att utveckla en rekommendationsalgoritm kan närmast jämföras med tävlingen som Netflix sponsrade under åren 2006–2009 och som hade som mål att ta fram ett så exakt rekommendationssystem för filmer som möjligt.¹⁵ Netflix hade börjat hyra ut DVD-filmer per post, och dess databas bestod vid den tiden av cirka 100 000 DVD-filmer. Bolaget utlovade en miljon dollar till den som kunde utveckla ett rekommendationssystem som skulle vara minst 10 procent bättre än algoritmen som användes då. Jämfört med Netflix hade Youtube ett mycket större urval av videor och mindre information om användarna.

Google började ta hjälp av maskininlärning för att ge videorekommendationer på Youtube år 2007. Bolagets egna forskare skapade till sist ett nätverk av videorna utifrån hur många gånger samma användare hade tittat på samma videor. Förmodligen var det förnuftigast att rekommendera videor som användare som tidigare visat intresse för samma videor hade tittat på. Utmaningen var det stora antalet videor, vilket gjorde det omöjligt att beräkna nya videorekommendationer för varje användare varje gång de bytte mellan olika videor. Å andra sidan skulle användarna inte bara rekommenderas de populäraste videorna, utan tanken var att de skulle kunna upptäcka nytt intressant innehåll.

Utvecklarna på Youtube löste utmaningen genom att skapa flera AI-modeller, använde sig av loggdata, som samlats in om användarna, som träningsdata för maskininlärning. Med hjälp av AI-modellerna kunde användarna erbjudas förslag på flera "relaterade videor" ('related videos'), som tagits fram utifrån videorna de precis tittat på. På så sätt skapades en av de tidiga rekommendationsalgoritmerna för sociala medier, som baserade sig på kollaborativ filtrering. Till skillnad från Facebook baserades rekommendationerna inte på användarnas sociala nätverk utan på ett nätverk av videor.

Youtube har varit pionjär även i utvecklingen av andra AI-funktioner än rekommendationsalgoritmer. År 2009 lanserade Youtube automatiska

undertexter, som baserade sig på taligenkänning, och år 2013 en funktion för ansiktsgenkänning för identifiering av människors och djurs – exempelvis katters – ansikten.^{16 17}

Det är oklart hur länge Youtubes tidigare AI-algoritm användes i videorekommendationer. År 2016 introducerades ändå en mer avancerad rekommendationsalgoritm som baserade sig på djupinlärning och neurala nätverk. Den nya AI-baserade algoritmen möjliggjorde effektivare innehållsmoderering till exempel för automatisk identifiering av innehåll som var olämpligt eller på något annat sätt skadligt för barn. Sedan 2017 har Youtube arbetat för att få algoritmen att lyfta fram innehåll som producerats av pålitliga medier. År 2019 vidareutvecklade Youtube algoritmen så att videor med tvivelaktiga påståenden rekommenderas i mindre utsträckning.^{18 19}

Tiktoks beroendeframkallande algoritm

Jämfört med flera andra sociala medieplattformar är Tiktok relativt nytt. Tiktoks systerapp Douyi lanserades i Kina 2016 och Tiktok i andra länder året därpå. I Finland ökade Tiktok i popularitet under åren 2019–2021.²⁰

Tiktoks moderbolag Bytedance är först och främst ett AI-bolag. En av bolagets första

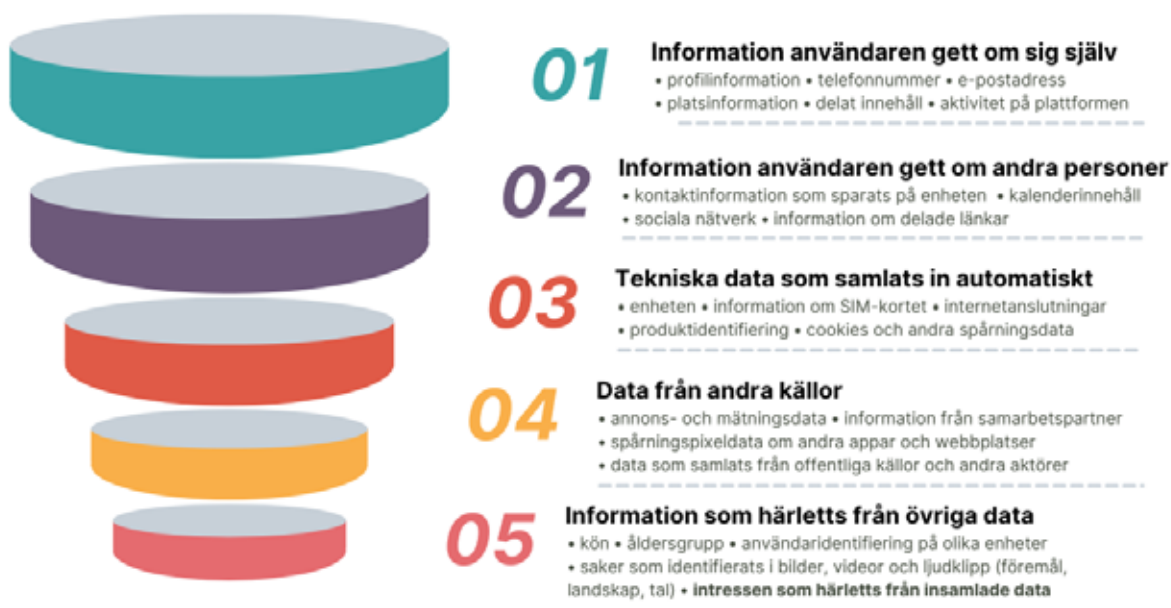
produkter år 2012 var appen Toutiao, som rekommenderade nyheter till kineserna med hjälp av maskininlärning och en användarspecifik, individanpassad rekommendationsalgoritm, som baserade sig på naturlig språkbehandling. År 2015 blev det möjligt att ladda upp korta videor i appen, och året därpå var Toutiao redan den mest populära plattformen för korta videor i Kina.²¹

Till skillnad från många andra sociala medieplattformar hade Bytedance från början tillgång till en kraftfull rekommendationsalgoritm, när bolaget lanserade Douyi och Tiktok. Det är värt att notera att "det nya" med Tiktok inte var de korta videorna som innehållsformat utan uttryckligen rekommendationsalgoritmen, som klarade av att fånga användarnas uppmärksamhet. Vid den tiden fanns det redan många populära appar för korta videoklipp, som Vine (2012–2017) och Snapchat, där *snappar* i videoform lanserades 2012 och *stories* 2013.^{22 23}

Tiktok kombinerade offentliga korta videoklipp med en AI-algoritm, som utifrån insamlade profileringsdata rekommenderade användarna intressanta videoklipp.

Kärnan i Tiktok är För dig-flödet (For you page), som består av algoritmens rekommendationer. Ett typiskt sätt att använda appen är att snabbt bläddra i flödet och titta på intressanta videor.

Nivåer för datainsamling på Tiktok



Eftersom beslutet att titta på en video fattas utifrån startbilden, titeln och de första sekunderna av videon är det av stor betydelse hur visuellt intressanta och uppseendeväckande videorna är. I enlighet med detta utgår algoritmen på Tiktok, så som på många andra sociala medier, inte från vilka profiler en användare valt att följa utan från vilka teman och vilken typ av videor användaren tidigare varit intresserad av.

Enligt Tiktok inverkar bland annat följande faktorer på vilka videor som lyfts fram i För dig-flödet:^{24 25}

- **Användarens interaktion: videorna som användaren gillat, delat, kommenterat, tittat på eller hoppat över samt profilerna hen börjat följa.**
- **Videodata: videoljud, hashtaggar, beskrivningstexter, visningar och publiceringsland.**
- **Användardata: enhetsinställningar, språk, plats, tidszon och datum och enhetstyp.**

Jämförelser med andra sociala medieplattformar har visat att datainsamlingen på Tiktok är mycket omfattande. Datainsamlingen på Tiktok kan delas in i fem nivåer enligt följande:²⁶

1. **Information som användaren gett om sig själv**
2. **Information som användaren gett om andra personer**
3. **Tekniska data som samlats in automatiskt**
4. **Data från andra källor**
5. **Information som härletts från övriga data**

Genom att jämföra data som samlats in om användaren och videodata görs det för varje video en bedömning av hur sannolikt användaren intresserar sig för den. Olika variabler viktas på olika sätt. Om användaren till exempel ser en video från början till slut indikerar det starkt att användaren intresserar sig för videor om ett visst ämne. En stor följarskara garanterar inte enbart att en innehållsskapares videor kommer före

andras i rekommendationssystemet. Användarens sociala nätverk, dvs. profilerna hen följer, har inte någon större betydelse för algoritmen på Tiktok. Algoritmen på Tiktok kan kallas ett rekommendationssystem som styrs av ett nätverk av intressen.

Tiktok-algoritmen är mycket reaktiv: användarens aktivitet på plattformen påverkar snabbt vilka videoklipp som visas för hen. Det gör i sin tur att Tiktok är ytterst beroendeframkallande för sina användare. Enligt interna dokument som läckts ut i offentligheten i samband med rättsprocesserna mot Tiktok i USA 2024 kunde användarna bli beroende av appen på 35 minuter och hamna i en algoritmiskapad filterbubbla så snabbt som på 20 minuter.²⁷

Korta videor intresserar särskilt unga

Tiktoks globala popularitet har bestämt riktningen för utvecklingen av sociala medier under de senaste åren. Precis som på många andra sociala medieplattformar berodde ökningen av antalet användare på Tiktok till en början särskilt på barn och unga, som snabbt tog till sig den nya appen.

Innan det blev populärt med Tiktok var de populäraste sociala medieplattformarna i Finland Whatsapp, Youtube, Snapchat och Instagram.Whatsapps och Snapchats popularitet har i huvudsak baserat sig på direktmeddelanden mellan användarna och grupper, medan Youtube blivit populärt på video- och Instagram på bildinnehåll. De korta klippen på Tiktok erbjöd användarna en ny typ av innehåll, jämfört med tidigare plattformar, och appen blev snabbt populär bland finländska ungdomar fr.o.m. 2019. Så tidigt som 2021 hade Tiktok redan blivit populärare än Instagram och lika populärt som Snapchat bland 13–16-åringarna. Tiktoks inträde på marknaden verkar trots det inte ha minskat ungdomarnas användning av andra populära plattformar.²⁸

Tiktoks popularitet har inte gått obemärkt förbi hos de övriga jättarna på sociala medier. Meta lanserade korta Reels-videor på Instagram

i augusti 2020 och Shorts-videor på Youtube i juli 2021. Bland videoplattformar kan Tiktok ses som den största utmanaren för just Instagram och Youtube – särskilt bland unga användare. I ljuset av det är det lätt att förstå att de också vill erbjuda samma slags korta videoklipp som Tiktok.

Plattformarna för korta videoklipp har förändrat och kommer även framöver att förändra användarnas beteende på sociala medier. Det här återspeglas bland annat i att korta videor på bara några år blivit ett mer populärt format för nyhetsinnehåll i många länder än långa videor eller direkta videosändningar. Enligt en studie från Reuters Institute (2024) tar människor i hela världen mest del av nyhetsvideor på Youtube, Facebook, Instagram och Tiktok. Bland personer under 35 år är Instagram och Tiktok särskilt populära. Tiktoks betydelse som nyhetskanal är tydlig också i länderna som ingår i det s.k. globala syd.²⁹

När det gäller nyhetsinnehåll aktualiseras frågan om huruvida de olika nyhetskällorna är lika pålitliga. Enligt ett flertal enkätundersökningar upplever användarna på Tiktok nyheterna de tar del av via plattformen som opålitliga. I en enkätundersökning gjord av Reuters Institute sade flest svarande att det var svårt att försäkra sig om nyheternas tillförlitlighet på Tiktok. Följande tre på listan var X, Facebook och Instagram. Resultaten från Finland motsvarade ungefär genomsnittet för de länder som var med i undersökningen.³⁰

Effekterna av Tiktok och boomen för korta videor har snabbt börjat märkas i Finland. Enligt en enkätundersökning från början av 2024 hade Tiktok blivit den populäraste nyhetskällan bland 13-18-åringar och 33 % ansåg att de fick ta del av allt nödvändigt nyhetsinnehåll via plattformen. De två följande nyhetskällorna på listan var kvällstidningarnas webbplatser och Instagram. Majoriteten av de svarande (56 %) ville helst ta del av nyheter i form av korta videor.³¹

Det är fråga om en snabb förändring, som leder till att ungas medieanvändning skiljer sig markant från medieanvändningen hos majoriteten av de vuxna. Det finns anledning till oro, eftersom

en relativt stor andel av Tiktok-användarna (35 %) inte kunde säga om de stött på desinformation³² på plattformen.

Det indikerar både att behovet av digital informationslitteracitet bland unga användare på sociala medier har ökat och att det är svårt att upptäcka desinformation på en plattform för kort videoinnehåll.

Datastyrda algoritmer får allt större makt

Även om den mest synliga effekten av Tiktok är den ökade populariteten av kort videoinnehåll är det bra att komma ihåg att framgången inte är de korta videorna i sig utan en kraftfull rekommendationsalgoritm. Därtill handlar det om i hur hög grad algoritmens innehållsrekommendationer styr användningen av plattformen.

Tiktoks – och tidigare också Youtubes – framgång har visat att en AI-driven rekommendationsalgoritm som baserar sig på profilering och ett nätverk av intressen kan vara effektivare på att fånga användarnas uppmärksamhet än en algoritm som är baserad på sociala nätverk och kollaborativ filtrering. För Tiktoks algoritm spelar användarnas sociala relationer ingen större roll. Det här utgör en betydande skillnad jämfört med algoritmerna för Facebook, Instagram och X (tidigare Twitter), eftersom de ursprungligen baserats på just användarnas sociala nätverk.

Meta har varit öppen om utvecklingen av sina rekommendationsalgoritmer. I konkurrensen med Tiktok har Meta under de senaste åren satsat på användning av AI bland annat i algoritmen för Reels på Instagram. I början av 2021 meddelade Meta att bolaget för dessa Reels utvecklat en AI-modell som baserar sig på kontinuerligt självstyrt lärande och på analyser av ljud och bild i videor.

Resultatet av bild- och taligenkänning kombineras med analysen av videornas titlar och beskrivningar, och utifrån dessa analyser kan AI-modellen ge allt exaktare videorekommendationer, som är i linje med användarens intressen. I

likhet med Tiktok samlar även Instagram in en stor mängd profileringsdata, dvs. signaler om användarens aktivitet, för att skapa en bild av användarens preferenser.^{33 34}

Meta har också ökat användningen av profileringsdata i algoritmen för Facebook-flödet. Det har lett till att allt fler av de inlägg som algoritmen väljer för en användares flöde är så kallade orelaterade inlägg, som alltså har lagts ut av innehållsskapare som användaren inte följer. Så sent som 2021 hade fortfarande i genomsnitt ca 12 % av alla inlägg i en användares flöde lagts ut av innehållsskapare som hen inte följde, men 2024 hade siffran stigit till ca 32 %. Samtidigt får användarna se allt färre av sina Facebook-vänners inlägg i sitt flöde. Med andra ord har användarnas egna val och sociala relationer allt mindre inverkan på vilka inlägg de ser på Facebook.³⁵

En intressant fråga är i vilken utsträckning algoritmerna på sociala medier uppfyller användarnas behov. Ett svar på frågan ges av en studie genomförd av Taylor och Choi, som undersökte hur väl innehållet på olika sociala medieplattformar motsvarade användarnas personliga önskemål. De som deltog i studien utvärderade hur väl algoritmerna på Facebook, Instagram, Twitter och Tiktok fungerade utgående från 15 påståenden.

Åtta av påståendena berörde algoritmernas önskade beteende och sju deras oönskade beteende. Enligt resultaten ansågs innehållet som valts ut av Tiktoks algoritm oftare främja och motsvara studiedeltagarnas önskemål än algoritmerna på Facebook, Instagram och Twitter. Facebooks algoritm upplevdes mest ge rekommendationer som inte motsvarade användarnas önskemål. Studien genomfördes under åren 2021–2022, dvs. innan Facebook ökade andelen inlägg publicerade av innehållsskapare som användarna inte följde bland algoritmens rekommendationer.³⁶ Å andra sidan är det bra att notera att deltagandet i en studie kring algoritmernas beteende förutsätter rätt avancerad algoritmisk litteracitet och en medvetenhet om hur algoritmer fungerar, för att deltagarna ska kunna

jämföra algoritmernas beteende i olika appar.

Trenden i utvecklingen av rekommendationsalgoritmer under de senaste åren kan sammanfattas så att datastyrda algoritmer används i en allt större utsträckning jämfört med socialt styrda algoritmer. Det kan i sin tur inom en snar framtid förväntas innebära ytterligare förändringar i hur sociala medier används. Samtidigt kommer datastyrda AI-baserade algoritmer att ha en allt starkare inverkan på miljarder användares beteende på sociala medier, på både individ- och gruppnivå.³⁷

AI-baserade hyperpersonaliserade rekommendationsalgoritmer leder till att algoritmer – snarare än användarna själva – har allt större makt över vad vi ser på sociala medier eller andra digitala plattformar. När allt kommer omkring utvecklas algoritmerna utifrån plattformägarnas, inte användarnas, önskemål. Å andra sidan tyder resultaten av den ovannämnda studien om algoritmerna på sociala medier på att datastyrda algoritmer bättre kan tillgodose användarnas behov än socialt styrda algoritmer.

Algoritmerna är svåra att reglera

Effekterna av digitala plattformar är inte begränsade till själva användarupplevelsen. Samtidigt är det nödvändigt att bedöma till exempel hur datainsamlingen på plattformarna påverkar användarnas integritet, hur sanningsenligt innehåll algoritmerna rekommenderar, hur väl plattformarna klarar av att bekämpa skadligt innehåll och hatretorik och i slutändan hur plattformarna påverkar användarnas välbefinnande som helhet. Många EU-förordningar tangerar också detta, såsom den allmänna dataskyddsförordningen (GDPR), förordningen om digitala tjänster (DSA) och AI-förordningen (AI-akten). Dataskyddsförordningen, som varit i kraft sedan 2018, har påverkat hur de digitala jättarna samlar in och använder sig av användarnas data. Till exempel Meta och Tiktok har ålagts GDPR-sanktioner till ett värde av hundratals miljoner euro på grund av brister i behandlingen av barns personuppgifter.³⁸

Förordningen om digitala tjänster har varit i kraft sedan februari 2024, och den har gett användarna av digitala tjänster ett flertal nya rättigheter. Enligt förordningen är det förbjudet att rikta reklam innehåll utifrån specifika personuppgifter som hälsouppgifter, politiska åsikter, etniskt ursprung eller övertygelser av olika slag. Därutöver är det förbjudet att rikta annonser till minderåriga baserat på profilering.³⁹

Förordningen förutsätter att webbtjänster öppet förklarar hur deras rekommendationssystem fungerar. Och mycket riktigt finns det en hel del information tillgänglig om hur algoritmerna på olika sociala medieplattformar fungerar. Samtidigt är det också ett faktum att informationen i många avseenden är ofullständig och slumpmässig. Till exempel Meta ger ut en stor mängd information om hur algoritmerna fungerar, men informationen om exempelvis viktningarna av eller de faktiska innehållsvolymer för Instagram- och Threads-flöden är rätt knapphändig. För dem som undersöker algoritmerna utgör dataläckor, studier och praktiska experiment fortfarande betydligt bättre informationskällor.

Förordningen om digitala tjänster innebar att stora webbplattformar blev skyldiga att erbjuda ett alternativ till rekommendationssystemet, som inte baserar sig på användarprofilering. Tabellen intill ger en översikt över hur algoritmerna

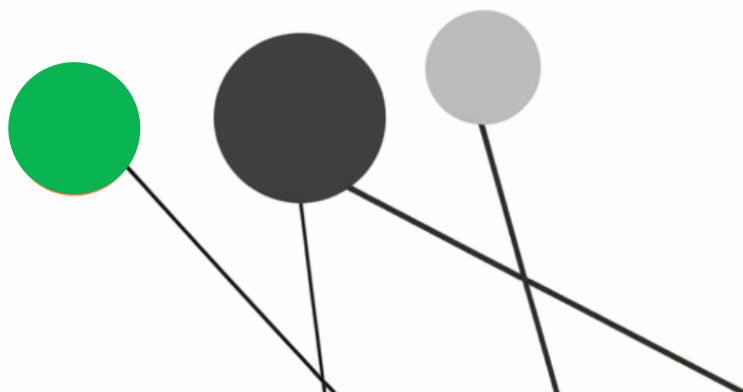
på Facebook, Instagram, Youtube, Tiktok och X normalt fungerar samt över alternativen till den profileringsbaserade algoritmen, som plattformarna erbjuder.

Till exempel på Youtube går det fortfarande relativt lätt att styra innehållet genom att följa intressanta innehållsskapare. Detsamma gäller Facebook, Instagram och X: på alla dessa plattformar är det möjligt att se vad ens eget sociala nätverk nyligen har publicerat. Däremot erbjuder Tiktok ett rätt märkligt alternativ till För dig-flödet: användaren får då se videor som generellt är populära inom hens region.⁴⁰ I praktiken går Tiktoks alternativ ut på att erbjuda ett slumpmässigt flöde, som knappast intresserar någon. På så sätt uppfyller den kraven i DSA, men nyttan för användarna är närmast obefintlig.

Tiktok är också i andra avseenden ett varnande exempel på en app som offentligt säger sig uppfylla kraven i förordningen om digitala tjänster men som i praktiken inte erbjuder sina användare de fördelar som förordningen haft som syfte att uppnå. Tiktok har bland annat haft stora brister i att ta bort olagligt innehåll och blockera minderåriga användare. Dessutom är kraven på transparens i förordningen inte till någon större nytta, om det inte går att försäkra sig om sanningshalten i bolagets rapporter.⁴¹

	Facebook	Instagram	Youtube	Tiktok	X
Andelen algoritm-rekommenderade inlägg som härstammar från konton användaren följer	Ca 68 %	60-70 %	Liten	Liten	50 %
Algoritmens rekommendationer från konton användaren inte följer	Ca 32 %	30-40 %	Upp till 100 %	Upp till 100 %	50 %
Införande av maskininlärning i algoritmens rekommendationer	2011	2016	2016	2016	2016
Alternativt flöde	Inlägg i kronologisk ordning från konton användaren följer	Inlägg i kronologisk ordning från konton användaren följer	Videor i kronologisk ordning från kanaler användaren prenumererar på	Populärt innehåll inom användarens region	Inlägg i kronologisk ordning från konton användaren följer

- 1 Wikipedia. (28.12.2024). Recommender system, https://en.wikipedia.org/w/index.php?title=Recommender_system&oldid=1265732754
- 2 Wikipedia. (12.12.2024). Collaborative filtering, https://en.wikipedia.org/w/index.php?title=Collaborative_filtering&oldid=1262596993
- 3 TikTok (19.6.2020). How TikTok recommends videos #ForYou, <https://newsroom.tiktok.com/en-gb/how-tiktok-recommends-videos-foryou-eu>
- 4 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 5 Facebook (27.4.2009). Facebook press room, tillgängligt via Archive.org-portalen: <https://web.archive.org/web/20090427160416/http://www.facebook.com/press/info.php?statistics>
- 6 Wallaroo. (2024). Facebook News Feed Algorithm History, hämtad 9.12.2024, <https://wallaroomedia.com/facebook-newsfeed-algorithm-history/#one>
- 7 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 8 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 9 Pönkä, H. (31.8.2016). Facebook meni rikki, <https://harto.wordpress.com/2016/08/31/facebook-meni-rikki/>
- 10 Pönkä, H. (31.10.2021). Infografik: Facebookin viha-reaktio ja algoritmin muutokset, <https://harto.wordpress.com/2021/10/31/infografikka-facebookin-viha-reaktio-ja-algoritmin-muutokset/>
- 11 Owen, L. (15.4.2019). One year in, Facebook's big algorithm change has spurred an angry, Fox News-dominated — and very engaged! — News Feed, <https://www.niemanlab.org/2019/03/one-year-in-facebooks-big-algorithm-change-has-spurred-an-angry-fox-news-dominated-and-very-engaged-news-feed/>
- 12 Knuutila, A. & Laaksonen, S-M. (2020). Viraali vihaiisuus ja tahmea nauru: tunteet ja algoritmit digitaalisessa vaalikampanjoinnissa, i: S Borg, E Kestilä-Kekkonen & H Wass (red.), Poliittikan ilmastomuutos: Eduskuntavaalitutkimus 2019. Justitieministeriets publikationer. Utredningar och anvisningar, Nr 2020:5, Justitieministeriet, <https://julkaisut.valtioneuvosto.fi/handle/10024/162429>
- 13 Meta. (2024). Facebook Transparency Center, Widely viewed content report: What People See on Facebook, Q3 2024, hämtad 11.12.2024, <https://transparency.meta.com/fi-fi/data/widely-viewed-content-report/>
- 14 Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D. & Aly, M. (2008). Video suggestion and discovery for youtube: taking random walks through the view graph, <https://static.googleusercontent.com/media/research.google.com/fi/pubs/archive/34407.pdf>
- 15 Wikipedia. (28.12.2024). Recommender system, https://en.wikipedia.org/w/index.php?title=Recommender_system&oldid=1265732754
- 16 Youtube (4.12.2009). Automatic Captioning in YouTube, <https://research.google/blog/automatic-captioning-in-youtube/>
- 17 Le, Q. V. (2013). Building high-level features using large scale unsupervised learning, <https://ieeexplore.ieee.org/document/6639343>
- 18 Youtube (15.9.2021). On YouTube's recommendation system, <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>
- 19 Youtube. (n.d.). Videorekommendationer: Hur fungerar Youtubes system för rekommendationer?, hämtad 11.12.2024, https://www.youtube.com/intl/ALL_se/howyoutubeworks/product-features/recommendations/#signals-used-to-recommend-content
- 20 Wikipedia. (31.12.2024). TikTok, <https://en.wikipedia.org/w/index.php?title=TikTok&oldid=1266416079>
- 21 Wikipedia. (7.9.2024). Toutiao, <https://en.wikipedia.org/w/index.php?title=Toutiao&oldid=1244468033>
- 22 Wikipedia. (26.12.2024). Snapchat, <https://en.wikipedia.org/w/index.php?title=Snapchat&oldid=1265349499>
- 23 Wikipedia. (21.11.2024). Vine (service), [https://en.wikipedia.org/w/index.php?title=Vine_\(service\)&oldid=1258710466](https://en.wikipedia.org/w/index.php?title=Vine_(service)&oldid=1258710466)
- 24 TikTok (19.6.2020). How TikTok recommends videos #ForYou, <https://newsroom.tiktok.com/en-gb/how-tiktok-recommends-videos-foryou-eu>
- 25 TikTok (n.d.). Så här rekommenderar TikTok innehåll, hämtad 15.12.2024, <https://support.tiktok.com/sv/using-tiktok/exploring-videos/how-tiktok-recommends-content>
- 26 Pönkä, H. (29.4.2024). Uusi selvitys TikTakin riskeistä ja raportti TikTakin Kiina-yhteyksistä, <https://harto.wordpress.com/2024/04/29/uusi-selvitys-tiktokin-riskeista-ja-raportti-tiktokin-kiina-yhteyksista/>
- 27 NPR. (11.10.2024). TikTok executives know about app's effect on teens, lawsuits documents allege, <https://www.npr.org/2024/10/11/g-si-27676/tiktok-redacted-documents-in-teen-safety-lawsuit-revealed>
- 28 DNA. (n.d.). koululaistutkimukset vuosilta 2017–2024, hämtad 28.12.2024, <https://corporate.dna.fi/medialle/dnan-tutkimukset/>
- 29 Reuters Institute. (2024). Digital News Report 2024, <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- 30 Reuters Institute. (2024). Digital News Report 2024, <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- 31 Uutismedian liitto. (29.1.2024). Uutisten viikon tutkimus: Tiktok on noussut nuorten tärkeimmäksi uutiskanavaksi, kiinnostus omaan elämään liittyviin uutisiin kasvussa, <https://www.uutismediat.fi/ajankohtaista/uutisten-viikon-tutkimus-tiktok-on-noussut-nuorten-tarkeimmaksi-uutiskanavaksi-kiinnostus-omaan-elamaan-liittyviin-uutisiin-kaavussa/>
- 32 Grönlund, M., Horowitz, M. & Lehtisaari, K. (22.1.2024). Luottamusta rakentamassa, luottamusta kokemassa: Mediapäättäjien ja journalistien yleisöjen näkemyksiä, DECA-hanke, https://cdn.prod.website-files.com/64123799970e0ddca93c317a/65af9e00c156d47a2e021807_Luottamusta%20rakentamassa%2C%20luottamusta%20kokemassa_taitettu_FINAL.pdf
- 33 Meta. (12.3.2021). Learning from videos to understand the world, <https://ai.meta.com/blog/learning-from-videos-to-understand-the-world/>
- 34 Meta. (13.12.2024). Instagram Reels Chaining-AI-system, <https://transparency.meta.com/features/explaining-ranking/ig-reels-chaining/>
- 35 Meta. (2024). Facebook Transparency Center, Widely viewed content report: What People See on Facebook, Q3 2024, hämtad 11.12.2024, <https://transparency.meta.com/fi-fi/data/widely-viewed-content-report/>
- 36 Taylor, S. H. & Choi, M. (2022). An Initial Conceptualization of Algorithm Responsiveness: Comparing Perceptions of Algorithms Across Social Media Platforms, <https://journals.sagepub.com/doi/10.1177/20563051221144322>
- 37 Stone, S. (19.10.2023). Mastering Social Media Algorithms: The Interest Graph vs the Social Graph, <https://www.theshelf.com/the-blog/social-media-algorithms-interest-graph-vs-social-graph/>
- 38 CMS LawTax. (n.d.). GDPR Enforcement Tracker, hämtad 30.12.2024, <https://www.enforcementtracker.com/>
- 39 Europeiska kommissionen. (n.d.). EU:s förordning om digitala tjänster, hämtad 30.12.2024, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_sv
- 40 Sitra. (3.4.2024). Så här stänger du av rekommendationsalgoritmerna på större sociala medier, <https://www.sitra.fi/sv/artiklar/sa-har-stanger-du-av-rekommendationsalgoritmerna-pa-store-sociala-medier/>
- 41 Pönkä, H. & Mikko, S. (28.10.2024). Mitä hyötyä digipalvelusäädöksestä on diginatiiville? Tarkastelussa Tiktok, Faktabaari



16. AI – hot eller möjlighet för demokratin?

MINNA ASLAMA HOROWITZ, HELSINGFORS UNIVERSITET

HOT ELLER MÖJLIGHET? När det gäller generativ AI – och också andra typer och tillämpningar av AI – är svaret: både och. Melvin Kranzberg, professor i teknikhistoria, var av den åsikten om all teknik redan på 1980-talet. Den så kallade Kranzbergs lag har också fått sitt namn efter honom: Ingen teknik är god eller ond, men inte heller neutral. Allt beror på hur, i vilket sammanhang och för vilket ändamål tekniken används.¹

Det här är bra att komma ihåg, när man stöter på antingen skrämmande hotbilder eller lovordande uppfattningar om AI. Detsamma tycker faktiskt också GenAI själv.

När man ställer frågan i rubriken till det generativa AI-verktyget ChatGPT får man följande svar:

ChatGPT påminner en om att alla inte har tillgång till AI och än mindre förmåga att använda AI. AI kan också lätt generera propaganda och desinformation och alltså på många sätt manipulera våra åsikter. Vi vet inte heller vilka principer som ligger bakom olika AI-applikationer.

Å andra sidan kan ChatGPT räkna upp många bra saker om sig själv. AI gör det lättare för medborgare att få tillgång till information och att vara aktiva. AI ger också stöd för beslutsfattande, till exempel genom att skapa olika scenarier eller förenkla komplexa texter. Generativ AI kan utgöra ett av verktygen som används för att upptäcka och bekämpa desinformation på nätet. Generativ AI kan dessutom hålla reda på och erbjuda opartiska och balanserade perspektiv på omdiskuterade frågor och på så sätt främja den offentliga debatten.

Svaren ovan kunde ChatGPT-versionen, som är öppen för alla, generera på nolltid på både engelska och svenska. I ljuset av demokratisk kommunikation är det fantastiskt att det är möjligt att översätta information rekordsnabbt till olika språk.

Å andra sidan säger en sammanfattning av för- och nackdelar ingenting om AI:s vidare konsekvenser, när det gäller den så kallade värdekedjan. För det första kan man vid planering av en applikation fundera över hur den inverkar på demokratin, till exempel medborgarnas grundläggande rättigheter. För det andra kan datainsamlingen, kvaliteten på data och träningen av data präglas av etiska principer och demokrati – eller också inte. Och till sist kan applikationens användningsområden avgöra huruvida den stärker eller urholkar demokratin.

Det finns utmaningar, det finns möjligheter

ChatGPT:s svar, även om de är sanna i sig, illustrerar AI:s största utmaningar och möjligheter för demokratin. I praktiken är alla fördelar ChatGPT själv räknar upp korrekta i sig, men samma fördelar kan också utgöra problem.

Exempel: Om AI är en ypperlig hjälpredda för oss är den det också för dem som sprider lögn, propaganda och hatretorik. Och: Om vi inte får veta hur AI fungerar och vilken typ av data ett AI-verktyg använder sig av, kan vi inte helt lita på de råd verktyget ger oss eller till exempel dem som ska ge medicinska diagnoser för att inte tala om politiska beslutsfattare. Kranzbergs lag kan alltså tillämpas också på AI: det finns utmaningar, det finns möjligheter.

Svaren visar också hur snäv uppfattning ChatGPT har om det stora, komplexa begreppet och de principer vi kallar för demokrati. Den förbiser så där bara till exempel integritetsskyddet och profileringen som görs med hjälp av exempelvis ansiktsgenkänningsteknik. På samma sätt nämns inte heller de ekonomiska faktorerna kring AI (dvs. big business) eller ojämlikheten som förknippas med dem. Vi kan till exempel använda ChatGPT:s gratisapp, men genom att betala kan vi få tillgång till ett bättre verktyg.²

Demokrati innebär att människan genom sitt arbete i den mån det är möjligt är aktiv i samhället, vilket bland annat nämns i FN:s allmänna förklaring om de mänskliga rättigheterna.³ Trots det nämner ChatGPT i sitt svar inte sådana utmaningar för demokratin som arbetsplatser som försvinner eller exempelvis oron för upphovsrätten.⁴ Verktyget nämner inte heller problem relaterade till hållbar utveckling och särskilt energiförbrukning, även om AI utgör en aldrig tidigare skadad energitjuv.⁵ Miljöfrågor är visst ändå i fokus i dagens samhälle – och viktiga för demokratin.

Eftersom AI redan finns överallt och är tillgänglig för allt fler har det skrivits tiotals deklARATIONER och regelverk kring etik och andra principer som respekterar demokratin. I de flesta av dem lyfts fram riskerna – för arbetstagare, integritet, nationell säkerhet, ekonomisk tillväxt – men det är färre som tar upp vanliga människors rättigheter och möjligheter.⁶

Dags att lära sig mer om AI

Men om man ska tro en annan tekniktänkare innebär varje ny teknik en förändring i hur vi förhåller oss till oss själva och varandra.

Redan för mer än ett halvt sekel sedan lade Marshall McLuhan märke till att den gemenskap som präglat historieberättandet kring lägereldar förändrades i och med böcker, eftersom man läste – om man kunde – dem för sig själv. Den digitala kommunikationens tidevarv gav enligt McLuhan upphov till så kallad neotribalism, där människor förenas över geografiska gränser av samma berättelser.⁷ Sociala medier är antagligen ett praktexempel på just det.

Den kunskap och de erfarenheter och kontakter vi får på sociala medier är emellertid inte helt gränslösa, utan algoritmerna gör vår värld snävare på nätet (se Harto Pönkä, kapitel 15 i den här guiden). Ger generativ AI alltså i fortsättningen upphov till en helt ny typ av grupperingar och samhällen? I ljuset av demokratin är den värsta hotbilden att vi lever livet i sällskap av skraddarsydda, virtuella AI-vänner, som alltid bekräftar vår världssyn, och att vi inte litar på någon information vi får i verkliga livet – och än mindre på varandra.

Därför är det dags att kräva att vi alla får lära oss inte bara om AI-kompetens utan också om AI:s inverkan på samhället och att vi alla får påverka hur, var och varför generativ AI används och av vem.⁸

¹ Melvin Kranzberg (1986) Technology and History: "Kranzberg's Laws". Technology and Culture Vol. 27, No. 3 (Jul, 1986). <https://doi.org/10.2307/3105385>

² Se tex. <https://openai.com/chatgpt/pricing/>.

³ Se UN Universal Declaration of Human Rights <https://www.un.org/en/about-us/universal-declaration-of-human-rights>; <https://fn.se/vi-gor/vi-utbildar-och-informerar/fn-info/vad-gor-fn/fns-arbete-med-manskliga-rattigheter/den-allmanna-forklaringen-om-de-manskliga-rattigheterna/>;

⁴ Se tex. Generative AI Has an Intellectual Property Problem, <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>; Tekoäly ja EU – myös tekoälyn on kunnioitettava tekijöiden oikeuksia. <https://www.teosto.fi/teostory/tekoaly-ja-eu-myoys-tekoalyn-on-kunnioitettava-tekijoiden-oikeuksia/> ; e.g.,

⁵ Se tex. AI has an environmental problem, <https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about>; Googlen päästöt karkasivat käsistä tekoälyn vuoksi, <https://yle.fi/a/74-20098239>

⁶ Se tex. EU:s AI-förordning (Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>; World Economic Forum The Presidio Recommendations <https://www.weforum.org/publications/the-presidio-recommendations-on-responsible-generative-ai/> och rapporten Governing AI for Humanity: Final Report skriven av FN:s expertgrupp på hög nivå. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

⁷ Marshall McLuhan (1964/1994 3rd ed.) Understanding Media. The Extensions of Man. <https://mitpress.mit.edu/9780262631594/understanding-media/>

⁸ Kirsi Hantula (publiceras senare under 2025) Kuinka turvata tiedollisia oikeuksia tekoälyn aikakaudella? SITRA; också e.g., Minna Horowitz, Hannu Nieminen, Katja Lehtisaari & Alessandro D'Arma (2024) Introduction. Epistemic Rights in the Era of Digital Disruption. Palgrave MacMillan. <https://link.springer.com/book/10.1007/978-3-031-45976-4>

17. AI och kreativitet – risk eller möjlighet?

ARI HAASIO, SEINÄJOKI YRKESHÖGSKOLA

KREATIVITET ÄR ETT MÅNGFASETTERAT begrepp som det inte finns någon entydig definition av. Generellt sett kan kreativitet anses innebära någonting nytt som har betydelse.¹ Kreativitet har en koppling till både vetenskap och konst, men den kan också vara till exempel en del av affärsverksamhet bland annat i form av nya, innovativa verksamhetsmodeller eller medarbetarnas förmåga att lösa problem på nya sätt. Därför kan man säga att kreativitet är detsamma som att hitta eller komma på en ny infallsvinkel.

Vad kan AI bidra med, och är den kapabel till att vara kreativ? Eller är det så att AI till och med begränsar vår kreativitet? Dagens AI-verktyg är alla exempel på svag AI. Svag eller snäv AI kan endast utföra uppgifter som människan gett den, medan stark AI klarar av att tänka självständigt, på samma sätt som människan. Hittills har det inte varit möjligt att utveckla stark AI.²

När man diskuterar betydelsen av och potentialen hos AI, när det gäller att främja kreativt arbete, är det viktigt att vara noga med om man med kreativitet avser en instrumentell användning av AI i en kreativ process eller det som AI genererar, såsom en bild som skapats av AI.

När man talar om för- och nackdelarna med AI är det alltid störst fokus på generativ AI, som kan användas för att producera innehåll som text och bilder. Det är bra att komma ihåg att generativ AI bara är en form av AI. Olika tillämpningar av AI är en del av vår vardag: avåkningsvarnare i bilar eller robotdammsugare är exempel på det. Även läkarvetenskapen använder sig av AI till exempel vid kirurgiska ingrepp.³

AI:s egna tankar

Låt oss först fråga ChatGPT hur den tycker att den kan främja en kreativ process.

AH: Hur kan AI främja kreativitet?

ChatGPT: AI som inspirationskälla: AI kan analysera stora mängder data, såsom konstverk, musik eller böcker, och föreslå nya idéer eller teman, som kan ge proffskonstnärer ny inspiration.

AI som verktyg: AI kan hjälpa till i en kreativ process, till exempel genom att skapa grafik, musik eller text. AI kan automatiskt skapa skisser eller melodier, som artister kan utveckla vidare.

AI som samarbetspartner: AI kan vara en samarbetspartner, som kreativa personer kan interagera med. Samarbetet kan i sin tur resultera i nya perspektiv och innovationer.

I sitt svar betonar AI sitt eget instrumentella värde: AI är ingen konstnär eller forskare utan ett verktyg som hjälper kreativa personer att få nya idéer, hantera stora datamängder och skissa på olika saker.

ChatGPT träffar här faktiskt mitt i prick. Om vi använder AI för att ta fram nya idéer och för att till exempel skissa på och strukturera upp ett projekt kan den vara till stor hjälp. Drängen ska dock aldrig få bli husbonde. Här lurar en stor fara, även om frestelsen är stor.

En lovsång till lättja?

Gör AI oss allt latare och dummare? Vill vi inte längre anstränga våra grå celler för att lösa problem, söka information eller skapa nytt innehåll?

Till exempel vid lärande kan AI utgöra ett problem: om en studerande låter AI göra hens uppgifter försämrar inlärningsresultaten. Problembaserat lärande är ett exempel på hur den egna kreativiteten kan användas i lärandet. Redan nu har ChatGPT:s ökade popularitet fått studerande att ty sig till AI. AI används bland annat för att skissa på lärdomsprov vid yrkeshögskolor, och en allt större del av gymnasiestuderande använder sig av AI-verktyg för att göra översättningsuppgifter. I de flesta fall är detta en björntjänst för ens eget lärande och kreativa tänkande.

Frestelsen med AI kan förklaras med klassiska Zipfs lag, dvs. principen om minsta ansträngning: man söker information där den är lättast att få tag på.⁴

Idéer och insikter?

Trots att den lata människan kan ty sig till AI för att klara av uppgifter som förutsätter kreativt tänkande och problemlösningsfärdigheter, är den inte nödvändigtvis alltid en dålig assistent. AI har många bra egenskaper, såsom förmågan att analysera stora datamängder, vilket gör att människan istället för rutinmässigt arbete kan fokusera på kognitiva processer.

Informationssökning är en viktig del av en kreativ process, när man skriver, forskar eller studerar. Även på den fronten kan AI vara till stor nytta, ifall man i sökandet efter lämpliga källor inte nöjer sig med verktyg som ChatGPT, som lätt kan hallucinera. Det har utvecklats AI-verktyg med syfte att underlätta informationssökningen, såsom IrisAI, som testats vid Östra Finlands universitet och Yrkeshögskolan Haaga-Helia.⁵

Testgruppen förhöll sig något skeptisk till IrisAI. Verkttyget hittade användbara källor, men testgruppen undrade hur heltäckande sökresultaten faktiskt var. Var det något väsentligt som utelämnades, något som en informatiker skulle ha hittat?

AI och litteratur

AI har gett konstnärer verktyg och möjligheter till ett nytt sätt att skapa konst. Ett gott exempel på det är olika slags stora videoskopulturner. Generativ AI har också fört med sig förändringar i processen för textproduktion. Främst har AI haft en inverkan på hur vi skriver essäer och uppsatser av olika slag, men det finns redan också fiktivt material och faktaböcker, som producerats med hjälp av AI. I framtiden kommer deras antal sannolikt att öka avsevärt.

Med hjälp av AI kan hela romaner skrivas. Ett av de mest kända exemplen är romanen *Tokyo-to Dojo-to* (engl. Tokyo Tower of Sympathy)⁶, som japanska Rie Kudan skrivit med hjälp av AI och som 2023 tilldelades litteraturpriset Akutagawa. Det finns också engelskspråkiga AI-verktyg ute

Det är författarnas jobb att skriva, och det bör fortsätta att vara så.

på marknaden som utvecklats enkom för att producera skönlitterär text. Människan slår fast genren och ett antal andra bakgrundsvariabler, och AI skapar manuset. Så som jag ser det inverkar AI, när den används så här, skadligt på litteraturen. AI producerar enkelt lätt, schablonmässig underhållning. Sådana verk är inte särskilt bra på att beskriva känslor eller visa prov på analytiskt tänkande.

Sudowrite⁷ gör det till exempel enkelt att skriva en roman: man ska välja en genre och ange ett ämne, och Sudowrite skriver 300 ord åt en. Verket ger förslag på fortsättningen, och man kan välja bland dem eller ge egna alternativ. Dessutom är det i Sudowrite möjligt att välja en stil som baserar sig till exempel på en viss författares berättarstil. Man kan alltså be AI skriva i samma stil som till exempel Agatha Christie. Och då gör den meningarna i texten «christieaktiga».

Handlar det i själva verket om kreativitet om en bok skrivs av Sudowrite eller ett motsvarande verktyg? Producerar sådana verktyg mångfasetterad, personlig och reflekterande text? Även om det är fråga om underhållning tror jag att de nuvarande AI-verktygen borde på sin höjd användas till att ta fram skisser. Det är författarnas jobb att skriva, och det bör fortsätta att vara så.

På svenska har det redan getts ut fiktion, som skrivits av AI. AI har dessutom använts i vissa faktaböcker främst för att strukturera innehållet och för att ta fram idéer för böckerna. På engelska har det också utkommit en hel del AI-fiktion, medan det på finska hittills endast har getts ut några AI-skrivna diktsamlingar.⁸

Ett verktyg för konstnärer

I konstkreter har AI väckt hetsig debatt. Debatten har till och med handlat om huruvida

AI-producerad konst är riktig konst.⁹ Precis som Sudowrite gör författare av amatörskribenter, gör AI det möjligt för vem som helst att skapa bilder. Upphovsrätten blir då i många fall ett problem, eftersom professionella konstnärers verk utnyttjas ofta utan tillstånd.

Återigen är det fråga om hur AI faktiskt används. Konstnären Pasi Rauhala har till exempel använt sig av AI i sina verk. Det har möjliggjort många sådana lösningar som annars skulle ha varit svåra att få till. Ett exempel på det är Rauhalas videoskulptur *Mycelium of the Future* i Aalto-universitetets lokaliteter.¹⁰

Givetvis gör AI det också möjligt för alla amatörkonstnärer att skapa bilder av högre kvalitet än clipart-bilder, men i slutändan är det människan som skapar konst, inte AI. Vilken kreativ process det än handlar om är AI ett verktyg som människan kan använda sig av på ett ansvarsfullt sätt. Kreativiteten börjar med människan, medan datatekniken kan bidra med nya dimensioner och öppna nya möjligheter. Någon självständig konstnär eller forskare kan tekniken tills vidare dock inte vara.

¹ Sternberg R. J. (2001). What is the common thread of creativity? Its dialectical relation to intelligence and wisdom. *American Psychologist* 2001; 56. s. 360-362.

² Haasio, A. (2025). Tekoöly ja kirjastot: Avain. (utg. 1/2025).

³ Chatterjee, S., Das, S., Ganguly, K., & Mandal, D. (2024). Advancements in robotic surgery: innovations, challenges and prospects. *Journal of Robotic Surgery*, 18 (1), 28.

⁴ Haasio, A., Harviainen, J. T. & Savolainen, R. (2019). Johdatus tiedonhankintatutkimukseen. Avain.

⁵ Haasio, 2025.

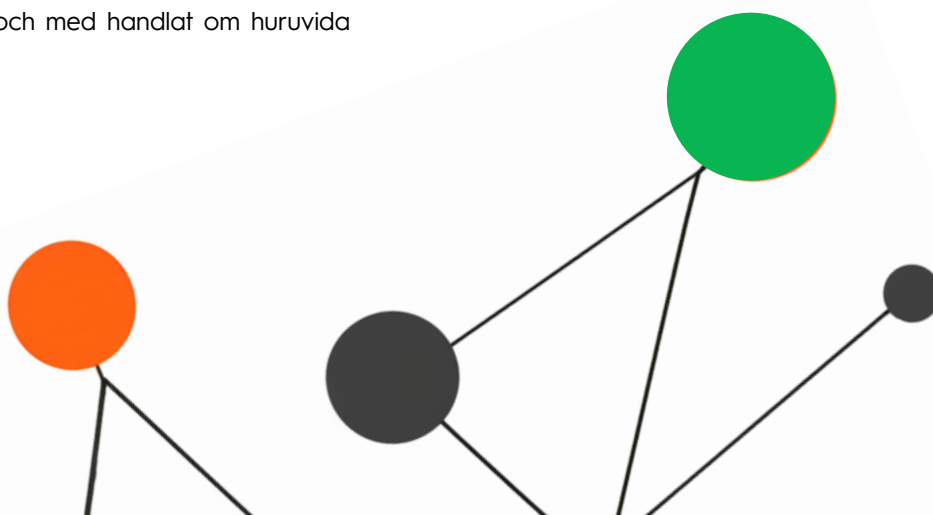
⁶ Yao, D. (5.2.2024). Author Admits tl. o Using AI to Write Award-winning Novel. *AI Business* 5.2024. <https://aibusiness.com/nlp/author-reveals-using-chatgpt-for-award-winning-novel>

⁷ Sudowrite. <https://www.sudowrite.com>.

⁸ Haasio, 2025.

⁹ Mineo, L. (2023). Writer, animator, architect, musician, and mixed-media artist detail potential value, limit of works produced by AI. *The Harvard Gazette* 15.8.2023. <https://news.harvard.edu/gazette/story/2023/08/is-art-generated-by-artificial-intelligence-real-art/>

¹⁰ Salonen, P. (2023). Tekoöly rynnii taiteeseen. *Taide* 3/2023.



18. Amerikanska AI-rekommendationer för lärare?

PATI RUIZ, SENIOR DIRECTOR OF EDTECH AND
EMERGING TECHNOLOGIES, DIGITAL PROMISE

TEMAT FÖR DEN HÄR ARTIKELN är de utmaningar som aktualiserats när man integrerat och tagit i bruk ny teknik som fortfarande utvecklas, såsom GenAI, i skoldistrikt i USA. Målet har varit att dela erfarenheter av samarbetet mellan skoldistriktet samt kunskap och resurser inom utbildningssektorn. På så sätt vill vi hjälpa skolledare och lärare att fatta faktabaserade beslut, som gagnar både deras elever och skolgemenskaper.

Bakgrund

Varje dag kommer det ut hundratals nya GenAI-applikationer på marknaden. Det finns dock inte ännu några belegg för vilka effekter de har eller till hur stor nytta de är i lärandet, vilket kan påverka hur olika riktlinjer anpassas på ett ändamålsenligt sätt.¹ Användarna har snabbt tagit till sig olika GenAI-verktyg: Enligt en relativt ny enkätundersökning bland 1002 K-12-lärare och 1000 elever i USA, som genomförts av Walto Family Foundation², använder 51 procent av lärarna ChatGPT. Många av de svarande (40 %) uppgav att de använde verktyget minst en gång i veckan. Trots att tekniken snabbt har börjat användas är många osäkra på vad GenAI är, hur den fungerar och hur den påverkar elever, familjer och lärare.

Införandet av GenAI-verktygen i skolvärlden aktualiserar flera etiska³ frågor⁴ samt frågor om teknikens miljöpåverkan. Därutöver måste man se till att AI-verktygen och -systemen uppfyller tillgänglighetskraven för alla elever och hela personalen.

För att kunna skapa sig en uppfattning om hur införandet av GenAI påverkar elever, undervisande personal och familjer runt om i USA har GenAI-arbetsgruppen sammankallat en stor grupp ledande personer inom utbildningssektorn. Målet är att tillsammans utreda GenAI:s roll i utbildningen och stärka utvecklingen av policyer som främjar ansvarsfull, etisk och effektiv användning av AI (Responsible, Ethical and Effective Responsible/Acceptable Use Policies, R/AUP).

Forskningsaktivitet

Nätverket Digital Promise's League of Innovative Schools är ett forum för innovativa utbildningsledare från 150 skoldistrikt och 38 delstater. Nätverket planerar, testanvänder och skalar lösningar, som

kan användas för att påverka inlärningsresultaten med särskilt fokus på marginaliserade elever. Diskussionerna den senaste tiden har lyft fram utmaningar kring AI och ny teknik, när det gäller undervisning och lärande, som utbildningsledare möter i sitt arbete. Projektgruppen bjöd in distrikten inom nätverket till en GenAI-arbetsgrupp.

Forskargruppen och medlemmarna i arbetsgruppen träffades en gång i månaden under läsåret 2023-2024 för att diskutera framstegen i de olika skoldistrikten. GenAI-arbetsgruppen skapade också två fokusgrupper under projektperioden.

Resultat

Målgrupperna, intervjuerna och månadsmötena lyfte fram förändringen i skoldistriktens fokusområden efter införandet av AI och annan ny teknik. Även om medlemmarna i GenAI-arbetsgruppen hade utvecklat strategier för att kartlägga användningen av GenAI på var sitt håll, förstod de att det fanns olika sätt i olika skoldistrikt att ta de första stegen och att det var viktigt att inte skynda på implementeringen.

Medlemmarna i GenAI-arbetsgruppen konstaterade att undervisande personal av ett flertal olika orsaker tvekade att använda sig av AI, såsom AI:s okända inverkan på inlärningsprocessen, de totala kostnaderna och i vissa fall också bristerna i AI-kompetens och tekniskt kunnande.

I diskussioner med ledarna för skoldistrikten visade eleverna att det fanns variation i hur väl de förstod hur generativ AI fungerade och att de var osäkra på hur AI kan användas så att den är till nytta. Personalrepresentanter berättade hur generativ AI används på deras arbetsplatser, men de var samtidigt oroad över hur snabbt tekniken som är i ständig förändring inverkar både på deras eget dagliga liv och på elevernas liv.

Fokusområden för GenAI-vägledning

I diskussionerna med distriktsledarna och skolmedarbetarna lyftes det fram att det behövs stödresurser, som gör det lättare att integrera AI och övrig ny teknik i olika utbildningsmiljöer. Utifrån dessa diskussioner samt Digital Promises AI Literacy Framework-referensramen⁵ och EdSafe AI Alliances

SAFE Benchmarks-jämförelsekriterierna⁷ har vi identifierat sex fokusområden för GenAI-vägledning:

1. AI-kompetens
2. säkerhet
3. etik
4. transparens
5. anvisningar för implementering
6. utvärdering och mätning av effekter.

Dessa fokusområden beskrivs närmare nedan.

AI-kompetens

I AI-kompetensen ingår kunskap om hur mediekompetens, digital litteracitet och kritiskt tänkande kan utvecklas, så att eleverna, lärarna och skoldistriktet förstår hur AI kan användas på ett tryggt och etiskt sätt. Behoven inom det här fokusområdet är gemensamma. AI-kompetensen definierades som de kunskaper och färdigheter som hjälper människan att kritiskt förstå, utvärdera och använda sig av AI-system och -verktyg för att på ett tryggt och etiskt sätt vara en del av en värld som blir alltmer digital⁸.

Även om AI-kompetensen inte behöver nämnas i rekommendationerna kan distriktsledarna föreslå riktlinjer för en sådan användning av GenAI-verktyg som skapar mervärde. Det rekommenderas att elevernas AI-kompetens inte utvecklas i form av separata kurser, utan de här färdigheterna och handlingsmodellerna bör integreras i läroplanen i olika årskurser i anslutning till den nuvarande undervisningen om mediekompetens, digital kompetens och informationshantering.

Säkerhet

Det som ska komma i första rummet är att skydda elevernas, lärarnas och hela skolans data och integritet samtidigt som eventuella cybersäkerhetsrisker tacklas och federala riktlinjer, såsom Family Educational Rights and Privacy Act (FERPA)¹⁰ och Children's Online Privacy Protection Rule (COPPA), tillämpas¹¹. Dessa riktlinjer förutsätter att vårdnadshavarna ger sitt samtycke, innan elevernas uppgifter delas, och riktlinjerna ska också fastställa en åldersgräns för användningen av AI-verktyg, till exempel 13 år för chattbottar.

Etik

Användarna av GenAI ska vara ansvarsfulla och agera rättvist och opartiskt. De ska också vara medvetna om att både det som människan själv skapat och det som GenAI:s genererat kan präglas av snedvridningar. Bristerna i rättvisa och etik samt snedvridningarna har blivit till några av de största orosmomenten vid användning av AI. Enligt arbetsgruppen är den undervisande personalen främst orolig över fusk, men personalen oroade sig också för hur de kan stötta de elever som riskerar att bli offer för snedvridningarna i AI-system och -verktyg.

Transparens

De som ansvarar för utbildningen bör öppet redogöra för vilka processer de använder sig av för att välja ut GenAI-verktyg för skolor och hur de utvecklar och anpassar sina riktlinjer. Som det framgår av Office of Educational Technology¹²: "Beslutsfattarna inom utbildningssektorn bör förstå hur olika AI-modeller fungerar i olika undervisningssituationer, så att de bättre kan förutse de begränsningar, problem och risker modellerna har."

Därtill bör det rekommenderas att både eleverna och den undervisande personalen är öppna med hur de eventuellt använt GenAI-verktyg.

Riktlinjer för implementering

Rekommendationerna för implementering av GenAI-verktyg bör också innehålla rekommendationer som hjälper elever och undervisande personal att använda AI-verktyg efter bästa förmåga och att skapa material som de själva och skolkollegorna gagnas av. Både elever och undervisande personal önskade sig klarhet i när det är tillåtet att använda AI och när det inte är det. Även om användarna själva i slutändan fattar beslut om huruvida de använder AI-verktyg behöver de trots det vägledning, särskilt då det är fråga om ny teknik.

Utvärdering och effekter

Det är viktigt att ge anvisningar om hur användarna bör utvärdera AI-system och -verktyg och effekterna av att de används. Skoldistriktet, den undervisande personalen, eleverna och hela skolmiljön bör även framöver fundera över huruvida enskilda verktyg faktiskt hjälper eller snarare stjälper.

I anvisningarna om utvärdering bör det också ingå instruktioner för vilka verktyg som utifrån ett visst användningsbehov kan användas och reflektioner över konsekvenserna av användningen av AI-verktyg. Användbara instruktioner för utvärdering av AI finns bland annat i artikeln AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology¹³.

Huvudresultatet av GenAI-arbetsgruppens projekt är dokumentet "Emerging Technology Guidance Language Deck"¹⁴. Det är ett dokument som uppdateras kontinuerligt och som innehåller exempel på möjlig användning av GenAI och övrig ny teknik i undervisningen.

Artikeln baserar sig på publikationen "A Position Paper for NSF Grant 2334525: Responsible, Ethical, and Effective Acceptable Use Policies for the Integration of Generative AI in US School Districts and Beyond"¹⁵.

¹ Lodge, J., Thompson, K. & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*. 39. 1-8. <https://doi.org/10.14742/ajet.8695>.

² Walton Family Foundation. (2023, March 1). ChatGPT Used by Teachers More Than Students, New Survey from Walton Family Foundation Finds <https://www.waltonfamilyfoundation.org/chatgpt-used-by-teachers-more-than-students-new-survey-from-walton-family-foundation-finds>

³ Nguyen, A., Ngo, H.N., Hong, Y., Dang, B., Nguyen, B. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technology* 28, 4221–4241. <https://doi.org/10.1007/s10639-022-11316-w>

⁴ Unesco. (2023). ChatGPT and Artificial Intelligence in higher education: Quick start guide. United Nations. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

⁵ Ruiz, P., Armstrong, A. L., Karim, S., Shell, A., Singmaster, H., Giang, M. (2024) Emerging Technology Acceptable Use Policy [Google slides]. Digital Promise. <https://bit.ly/DPAUPLanguage>

⁶ Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

⁷ EDSAFE AI Alliance. (2024). Safe benchmarks framework. EDSAFE AI. <https://www.edsafeai.org/safe>

⁸ Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

⁹ Digital Promise <https://digitalpromise.org/initiative/artificial-intelligence-in-education/ai-literacy/>

¹⁰ FERPA <https://studentprivacy.ed.gov/faq/what-ferpa>

¹¹ COPPA <https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa>

¹² Office of Educational Technology. (2023). Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations, U.S. Department of Education.

¹³ Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

¹⁴ Digital Promise Emerging Technology Guidance Language Deck <https://digitalpromise.org/items/f84dedc7-762c-460e-b9ba-eb4ba28a9b51>

¹⁵ Pati Ruiz, Ed.D., Sana Karim, Amanda LaTasha Armstrong, PhD, Alison Shell, PhD, Babe A Position Paper for NSF Grant 2334525: Responsible, Ethical, and Effective Acceptable Use Policies for the Integration of Generative AI in US School Districts and Beyond" Liberman. https://stelar.edc.org/sites/default/files/2024-09/NSF2334525_20Ruiz%20RAPID_%20Responsible%2C%20Ethical%2C%20and%20Effective%20Acceptable%20Use%20Policies%20for%20the%20Integration%20of%20Generative%20AI%20in%20US%20School%20Districts%20and%20Beyond.pdf

19. GenAI, upphovsrätten och lärare

KARI KIVINEN, FAKTABAAARI

Innehållet i materialet som använts för att träna populära språkmodeller är åtminstone för närvarande höljt i dunkel.

ANVÄNDNINGEN AV AI-VERKTYG ger upphov till både etiska och upphovsrättsliga bekymmer. För att lärare samt elever och studerande med tillförsikt samt kritiskt och tryggt ska kunna använda sig av ny och framväxande teknik bör de ha grundläggande kunskaper om och en samsyn på hur GenAI kan användas i undervisningen med hänsyn till upphovsrättsliga frågor.

Med hjälp av generativa AI-verktyg är det möjligt att skapa text, bilder, videor och musik. Bakom det kreativa material dessa verktyg genererar finns emellertid en hel del upphovsrättsliga osäkerhetsfaktorer och en lagstiftning som har svårt att hålla jämna steg med den snabba utvecklingen.

Träningen av stora språkmodeller förutsätter enorma mängder text, kod, bilder och data, varav en stor del kan vara upphovsrättsskyddat material. Om den här typen av data samlas in utan tillstånd finns det en risk för att det sker intrång på någons upphovsrätt.

Innehållet i materialet som använts för att träna populära språkmodeller är åtminstone för närvarande höljt i dunkel. EU:s AI-förordning, som träder i kraft 2026, förutsätter att de allmänt tillgängliga AI-systemen i framtiden ska följa EU:s upphovsrättslagstiftning och ge sammandrag av materialet som använts vid träning av systemens data. Den närmaste framtiden får utvisa hur detta kommer att genomföras. Eftersom vi inte ännu vet vilken information som använts för att träna AI-modeller är det svårt att bedöma eventuella intrång på upphovsrätten.

Många rättsfall

AI-verktyg kan producera material som imiterar eller i hög grad liknar ett upphovsrättsskyddat verk. Flera konstnärer och förläggare har lämnat in stämningsansökningar, där de anklagar AI-plattformar för olaglig kopiering och förbjuder användningen av sina verk vid insamling av data och vid träning av AI-verktyg. Fallet som fått mest medieuppmärksamhet är New York Times stämning mot OpenAI. OpenAI försvarar sig med ett intressant svar:

”Eftersom upphovsrätten numera i praktiken gäller alla mänskliga uttrycksformer – inklusive blogginlägg, foton, foruminlägg, delar av programkod och regeringens dokument – är det omöjligt att

träna dagens ledande AI-modeller utan att använda upphovsrättsskyddat material.”¹

Besluten i dessa rättsfall kommer att innebära ändringar i och uppdateringar av upphovsrättslagstiftning och -praxis under de närmaste åren. Det är tills vidare relativt oklart hur vi kan förhindra att AI-plattformar gör intrång på upphovsrätten och vilka påföljder som borde tillämpas och av vem.

Den viktigaste frågan för användarna av AI-verktyg är vem som äger det som AI producerat. Kan traditionell upphovsrätt tillämpas på sådant material, eller behövs det nya lösningar?

För närvarande skyddar upphovsrätten inte ett verk, om det har producerats av AI och om en människa inte till en väsentlig del eller inte alls har varit involverad i skapandet av verket.

Försök att förhindra datainsamling

I synnerhet stora förlag och mediehus har börjat förhindra insamlingen av upphovsrättsskyddat innehåll på sina webbplatser. Det att insamlingen av träningsdata förhindras kan innebära att dessa webbplatser synlighet i sökmotorer begränsas och att lagliga sökmotorer hindras från att indexera innehållet på webbplatserna.

Enligt Wired² har 88 procent av de ledande nyhetsbyråerna i USA förhindrat datainsamlingen. I februari 2024³ rapporterade Reuters Institute att nästan hälften (48 %) av de ledande nyhetssajterna i tio länder blockerat OpenAI:s datainsamlare. För användarna av GenAI är detta i viss mån problematiskt, eftersom aktuellt innehåll av hög kvalitet då ofta inte tas med i träningsmaterialet. För att försäkra sig om tillgången till träningsmaterial av hög kvalitet har AI-plattformarna börjat ingå licensavtal med olika innehållsproducenter. Detta kommer förmodligen att bli vanligare inom den närmaste framtiden.

Upphovsrätt och utbildning

Generativa AI-plattformar kan samla in data för sina databaser från alla promptar och instruktioner för att ytterligare träna sina språkmodeller. Detsamma kan inträffa till exempel när

studenters arbeten skannas in med program för plagiatkontroll.

Speciellt med känsliga uppgifter bör man vara försiktig, och de ska under inga omständigheter matas in i ett AI-verktyg. Därför är det av största vikt att välja att använda AI-verktyg som tar hänsyn till dataskyddet. Verktygen bör exempelvis kunna garantera att data från promptar inte lagras eller återanvänds för ytterligare träning av verktygen. Det är också viktigt att komma ihåg att största delen av studenternas arbeten består av upphovsrättsskyddat material. Utbildningsanordnare gör därför klokt i att försäkra sig om att AI-verktygen som används också följer EU:s upphovsrättslagar.

EUIPO har gett ut en interaktiv infografik om temat "Hur kan lärare hjälpa elever att undvika intrång på upphovsrätten när de använder GenAI?"

1. Berätta för dina elever om upphovsrätten.
2. Förklara att AI-genererat material kan göra intrång på existerande upphovsrätt, om det i materialet används upphovsrättsskyddat material.
3. Undervisa om hur man hänvisar till källmaterial på ett ändamålsenligt sätt och förklara på vilket sätt man kan referera eller citera innehåll som producerats av GenAI.
4. Stärk elevernas kritiska tänkande, så att de lär sig verifiera sanningshalten i det som GenAI producerat.
5. Uppmuntra eleverna att skapa unika verk och att använda AI som ett verktyg för inspiration, inte som en ersättning för egen kreativitet.
6. Skapa interaktiva lektioner, där eleverna kan öva på att identifiera, använda och referera AI-producerat innehåll på korrekt sätt.

¹ OpenAI (hämtad 12.12.2024) OpenAI faces multiple lawsuits over its use of copyrighted articles, books, and art to train its generative AI tools. <https://www.euronews.com/next/2024/01/09/openai-says-its-impossible-to-train-ai-without-copyrighted-materials>

² Wired (2024) Most top news sites block AI bots. Right-wing media welcomes them. <https://www.wired.com/story/most-news-sites-block-ai-bots-right-wing-media-welcomes-them/>

³ Reuters Institute (2024) How many news websites block AI crawlers. <https://reutersinstitute.politics.ox.ac.uk/how-many-news-websites-block-ai-crawlers>

20. AI-verktyg och dataskydd

HARTO PÖNKÄ, INNOWISE

I EUROPEISKA UNIONEN REGLERAS behandlingen av personuppgifter och dataskyddet för register huvudsakligen av EU:s allmänna dataskyddsförordning (GDPR).¹ Många andra förordningar, såsom EU-förordningar och nationella lagar, innehåller också bestämmelser om behandling av personuppgifter och integritetsskydd.

Till exempel reglerar lagen om grundläggande utbildning i Finland inte direkt vilka personuppgifter som specifikt ska behandlas i utbildningskontext, men skyldigheterna som nämns i lagen utgör

grunden för kommuners elevregister och mer allmänt för vilka personuppgifter som överlag kan behandlas i den grundläggande utbildningen.² Situationen är likadan på övriga utbildningsstadier.

Utbildningsanordnare ska inom ramen för de lagar som styr verksamheten besluta om och planera vilka personuppgifter som behandlas och hur.

Utbildningsanordnare ska inom ramen för de lagar som styr verksamheten besluta om och planera vilka personuppgifter som behandlas och hur. En enskild lärare måste följa utbildningsanordnarens riktlinjer och anvisningar. När syftet med behandlingen av personuppgifter är att anordna utbildning får

uppgifterna inte användas för andra ändamål såsom marknadsföring, programvaruleverantörers produktutveckling eller träning av fritt tillgängliga AI-modeller. I digitala lärmiljöer, AI-applikationer och övriga digitala plattformar ska dataskyddsbestämmelserna beaktas speciellt noggrant, så att behandlingen av personuppgifter följer lagen.

Dataskyddet beaktas redan vid planering av verksamhet

I utbildningen är det utbildningsanordnaren som är den personuppgiftsansvariga – alltså till exempel en kommun eller den instans som ansvarar för verksamheten vid en privat utbildningsinstitution. Dataskyddet bör beaktas redan i planeringen av verksamheten.

En av de första frågorna som ska avgöras är den rättsliga grunden för behandlingen av personuppgifterna. Den mest naturliga och också rekommenderade rättsliga grunden i utbildningen, som tas upp i dataskyddsförordningen, är rättslig förpliktelse, även om den inte heller är helt problemfri.³ I kommersiell eller på annat sätt

avtalsbaserad utbildning baseras behandlingen av personuppgifter i sin tur på avtalen i fråga.

Något som också är väsentligt när behandlingen av personuppgifter planeras är att göra en riskbedömning på förhand och att besluta om skyddsåtgärder utifrån den. Skyddsåtgärder kan vara antingen tekniska eller organisatoriska. Syftet med skyddsåtgärder är fram för allt att förhindra att uppgifterna används för olagliga ändamål och att säkerställa att de registrerades rättigheter respekteras i enlighet med GDPR. Till exempel handlar det å ena sidan om att personuppgifterna inte läcks ut till utomstående och att de inte används för andra ändamål än för att anordna utbildning och å andra sidan om att den registrerade kan utöva sina rättigheter till exempel för att få transparent information om hur personuppgifterna behandlas och för att granska och rätta sina uppgifter.

På digitala plattformar handlar skyddsåtgärder till exempel om användarnamn, inloggning, kryptering av nätverksanslutning, säker lagring och säkerhetskopiering av uppgifter, åtkomstnivåer, raderande av uppgifter inom utsatt tid, bekämpning av sabotageprogram, säkerhetsuppdateringar och så vidare. Eftersom

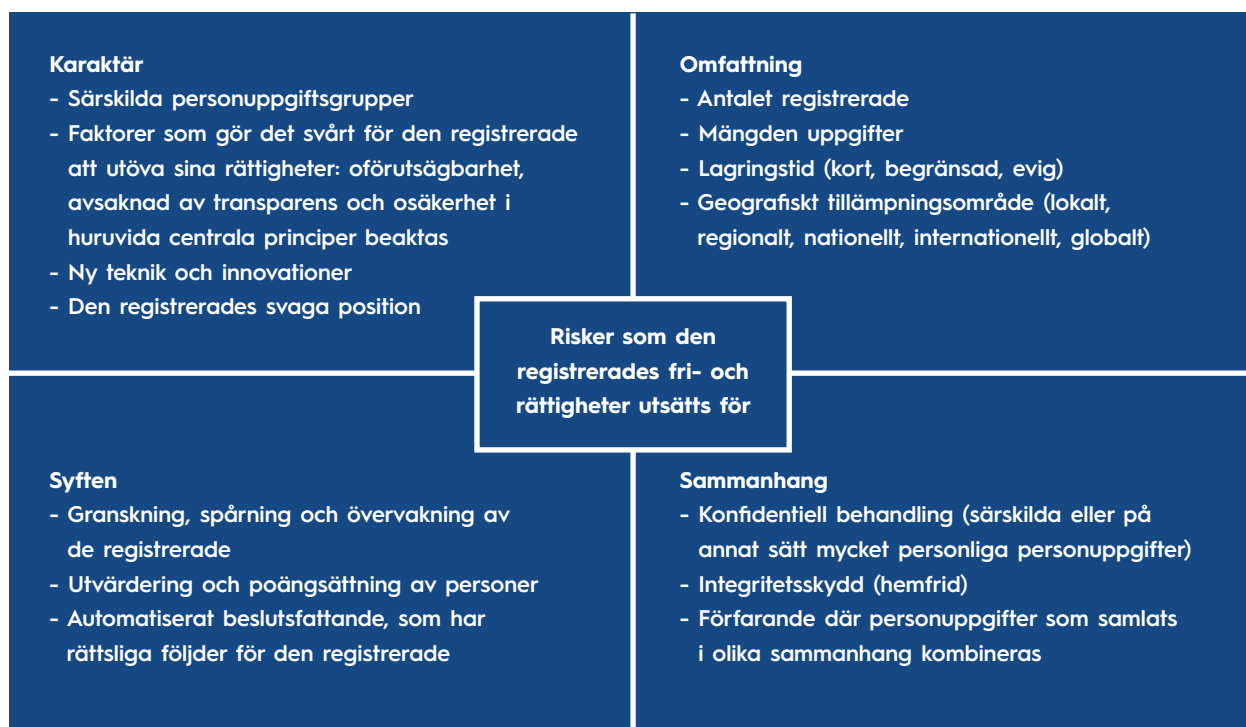
applikationerna används av människor räcker det inte enbart med tekniska skyddsåtgärder, utan därtill behövs det organisatoriska åtgärder såsom anvisningar, personalutbildning och manuell övervakning.

Webbtjänster och applikationer är ofta komplexa vad gäller både de faktiska funktionerna och dataskyddet. I AI-applikationer av olika slag ökar den tekniska komplexiteten ytterligare. AI-baserade funktioner läggs också till i applikationer där de tidigare inte har funnits.

Till exempel har generativa AI-funktioner lagts till i vanliga program för bildredigering. I ljuset av bedömningen av riskerna för dataskyddet innebär detta vissa utmaningar, eftersom det kan visa sig vara arbetskrävande också i bekant programvara att gå igenom alla funktioner och egenskaper som behöver utvärderas.⁴

Bedömning av dataskydd i applikationer

GDPR förutsätter att den personaluppgiftsansvariga gör en bedömning av riskerna vid behandling av personuppgifter och vidtar skyddsåtgärder utifrån de identifierade riskerna. Riskbedömningen ska beakta både hur allvarliga de potentiella



konsekvenserna för den registrerade är och hur sannolikt de inträffar. Om risken bedöms vara hög ska den personuppgiftsansvariga göra en konsekvensbedömning kring dataskyddet i enlighet med den allmänna dataskyddsförordningen. Konsekvensbedömningen är ett förfarande som fastställs i GDPR och som går ut på en systematisk genomgång av behandlingen av personuppgifter och en kartläggning av möjliga hot, så att det kan identifieras effektiva skyddsåtgärder.^{5 6}

En oförsiktig implementering av applikationer i undervisningen kan innebära att dataskyddet försummas och leda till att personuppgifter utsätts för risker. Därför måste AI-applikationer utvärderas i ljuset av dataskyddet på samma sätt som andra digitala plattformar, innan de tas i bruk i undervisningen.

Både utbildningsanordnaren i egenskap av den personuppgiftsansvariga och enskilda lärare kan hållas ansvariga för eventuella förseelser, ifall de inte har följt arbetsgivarens anvisningar om applikationer som ska användas och dataskyddsprinciper som ska följas. Innan en lärare börjar använda ett AI-verktyg eller en annan applikation är det viktigt att utreda vilka anvisningar och riktlinjer kring dataskyddet det finns inom organisationen.

Vanligtvis är det leverantörerna för inlärningsplattformar eller applikationer som är personuppgiftsbiträden, och då begränsas deras ansvar av gällande avtal. I den allmänna dataskyddsförordningen avses med personuppgiftsbiträdet den instans som i den personuppgiftsansvarigas räkning behandlar personuppgifter – till exempel upprätthåller en applikation i enlighet med ett serviceavtal.⁷

Vid riskbedömning kan man ta hjälp av följande schema, som innehåller faktorer som ökar riskerna och som har att göra med vad behandlingen av uppgifterna har för karaktär, omfattning, sammanhang och syften.⁸

När det är fråga om digitala plattformar och AI-verktyg som används i undervisningen ökar riskerna ofta till exempel i och med faktorer som:

- **Särskilda personuppgifter (t.ex. hälsouppgifter)**
- **Sekretessbelagda uppgifter (t.ex. pedagogiska dokument och verbala omdömen om en persons egenskaper)**
- **Den registrerades svaga position (minderåriga, personer i behov av särskilt stöd, eleven i relation till utbildningsanordnaren)**
- **Positionsdata (möjliggör systematisk spårning)**
- **Stort antal (t.ex. barn i läropliktig ålder i en viss kommun)**
- **Lång lagringstid (t.ex. läropliktstiden på nio år)**
- **Konfidentiella uppgifter (t.ex. uppgifter om familjemedlemmar och övrigt privatliv)**

När två av kriterierna ovan uppfylls kan det vara fråga om en situation där personuppgifterna utsätts för hög risk. Därtill ska det göras en konsekvensbedömning i flera särskilt nämnda fall. Sådana kan vara till exempel användning av ny teknik, användning av positionsdata, behandling av känsliga eller mycket personliga uppgifter samt profilering och automatiserat beslutsfattande, som kan ha betydande konsekvenser för den registrerade.

I utbildningssammanhang är det ofta så att kriterierna för hög risk kan uppfyllas i vissa enskilda fall. Till exempel kan elevers särskilda personuppgifter och konfidentiella uppgifter sparas i en applikation, trots att det inte varit utbildningsanordnarens avsikt, om eleverna på eget initiativ inkluderat sådana uppgifter i sina uppgiftssvar. Om applikationen också innehåller AI-baserade självriktade uppgifter kan det vara fråga om behandling av personuppgifter som involverar profilering och automatiserat

Elevers särskilda personuppgifter och konfidentiella uppgifter kan sparas i en applikation, trots att det inte varit utbildningsanordnarens avsikt

beslutsfattande. Utbildningsanordnaren ska därför se över behandlingen av personuppgifter i olika situationer, där AI-funktionerna på en inlärningsplattform kommer åt uppgifter som eleverna och lärarna sparar på plattformen. Å andra sidan måste lärare vara noga med att sådana AI-funktioner inte oavsiktligt är i bruk, om de inte först har utvärderats.

I molntjänster, som de flesta AI-verktyg också är, kan olika typer av personuppgifter sparas på servrar som tjänsten använder sig av. Sådana uppgifter kan vara användarnamn, filer och annat material samt uppgifter i en aktivitetslogg eller en annan typ av logg, som samlas in när molntjänsterna används. Vid riskbedömningar av molntjänster är det viktigt att speciellt ta hänsyn till att personuppgifterna behandlas transparent, att uppgifterna minimeras, att lagringsperioden begränsas och att uppgifterna behandlas konfidentiellt.

En kartläggning av hur data som innehåller personuppgifter behandlas förutsätter en genomgång av tjänsteleverantörens avtal, dataskyddsbeskrivningar och övriga dokument. Många programvaruföretag har utvecklat sina applikationer främst för marknader utanför EU, vilket gör att kraven i GDPR inte nödvändigtvis beaktats i dem. Även om en app påstås följa GDPR är det ännu ingen garanti för att funktionerna i applikationen uppfyller den nivå av dataskydd som gäller för utbildningssammanhang.

Molntjänsterna aktualiserar också frågan om var personuppgifterna lagras geografiskt och huruvida de överförs till länder utanför EU och EES. Till exempel i Microsofts, Googles och Adobes webbapplikationer kan personuppgifter överföras till länder utanför EU, där dataskyddet inte uppfyller de krav som fastställs i GDPR. I serviceavtalet för en enskild applikation kan det nämnas till och med flera tiotals underbiträden i tiotals olika länder. I en del applikationer kan lagringen av data begränsas till EU-området till exempel med en tilläggstjänst eller genom att kunden köper en dyrare licens. Om detta inte är möjligt ska andra skyddsåtgärder utredas.

Observationer av dataskydd i AI-applikationer

Under det senaste året har jag varit involverad i flera konsekvensbedömningar av AI-applikationer och andra digitala plattformar som används i undervisningen. Nedan går jag igenom några av mina observationer som jag gjort i det arbetet. Det som arbetet gett vid handen är att jag alltid rekommenderar att man gör upp en noggrann plan och gör en riskbedömning innan man tar i bruk AI-applikationer eller -funktioner.

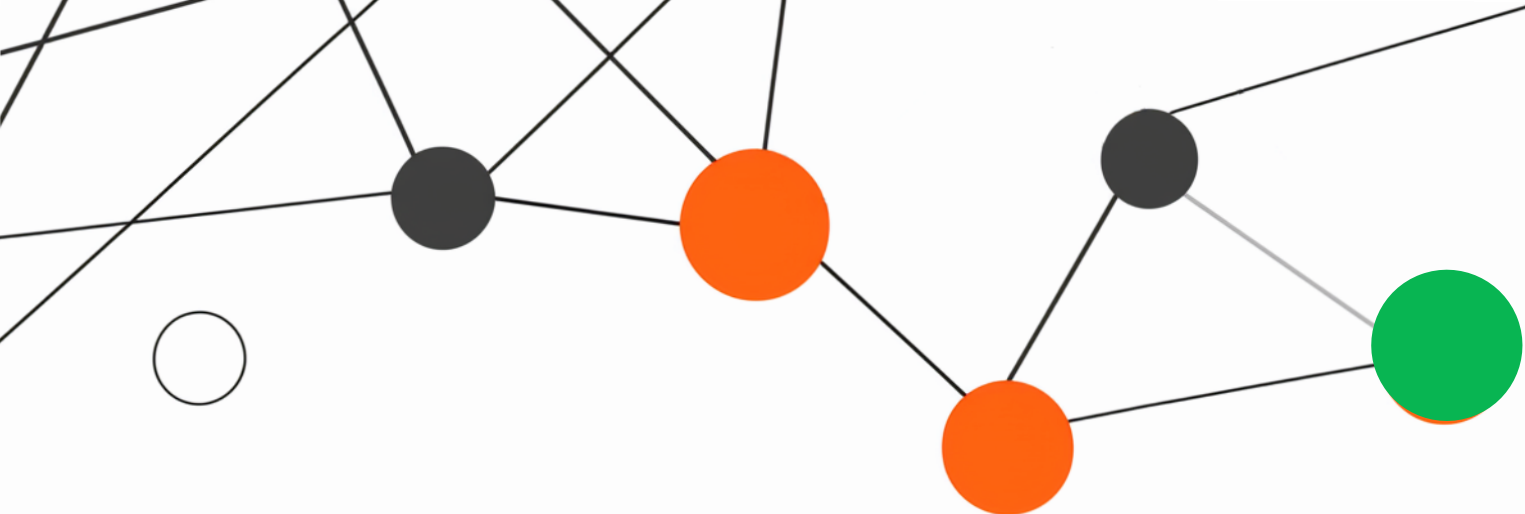
Observation 1: Det kan finnas brister i dataskyddet för generativa AI-applikationer

Generativa AI-applikationer som snabbt tagits i allmänt bruk är vanligtvis gratis att använda och deras funktioner är avsedda för konsumentbruk. När de används i undervisningen är det bra att göra en bedömning av huruvida de är avsedda för privata användare eller organisationer.

I undervisningen kan minimikravet för dataskyddet i chattbotar som baserar sig på stora språkmodeller anses vara att de instruktioner eller promptar som matas in i en bot inte används till att träna språkmodellen i fråga. Personuppgifter som eventuellt ingått i promptar ska inte användas till att utveckla AI-applikationen. Om sådana saker inte framgår av dokumentationen i applikationen är det ett tecken på bristande dataskyddskompetens hos tjänsteleverantören.

Det kan finnas brister också i säkerhetsfunktionerna i applikationerna. Om det i applikationen till exempel är möjligt att dela AI-generad text eller bild via en delningslänk kan en sådan funktion lätt innebära att personuppgifter läcks ut till utomstående antingen av misstag eller avsiktligt till exempel i mobbningssyften. Utbildningsanordnaren ska kunna övervaka hur tjänsten används – inklusive delningsfunktionerna – för att kunna ingripa vid missbruk och avlägsna olämpligt innehåll. Det är önskvärt att alltför omfattande delningsfunktioner kan stängas av.

Bristerna kan också gälla helt grundläggande dataskydd som möjligheten att få ut transparent information om behandlingen av personuppgifter



eller att granska egna uppgifter. Åtminstone användarna på administratörsnivån bör ha möjlighet att granska eller ladda ner användarnas uppgifter i digitalt format, så att utbildningsanordnaren ska kunna fullgöra sitt ansvar som personuppgiftsansvarig.

Observation 2: Generativ AI kan ha för omfattande rättigheter att komma åt lagrade uppgifter

I molntjänster som är avsedda för utbildningsanordnare kan det vara möjligt att tillåta generativ AI, som chattbotar, att komma åt alla filer och andra data som sparats i alla organisationens molntjänster. I sådana fall finns det en stor risk för att personuppgifter med koppling till undervisning används genom generativ AI för ändamål som inte har med undervisningen att göra.

Överlag är det viktigt att vara noga med vilka uppgifter generativ AI ges tillgång till – eller matas med i form av promptar. En bra tumregel är att ge AI tillgång endast till vissa typer av data som man är säker på inte innehåller personuppgifter. Ifall man vill använda AI för behandling av elevernas personuppgifter är det nödvändigt att göra en riskbedömning, i enlighet med beskrivningen ovan, så att man kan vara säker på att behandlingen följer gällande dataskyddsbestämmelser.

Observation 3: Eleverna kan få möjlighet att använda generativ AI utan avtal om behandling av personuppgifter

Samma tjänsteleverantörer erbjuder både molntjänster som är avsedda för utbildningsinstitutioner och generativa AI-

applikationer som i sin tur är fritt tillgängliga för konsumenter, och det leder vidare till att det är oklart vad webbtjänsterna som erbjudits utbildningsinstitutionerna egentligen innehåller. Med användarnamn som tillhandahålls av en utbildningsinstitution kan det vara möjligt att också använda generativa AI-applikationer som inte ingår i avtalet som utbildningsinstitutionen ingått med tjänsteleverantören. Molntjänster erbjuder vanligtvis flera alternativa licenser, och i en del av dem ingår en AI-applikation och i andra inte. Då beror det på licensen huruvida det är utbildningsanordnaren eller tjänsteleverantören som är den personuppgiftsansvariga för applikationen.

Utbildningsanordnaren gör därför klokt i att kontrollera inställningarna för molntjänsterna när de tas i bruk, så att eleverna med sina inloggningsuppgifter inte kan använda AI- och andra webbapplikationer, som inte omfattas av avtalet med tjänsteleverantören när det gäller behandlingen av personuppgifter.

Observation 4: Avtal om behandling av personuppgifter är inte alltid heltäckande

Oavsett vilken applikation det är fråga om ska utbildningsanordnaren kontrollera huruvida det är möjligt att ingå ett avtal om behandling av personuppgifter med tjänsteleverantören eller om det ingår i serviceavtalet. Om så inte är fallet ska elevernas personuppgifter inte sparas i applikationen utan ett separat begärt samtycke till den typen av utlämnande och användning av personuppgifter. Läraren får med fördel kontrollera utbildningsanordnarens riktlinjer för hur ett samtycke ska begäras och huruvida lärare får begära samtycken till användning av applikationer.

I tjänsteleverantörens standardavtal står det inte nödvändigtvis vilka alla personuppgifter som ska behandlas och för vilka ändamål. Då är det möjligt att uppgifterna som sparas i tjänsten kan användas till exempel vid produktutveckling, vid träning av AI-modeller eller för tjänsteleverantörens egna ändamål. Om det i applikationen ingår funktioner som förutsätter behandling av personuppgifter men som inte omfattas av avtalet, kan leverantören till den delen anses vara den personuppgiftsansvariga. Elevernas uppgifter hamnar i så fall hos utomstående parter och används för ändamål som inte har med undervisningen att göra.

Utbildningsanordnaren ska alltid kontrollera att ett gällande avtal följer den allmänna dataskyddsförordningen och att alla personuppgifter behandlas endast för de ändamål som den personuppgiftsansvariga fastställt.¹⁰

Ett varnande exempel på detta gäller användningen av tjänsten Google Workspaces for Education, som år 2021 ledde till en anmärkning utfärdad av Dataombudsmannen.¹¹ Beslutet har inte ännu vunnit laga kraft.

Dataskyddsmyndigheten i Danmark har också påpekat viss problem med användningen av Googles molntjänster inom kommunal grundläggande utbildning.¹²

Gyllene regler för användning av AI i undervisningen

Om AI-applikationer används i undervisningen är det viktigt att beakta å ena sidan utbildningsanordnarens skyldigheter som personuppgiftsansvarig och å andra sidan lärarnas tillräckliga dataskyddskompetens, så att man kan agera rätt i olika situationer i praktiken. Det senare förutsätter att dataskydd och användning av AI beaktas i lärarfortbildningen.

Dataskydd kan kännas som en komplex helhet, men i undervisningen kan man beakta det genom att börja med enkla grundläggande saker. De härrör i sin tur från dataskyddsprinciperna som utgör kärnan

i den allmänna dataskyddsförordningen.

Exempel på dessa principer är principerna om användningsändamålet och uppgiftsminimering, som slår fast att personuppgifter endast får användas för syften som slagits fast på förhand och att uppgifterna som behandlas ska begränsas till ett minimum. En tredje princip som det är bra att komma ihåg är transparensprincipen, som i korthet går ut på att alla ska kunna veta vilka uppgifter om dem som behandlas och hur. Utgående från dessa tre principer är det tryggt att börja använda också AI-applikationer.

De gyllene reglerna nedan gäller för användning av AI i undervisningen i ljust av dataskyddet. Reglerna är avsedda för lärare, så att de kan beakta dataskyddet i undervisningssituationer i praktiken:

¹ Europaparlamentets och rådets förordning (EU) 2016/679 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter och om upphävande av direktiv 95/46/EG (allmän dataskyddsförordning), <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=CELEX:32016R0679>

² Lag om grundläggande utbildning 628/1998, <https://www.finlex.fi/sv/lagstiftning/1998/628>

³ Silvennoinnen, E., Tedre, M. & Valtonen, T. (2024). Datafikoituvu peruskoulu – tasapainoilua lapsen henkilötietojen suojan ja opetuksen tavoitteiden välillä, *Lakimies*, 122(5), 655–678, <https://journal.fi/lakimies/article/view/143755>

⁴ ibid.

⁵ Arbetsgruppen för skydd av personuppgifter. (2017). Riktlinjer om konsekvensbedömning avseende dataskydd och fastställande av huruvida behandlingen "sannolikt

leder till en hög risk" i den mening som avses i förordning (EU) 2016/679, <https://tietosuoja.fi/documents/6927448/8316711/Riktlinjer+om+konsekvensbedomning+avseende+dataskydd.pdf>

⁶ Dataombudsmannens byrå. (2021). Instruktioner för konsekvensbedömning avseende dataskydd, <https://tietosuoja.fi/documents/6927448/66036250/TVA+ohje.pdf?ff0b6e1b-5b89-e85e-a2e5-6c4bd4c0ccfc/TVA+ohje.pdf?i=1639729535787>

⁷ Europaparlamentets och rådets förordning (EU) 2016/679 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter och om upphävande av direktiv 95/46/EG (allmän dataskyddsförordning), <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=CELEX:32016R0679>

⁸ Dataombudsmannens byrå. (n.d.). Bedöm riskerna och planera åtgärderna för att genomföra dataskyddet, hämtad 21.2.2025, <https://tietosuoja.fi/sv/bedom-riskerna>

⁹ DigiFinland. (26.6.2024). Cirrus-hanke: Tapausesimerkit, <https://digifinland.fi/wp-content/uploads/2024/06/Cirrus-Tapausesimerkit-2024-v1.0.pdf>

¹⁰ Europeiska dataskyddsstyrelsen. (2021). Riktlinjer 07/2020 angående begreppen personuppgiftsansvarig och personuppgiftsbiträde i GDPR, version 2.0, https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_sv

¹¹ Dataombudsmannen. (30.12.2021). Henkilötietojen käsittelyn laimukaisuus ja siirto kolmansiin maihin koulun opetusohjelman käytössä, <https://www.finlex.fi/fi/viranomaiset/tsv/2021/20211503>

¹² Dataombudsmannens byrå. (2.2.2024). Den danska dataskyddsmyndigheten utfärdade ett beslut om användningen av Googles programvara i grundskolan, <https://tietosuoja.fi/-/tanskan-tietosuojaviranomainen-antoi-paatoksen-googlen-ohjelmistojen-kaytosta-peruskouluissa>

Gyllene regler för användning av AI i undervisningen

Gör så här...

- Lär dig om möjligheterna med AI
- Använd endast tillåtna applikationer i undervisningen
- Använd personuppgifter endast i undervisningssyfte
- Ha koll på riktlinjer och anvisningar kring dataskydd
- Be om hjälp vid dataskyddsproblem

och inte så här...

- Lita inte blint på AI-applikationer
- Ge inte personuppgifter till AI-applikationer
- Lägg inte ut elevernas uppgifter utan tillstånd
- Bryt inte mot gällande anvisningar
- Låt inte bli att berätta om problem som identifierats

21. Bilaga: Faktabaaris anvisningar till elever

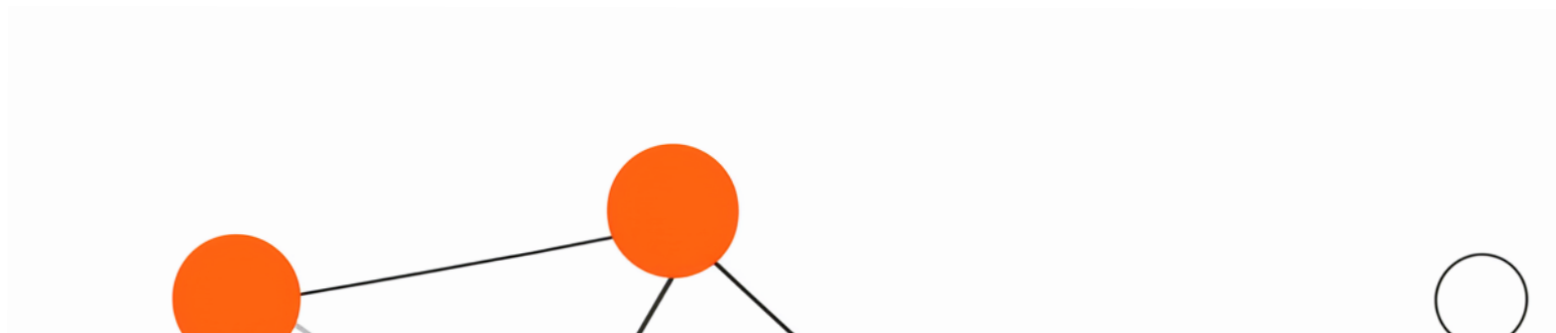
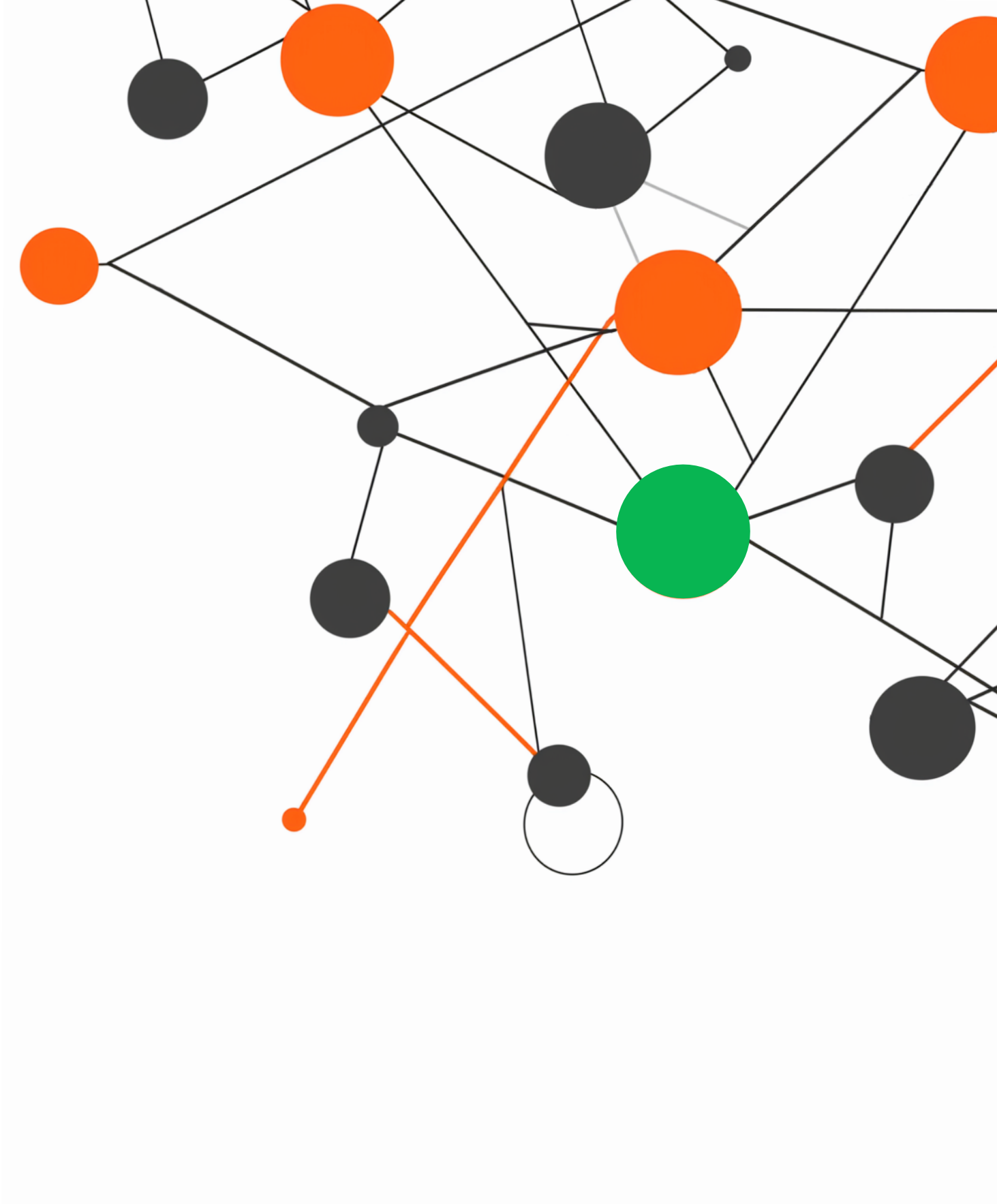
KARI KIVINEN, FAKTABAAARI

AI är ett verktyg – använd det ansvarsfullt!

- **Kontrollera reglerna:** Ta reda på vilka AI-verktyg som är tillåtna i skolan och i vilka situationer.
- **Var kritisk – kom ihåg källkritiken:** AI kan göra misstag. Granska alltid det som AI genererat och källhänvisningar som AI gett och jämför informationen AI ger med andra källor.
- **Var ärlig – ange källorna:** Förklara öppet vad som gjorts med hjälp av AI på samma sätt som du anger alla andra källor. Ibland är det bra att öppet berätta vilken typ av promptar som har använts.
- **Följ instruktionerna:** Följ instruktionerna du får av läraren.
- **Skydda din egen integritet:** Mata aldrig in konfidentiell information i ett AI-verktyg.
- **Var försiktig:** Följ säkerhetsanvisningarna.
- **Be om hjälp:** Om du är osäker kan du be om råd.
- **Hjälp andra:** Dela med dig av din kunskap till andra.
- **Sätt gränser:** Använd AI-applikationer med måtta, så att användningen av apparna inte tar upp all din tid. Kom också ihåg att AI använder mycket energi – använd den som stöd endast i viktiga uppgifter.
- **Agera ansvarsfullt:** Se till att du inte gör intrång på upphovsrätten eller plagierar andras arbete.

Kom ihåg:

- AI ersätter inte din egen förmåga att tänka. Artificiell intelligens stöder din egen intelligens!
- AI kan vara ett användbart verktyg, men det måste användas ansvarsfullt.
- AI är inte en människa, och den kan inte tänka självständigt.
- Det är du som ansvarar för hur du använder AI och det som AI genererar.



STOP TÄNK KONTROLLERA



peli.faktabaari.fi