

När AI träder in i naturvetenskaps- och teknikklassrummen: Etiska dilemman för lärare

Karin Stolpe
Gunnar Höst

Linköpings universitet
Linköpings universitet och NATDID

Artificiell intelligens (AI) förekommer allt mer i diskussioner om skola och utbildning. En utmaning för lärare är att följa med i den tekniska utvecklingen och bristen på adekvat kompetensutveckling. Men även etiska frågor är viktiga, något som diskuteras relativt lite i relation till den svenska skolan. Etiska frågor handlar bland annat att som lärare gå bortom färdiga råd och själv reflektera etiskt kring sin användning av AI i klassrummet. Syftet med artikeln är att lyfta forskning som presenterar etiska principer relevanta för AI i skolan på olika organisatoriska nivåer. Vi ger utifrån detta exempel på etiska dilemman i relation till AI, eleven, läraren och skolan.

Inledning

Artificiell intelligens (AI) förekommer allt mer i diskussioner om skola och utbildning. När AI-verktyget ChatGPT lanserades 2022 väckte det inledningsvis en stor oro för hur AI skulle kunna inverka på elevernas kunskaper och risken för att de ska fuska. När det handlar om lärares egen användning av AI har diskussionen ofta handlat om utmaningen att följa med i den tekniska utvecklingen och bristen på adekvat kompetensutveckling.

Skolverket genomförde under våren 2024 en undersökning över gymnasielärares användning och hantering av AI [1]. Med en enkät har Skolverket undersökt om lärarna använder AI-tjänster för utformande av lektionsuppgifter, för att undervisa om AI, för bedömning av elevuppgifter eller i samband med kompetensutveckling. I rapporten framkommer det att användningen fortfarande är sparsam.

Lärarna lyfter fram att de saknar egen kunskap som ett av skälen till den sparsamma användningen. Andra skäl rör bristande tillförlitlighet hos verktygen, vilket ger en risk att eleverna lär sig felaktiga faktakunskaper, och en oro för att eleverna ska bli mer passiva och mer beroende av teknik. Det framkommer också ett exempel från en lärare som helt enkelt tycker att det är roligt att själv planera sina lektioner som en del av läraryrket.

De tekniska och kunskapsmässiga aspekterna beskrivs således i Skolverkets rapport. Under ytan hos resultaten, och hos riktlinjer och rekommendationer såsom Skolverkets råd kring användning av AI [2], ligger etiska avvägningar och frågeställningar. Men trots att etiska frågor nu alltså aktualiserats i samtalet, diskuteras de etiska aspekterna kring AI relativt lite i relation till den svenska skolan. En utmaning för lärare är därför att gå bortom färdiga råd, för att själv reflektera etiskt kring sin användning av AI i klassrummet.

Den här texten syftar till att adressera denna utmaning genom att lyfta forskning som presenterar etiska principer relevanta för AI i skolan på olika organisatoriska nivåer. Utifrån forskningslitteratur och policydokument kommer vi sedan att lyfta några exempel på etiska dilemman som kan fungera som utgångspunkt för att diskutera AI-användning i skolan.

Etiska principer för AI i skolan

Etiska reflektioner handlar om resonemang och avvägningar i frågor där olika värden och intressen står i konflikt med varandra. Ett sätt att förtydliga områden med centrala etiska frågeställningar är genom ramverk som identifierar viktiga principer. När det gäller AI har flera sådana ramverk presenterats av olika aktörer (t.ex. UNESCO [3]). I en studie gick Jobin med flera igenom inte mindre än 84 sådana ramverk för att identifiera etiska principer som är viktiga att förhålla sig till [4]. Forskarna kom fram till följande elva etiska principer för användning och utveckling av AI i samhället i stort:

- **Transparens:** Det är viktigt att det är tydligt hur AI-system fungerar, var de finns och hur data samlas in.
- **Rättvisa:** AI-system ska inte bidra till att förstärka fördomar (bias) om eller bidra till diskriminering av exempelvis utsatta eller marginaliserade grupper. Rättvis tillgång till AI-system är också viktigt, och AI-stödda beslut behöver kunna utmanas.
- **Icke-skadlighet:** Individer eller grupper ska inte fara illa på grund av AI-system. Sådana effekter kan exempelvis handla om kränkning, kroppsskada eller cyberkrigföring.
- **Ansvar:** Ansvarsförhållanden vid användande av AI behöver klargöras, så att användare skyddas.
- **Skyddande av privatliv och integritet:** Individens privatliv ska respekteras, till exempel genom att upprätthålla dataskyddsförordningen.
- **Välgörande:** AI-system ska bidra positivt till människors välmående, till exempel genom att bidra till fred eller ekonomisk utveckling.
- **Frihet och autonomi:** Människors frihet att välja själva och att bestämma över sina liv ska inte motverkas av AI. Det kan också handla om frihet från till exempel övervakning och manipulering.
- **Förtroende:** Forskning och utveckling inom AI behöver gå att lita på.
- **Hållbarhet:** Utveckling och användning av AI-system behöver minimera sitt ekologiska avtryck, och kan bidra till en hållbar utveckling både socialt och miljömässigt.
- **Värdighet:** AI-system ska bibehålla eller öka mänsklig värdighet.
- **Solidaritet:** Fördelar och möjligheter med AI behöver spridas så att inte den sociala sammanhållningen hotas, till exempel när det gäller jobb som försvinner på grund av AI.

Resultatet pekar ut viktiga områden för etiska bedömningar, även om alla principer inte uttryckligen fanns i alla ramverken. De togs heller inte upp på samma sätt i olika ramverk eller ledde till samma rekommendationer. Forskarna såg dessutom att de första fem principerna (Transparens, Rättvisa, Icke-skadlighet, Ansvar och Skyddande av privatliv och integritet) fanns i mer än hälften av ramverken. Det tolkar de som att tankarna kring etik och AI konvergerar runt dessa principer.

Allmänna principer av det här slaget pekar alltså ut viktiga områden för etisk reflektion som kan vara relevanta i många sammanhang. Men det finns också etiska principer som gäller mer specifikt när det gäller utbildning. Det kom Adams med flera [5] fram till när de gick igenom fyra etiska ramverk specifikt framtagna för AI i utbildning från organisationer som World economic forum och Unicef. Förutom de elva principerna i punktlistan ovan, hittade forskarna fyra områden som särskilt behöver bevakas när det gäller utbildningssektorn [5]:

- **Beakta pedagogisk lämplighet:** Personer inom utbildningssektorns alla nivåer behöver uppmärksamma och avväga på vilket sätt AI kan underlätta utveckling och lärande för elever, men också hur det kan leda till att annan utveckling och lärande minskar.
- **Bevaka barns rättigheter:** Barn är särskilt sårbara, och barns rättigheter är därför centrala och viktiga att skydda även när det gäller AI.
- **Utveckla AI-litteracitet:** En väl utvecklad AI-litteracitet är viktig för såväl elever som lärare och andra personer inom utbildningssektorn. AI-litteracitet inkluderar förmågor, begrepp och etiska överväganden som omger skapande och användning av AI. Här ingår till exempel

att kunna vara en skicklig användare av AI, men också att förstå fördomar i data och vilka rättigheter man som användare av AI har.

- Skydda lärares välmående: Lärares välmående är komplext och inbegriper såväl den fysiska och mentala hälsan, som förmågan att kunna hitta en jämvikt mellan resurser och de krav som ställs på lärare.

Forskningen har alltså identifierat viktiga etiska principer för användning av AI generellt och specifikt inom skolans värld. Som framgår av beskrivningen av principerna är de allmänt hållna och behöver sättas i en kontext för att det ska gå att använda dem för vägledning i valsituationer och avvägningar. I det följande resonerar vi därför om olika exempel på hur de etiska principerna kan vara relevanta för den svenska skolans undervisning inom naturvetenskap och teknik.

AI i det svenska Nv/Te-klassrummet kan innebära etiska dilemman

Genom att applicera de etiska principerna från ramverk såsom de vi beskrivit ovan på en kontext eller situation visar det sig ofta att olika principer eller värden ställs mot varandra. De etiska dilemman som då uppstår har inga enkla svar, men kan vara användbara för att tydliggöra vilka val och konsekvenser som är relevanta. Vi lyfter här tre olika områden där etiska dilemman kan uppstå som är relevanta för lärare i den svenska skolans undervisning om naturvetenskap och teknik.

Eleven och AI - Hur skyddar vi barnens integritet samtidigt som vi ger dem kunskap om AI?

I Sverige är FN:s konvention om barnets rättigheter lag. I barnkonventionen ingår bland annat att alla barn ska ha samma rättigheter och lika värde. Barnets bästa ska beaktas vid alla beslut som rör barn och alla barn ska ha rätt att uttrycka sin mening och få den respekterad. Barnkonventionen är därför tillämplig i diskussionen om etiska principer för AI i utbildning [5]. Artikel nummer 16 beskriver exempelvis att varje barn har rätt till ett privatliv, och artikel 17 säger att barn ska ha rätt till information via till exempel internet. Staten ska uppmuntra att det skapas material som är av värde för barn och som inte är skadligt. Varje barn har även rätt till utbildning. Men när AI inkluderas i utbildningen uppstår ett dilemma. Å ena sidan behöver elever bekanta sig med AI-verktyg för att förstå dessa och kunna samarbeta med dem för att förstärka sina egna individuella förmågor. Å andra sidan behöver elevernas integritet skyddas. Det behöver vara transparent för eleverna hur data delas och används när de använder AI-verktyg [6].

I policydokument för AI i skolan diskuteras bland annat att barn har rätt att förstå vad AI är, och att de har rätt att få handledning i hur det fungerar. UNICEF föreslår till exempel att det är centralt att informationen som når barn är åldersadekvat. Adams med flera [5] har noterat att det i sådana policydokument ofta saknas en diskussion om rätten att få data som samlats in om en person raderad. När en elev använder ett AI-verktyg kan hen avslöja personlig information. Eftersom de flesta AI-verktyg idag kräver att man är inloggad, är det möjligt att spåra informationen tillbaka till individen. Det handlar kanske ofta om att eleven ställer skolrelaterade frågor, som vilket ämne som bildas när klorgas reagerar med vätgas. Men genom frågor som till exempel om det är normalt att som kille bli kär i en annan kille, eller hur man enklast blir av med fett runt magen, så kan information som egentligen är privat för barnet både sparas i en chatthistorik, och delas med AI-verktyget. Beroende på hur frågorna är ställda riskerar barnet att oavsiktligt dela med sig av information som inte var meningen att dela, och som också kan vara svårt att sedan ta tillbaka. Det är därför viktigt att elever, oavsett ålder, får kunskap om hur AI-verktygen fungerar när det gäller hantering av den delade informationen. Dessutom kan skolor ha upphandlade programvaror som skyddar användarnas integritet och inte delar med sig av data.

Utöver rätten till personlig integritet kan man också fråga sig om elever (och deras vårdnadshavare) har rätt att önska att inte använda AI-verktyg inom ramen för skolan. Vilket

ansvar har lärare att förbereda elever för att leva i morgondagens samhälle, där AI kanske är en integrerad del [5]? Till exempel lyfts principen om inkludering, det vill säga att AI-systemen ska vara lika tillgängliga för alla, oavsett ekonomiska förutsättningar, demografi, kultur och funktionsnedsättningar fram som viktig inom utbildningssektorn [7].

Adams med flera menar att om man ser arbete med AI-verktyg som ett samarbete mellan människa och AI, snarare än att det bara är en användning av AI, så blir andra frågor relevanta [5]. AI kan då förstås som att det kompletterar eller förstärker mänskliga förmågor, snarare än att AI tillskansar sig och ersätter mänskliga förmågor. Utifrån det här perspektivet blir det alltså viktigt att alla elever får en möjlighet att lära sig att samarbeta med AI som ett led i utbildningen.

Läraren och AI - Vilken roll spelar lärares professionella omdöme?

AI ger lärare möjligheten att konstruera en lektionsplanering [8]. Med hjälp av en instruktion (prompt) som beskriver vad läraren är ute efter kan denne inom loppet av några sekunder få en lektionsplanering bestående av en introduktion, en föreläsningsdel med förslag på begrepp som ska introduceras, och vilka exempel och bilder som läraren kan använda sig av. Det går också att få förslag på elevuppgifter som eleverna kan arbeta med enskilt eller i grupper. Dessutom är det möjligt att skapa hela terminsplaneringar på samma sätt.

Vid en första anblick kan detta tyckas attraktivt. AI-verktygen skulle kunna hjälpa till att minska lärares arbetsbörda i det här avseendet. Den tiden skulle istället kunna användas till att stötta elever som behöver mer hjälp. När lärare skapar den här typen av lektionsplaneringar behövs dock ett professionellt omdöme för att bedöma såväl att innehållet är korrekt, samt att det följer den svenska läroplanen.

De språkmodeller som används av exempelvis ChatGPT eller Copilot bygger på statistiska modeller [9]. Det svar som är mest sannolikt korrekt är det som kommer att dyka upp som svar. Det innebär att om modellen är tränad utifrån internationella data, där data från Sverige endast utgör en minimal del, så finns det stor risk för att andra länders läroplaner kommer att statistiskt dominera materialet, och därmed få en stark inverkan på svaret.

I ett ämne som biologi finns idag en ganska stor likformighet mellan olika länders läroplaner, vilket gör att de lektionsplaneringar som skapas där torde ligga nära den svenska läroplanen. Men ett ämne som teknik, som i princip inte existerar i andra länder än Sverige, behöver man som lärare vara mer uppmärksam på vilka svar man får. Här blir också sättet att ställa frågan, alltså hur man promptar, central. Om man ber om en lektionsplanering bör man således vara så specifik som möjligt med vilket innehåll man vill att lektionen ska ha. En annan aspekt är behovet av att lärare är självständiga i relation till AI-verktygen. Det kan finnas en risk att man låter sig styras alltför starkt av de förslag som man får av AI, istället för att utgå från egna värderingar och det egna professionella omdömet [10].

I de policydokument om AI som Adams med flera [5] har analyserat, lyfts även att AI kan utgöra en potentiell risk för lärares välmående. Genom att introduktionen av AI ställer nya krav på lärare och lärares arbete kan deras arbetsbelastning öka. Det finns också en farhåga att användandet av AI-verktyg i själva verket kommer att leda till att lärare lägger mer tid på att planera lektioner, värdera dessa, och skriva nya instruktioner. Den tiden kommer i så fall att minska den tid som lärare har för att bygga relationer med eleverna.

Dilemmat som framkommer blir således det mellan att å ena sidan spara tid, å andra sidan behöva ägna tid åt att granska det material som framkommer från AI, skriva fungerande promptar och lära sig de nya verktygen. Lärares professionella omdöme ställs på spel.

Skolan och AI - Vilken framtid bidrar skolorna till?

Inom utvecklingen av AI-system lyfter man ofta fram AI som en viktig faktor för att lösa klimatkrisen (t.ex [11]). AI ses som en möjlighet att lösa komplexa problem genom att utveckla mer hållbara energisystem eller underlätta forskningen om nya och mer miljövänliga material. Samtidigt finns en medvetenhet bland forskare inom AI att systemen har stort klimatavtryck [12]. Det gäller energiåtgången för att driva systemen såväl som producera material och tillverka av de komponenter som används i datorerna [3].

Det finns alltså ett dilemma mellan att å ena sidan utveckla AI-system som är tillräckligt effektiva för att kunna hjälpa till att lösa klimatfrågan, å den andra sidan att AI-systemen driver på klimatkrisen genom sin stora energiåtgång. All AI-användning bör således ta i beaktande frågan om hållbarhet [7]. I skolans undervisning om hållbar utveckling kan detta dilemma bli extra påtagligt. Hur ska undervisningen både kunna leda till att eleverna får möjlighet att utveckla sin AI-kompetens, och samtidigt inte bidra till en ökad användning av AI-system som kräver stora mängder energi? Det här dilemman kan i sig självt utgöra underlag för intressanta diskussioner i klassrummet.

Den mesta forskning som görs inom området handlar om att minska energiåtgången i själva träningsfasen, alltså i den fas där modellerna tränas på data [12]. Det finns fortfarande väldigt lite forskning som studerar hur stor påverkan som användandet av AI-systemen har på klimatet. Det finns dock källor som pekar på att energiåtgången i samband med användning vida överstiger träningsfasen [13]. När serverarna arbetar omvandlas dessutom elektrisk energi till värme. För att serverarna inte ska överhettas behöver de alltså kylas, vilket i dagsläget görs med vatten, något som också har potentiella konsekvenser för hållbar utveckling.

Så är AI-användning förenligt med målet om en hållbar utveckling, ett minskat klimatavtryck, och en större social rättvisa? Vi befinner oss för tillfället i en utveckling som går i rasande takt. Allteftersom modellerna har tränats, kan tekniken kanske också hjälpa oss med lösningar på de här komplexa problemen, precis som många AI-forskare hoppas. AI har på så sätt en tvådelad roll, å ena sidan är AI en potentiell lösning på klimatkrisen, å andra sidan bidrar AI till en ökad klimatpåverkan [14]. Ett av de största problemen idag enligt Vieth-Ditlmann [13] är att det saknas tillförlitliga data på de faktiska utsläppen. Utan sådan data står politiska beslutsfattare utan möjlighet att kunna sätta press på industrin.

Samtidigt lyfter också forskare att vi kan hamna i det som brukar benämnas som "Jevons paradox" [14]. William Stanley Jevon var en brittisk ekonom som levde under 1800-talet. Jevons paradox bygger på idén om att teknisk utveckling ökar effektiviteten vilket leder till en minskad resursåtgång. Det i sin tur tenderar att medföra ökad, snarare än minskad, konsumtion av resursen i fråga. Så här skriver Nordgren: "förbättrad energieffektivitet genom användning av AI skulle kunna öka kravet på AI-applikationer i andra samhällsområden och därigenom öka den totala användningen av AI i samhället" ([14], sid. 4, vår översättning). En åtgärd som Nordgren lyfter fram som en del av lösningen är att minska användningen av AI där den inte behövs. Inom sjukvård, eller som ett verktyg för att minska klimatpåverkan, är energiåtgången acceptabel. Men att använda AI för ren underhållning kan vara mer tveksamt.

Avslutande diskussion

Användandet av AI-verktyg i utbildningssammanhang är än så länge ett relativt nytt fenomen. Trots det är det inte längre en fråga om AI kommer att förändra utbildning, utan snarare *hur*. I artikeln har vi lyft fram forskning om vilka etiska dimensioner som vi behöver ta hänsyn till i denna utveckling, och gett exempel på några av de etiska dilemman som kan uppstå i relation till AI i skolan. Vi hoppas att detta kan ge inspiration till lärare och andra verksamma i skolan att själva reflektera över vilka värden som står mot varandra i olika valsituationer.

Även om vi bara har plockat upp några få etiska dilemman med relevans för skolan så finns det andra frågor som också behöver diskuteras. Det handlar till exempel om frågor om demokrati, källkritik och likvärdighet. Dessutom lever vi i ett samhälle där AI hela tiden utvecklas. Det enda vi vet om framtiden är att den inte kommer att likna hur det ser ut idag. Därför behöver elever förberedas för en värld i förändring, där de måste vara beredda att lära sig nya AI-verktyg, och nya sätt att tänka kring lärande och kunskap.

Sådana frågor kan också fungera som innehåll i undervisningen. Genom att inkludera elever i diskussioner om AI och hur dessa verktyg ska hanteras i skolan får eleverna också agens över sin egen roll. Det i sin tur kan även leda till att elevernas AI-litteracitet stärks och att de kan utveckla sin förmåga att resonera etiskt också över sin egen användning av AI utifrån sitt skolarbete.

En fråga som vi än så länge inte har berört i den här texten är den om att de flesta tillgängliga språkmodellerna, såsom ChatGPT, Copilot och andra, ägs och drivs av kommersiella företag. Dessa företag delar inte nödvändigtvis idén om gratis, transparent och demokratisk utbildning [9]. Även om det här inte är en fråga som en enskild lärare kan påverka, kan den behöva diskuteras i kollegiet.

En farhåga som ibland lyfts är att AI ska komma att ersätta lärare. Vad vi kan se finns det inget som tyder på det i nuläget. Det kanske snarare är tvärtom, att det är viktigare än någonsin med kritiskt tänkande lärare som kan fungera som mänskliga bollplank till elever [9]. Allteftersom AI utvecklas kommer verktygen kunna användas för individanpassad undervisning, bedömningsstöd för såväl prov som betygssättning och prediktion av studieval baserat på både prestationer och bakgrundsvariabler. Det är lärares professionella omdöme och de etiska diskussioner som förs i relation till denna utveckling som blir avgörande för var vi kommer att hamna i framtiden.

Författare



Figur 1. Karin Stolpe och Gunnar Höst.

Karin Stolpe arbetar som forskare och lärarutbildare vid Linköpings universitet. Hennes forskning handlar främst om lärares undervisning i teknik och naturvetenskapliga ämnen.

Gunnar Höst är huvudredaktör för ATENA Didaktik och biträdande föreståndare för NATDID, Nationellt centrum för naturvetenskapernas och teknikens didaktik. Han är också universitetslektor

i visuellt lärande och kommunikation vid Institutionen för teknik och naturvetenskap, Linköpings universitet.

Referenser

1. Artificiell intelligens i undervisningen. En lägesbild över gymnasielärares användning och hantering av AI under vårterminen 2024. Skolverket; 2024.
2. Råd om AI, Chat GPT och liknande verktyg. Skolverket. 2024; <https://www.skolverket.se/skolutveckling/inspiration-och-stod-i-arbetet/stod-i-arbetet/rad-om-ai-chat-gpt-och-liknande-verktyg>. Åtkomst september 2024.
3. UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2022. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
4. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*. 2019;1(9). <https://doi.org/10.1038/s42256-019-0088-2>
5. Adams C, Pente P, Lemermeyer G, Rockwell G. Ethical principles for artificial intelligence in K-12 education. *Computers and Education: Artificial Intelligence*. 2023;4. <https://doi.org/https://doi.org/10.1016/j.caeai.2023.100131>
6. Holmes W, Bialik M, Fadel C. Artificial intelligence in education. I: *Globethics*. 2023.
7. Nguyen A, Ngo HN, Hong Y, Dang B, Nguyen B-PT. Ethical principles for artificial intelligence in education. *Education and Information Technologies*. 2023;28(4). <https://doi.org/10.1007/s10639-022-11316-w>
8. van den Berg G, du Plessis E. ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*. 2023;13(10). <https://doi.org/10.3390/educsci13100998>
9. Klemke R, Jarodzka H. Locked in generative AI: The impact of large language models on educational freedom and teacher education. I: Searson M, Langran E, Trumble J, redaktörer. *Exploring new horizons: Generative artificial intelligence and teacher education*. Association for the Advancement of Computing Education (AACE); 2024. <https://www.learntechlib.org/p/223928/>
10. Akgun S, Greenhow C. Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*. 2022;2(3). <https://doi.org/10.1007/s43681-021-00096-7>
11. Artificial intelligence saving the natural world. *Föreanta nationerna*. 2024; <https://www.un.org/en/desa/artificial-intelligence-saving-natural-world>. Åtkomst september 2024.
12. Verdecchia R, Sallou J, Cruz L. A systematic review of Green AI. *WIREs Data Mining and Knowledge Discovery*. 2023;13(4). <https://doi.org/https://doi.org/10.1002/widm.1507>
13. Vieth-Dietlman K. Sustainable AI: a contradiction in terms?. *Algorithm watch*. 2024; <https://algorithmwatch.org/en/sustainable-ai-explained/>. Åtkomst september 2024.
14. Nordgren A. Artificial intelligence and climate change: Ethical issues. *Journal of Information, Communication and Ethics in Society*. 2023;21(1). <https://doi.org/10.1108/JICES-11-2021-0106>