Cavallin, C. (2026). Kunstig intelligens og tillit: Utfordringer for norske grunnskolelæreres profesjonsfaglige digitale kompetanse. I L. I. Hognestad, P. I. Kjærgård & D. Ø. Østereng (Red.), *Danning og samfunnsansvar: Utdanning og journalistikk i en digital tid* (s. 47–65). Fagbokforlaget. https://doi.org/10.55669/oa771202

2

Kunstig intelligens og tillit

Utfordringer for norske grunnskolelæreres profesjonsfaglige digitale kompetanse

Clemens Cavallin

This work is licensed under the Creative Commons Attribution 4.0 International License.
To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

Sammendrag: Dette kapittelet utforsker hvordan kunstig intelligens (KI) utfordrer tillitsbaserte relasjoner i den norske grunnskolen, og i hvilken grad gjeldende styringsdokumenter legger til rette for læreres profesjonsfaglige digitale kompetanse (PfDK) i møte med dette. Mens KI radikaliserer behovet for mediekritikk, argumenteres det for at et ensidig fokus på kritisk dekonstruksjon kan svekke den grunnleggende tilliten som all utdanning og danning bygger på. Gjennom en teoretisk drøfting av tillitsbegrepet og en analyse av læreplanverket (LK20) og Utdanningsdirektoratets ressurser, belyser teksten spenningen mellom kritisk tenkning og tillit i en ny digital virkelighet, hvor grensen mellom menneskelig og maskinell kommunikasjon blir stadig mer flytende.

Nøkkelord: Kunstig intelligens, LK20, PfDK, danning, tillit, grunnskole

Abstract: This chapter explores how artificial intelligence (AI) challenges trust-based relationships in Norwegian primary and lower secondary schools and the extent to which current policy documents support teachers' professional digital competence (PfDK) in facing this challenge. While AI radicalizes the need for media critique, it is argued that a one-sided focus on critical deconstruction may weaken the fundamental trust upon which all education and formation (*Bildung*) are built. Through a theoretical discussion of the concept of trust and an analysis of the national curriculum (LK20) alongside resources from the Norwegian Directorate for Education and Training, the text highlights the tension between critical thinking and trust in a new digital reality, where the boundary between human and machine communication becomes increasingly fluid.

Keywords: Artificial intelligence, LK20, PfDK, Bildung, trust, primary and lower secondary school

Innledning

Norge er, sammen med Danmark, blant landene i verden med høyest grad av sosial tillit – henholdsvis 72 og 74 prosent i 2022 (Ortiz-Ospina et al., 2024). Likevel står Norge, sammen med resten av verden, overfor en av de mest alvorlige truslene mot tillitsbaserte relasjoner, spesielt de som formidles via ulike medier: de nye multimodale formene for kunstig intelligens. Ettersom KI i prinsippet kan simulere alle former for menneskelig kommunikasjon, når behovet for en velfungerende mediekritikk et dramatisk høydepunkt. Man kan ikke lenger være sikker på at tekst, stemme, bilde eller video er «ekte». Dette problematiserer den grunnleggende tilliten til personer, institusjoner og samfunnsverdier som kreves for utdanning og danning (Bøhn, 2024).

Dette utgjør allerede et reelt problem i norske skoler, fremfor alt når det gjelder eksamensformer og innleveringer (Elstad & Eriksen, 2025, s. 91). Lærere er usikre på om elevene faktisk har skrevet tekstene de leverer inn, eller selv har løst de matematiske problemene de har arbeidet med hjemme. Den nye teknologien svekker altså generelt tilliten mellom lærer og elev. Det hjelper ikke å levere inn håndskrevne oppgaver på papir hvis arbeidet ikke er utført under tilsyn. Dermed blir undervisning i og øvelse på en radikalisert mediekritikk nødvendig – noe som lærere i skolen ikke er trent i, og som, når dette kapittelet skrives, ikke utgjør en integrert del av de norske lærerutdanningene.

En diskusjon om KI i skolen bør derfor ikke kun dreie seg om varianter av plagieringsproblematikk eller om hvordan teknologi kan virke forstyrrende. Den må berøre de grunnleggende tillitsrelasjonene som læring og evaluering er avhengige av (Grierson, 2018; Qin et al., 2020; Tschannen-Moran, 2017). Å fjerne teknologien fra skolen kan føre til gode resultater for læring og trivsel på kort sikt, men det innebærer samtidig et skritt i retning av økt kontroll og mistillit. De digitale ferdighetene og arbeidet med kritisk tenkning i norske skoler må derfor videreutvikles for å kunne håndtere denne nye situasjonen, samtidig som man bevarer og dyrker de tillitsbaserte relasjonene.

For å navigere i dette spenningsfeltet undersøker dette kapittelet hvordan norske styringsdokumenter og reguleringer legger til rette for grunnskolelærerens utvikling av en profesjonsfaglig digital kompetanse (PfDK) som inkluderer tillitsspørsmål. Dette er en avgjørende problemstilling i møte med den stadig mer komplekse relasjonen mellom kunstig intelligens og mellommenneskelig tillit. Hensikten med kapittelet er å avdekke sentral tematikk og

tendenser i styringsdokumenter og statlige videreutdanningsressurser, med et spesifikt blikk for relasjonen mellom tillit og kritikk. Etter en drøfting av sentrale begreper rettes analysen mot det nå gjeldende læreplanverket LK20, deretter mot Utdanningsdirektoratets (Udir) nettressurser og til slutt mot selve rammeverket for PfDK-arbeidet.

Tillit

Før vi analyserer styringsdokumentene, er det nødvendig å se nærmere på tillitsbegrepet og noen assosierte begreper. Tillit er grunnleggende både for utdanning og danning (Fisher & Tallant, 2020; Mitchell et al., 2018; Tschannen-Moran, 2017), og innebærer en sårbarhet for tillitsgiveren (eng. *trustor*) overfor tillitsmottakeren (eng. *trustee*) (Grimen, 2009, s. 60–71; Schoorman et al., 2007). Dette er spesielt markert når eleven er mindreårig. Barnet må ha tillit til at det læreren sier, og at det som formidles i læremidlene, er sant; ellers blir læring tilnærmet umulig. Også læreren må stole på eleven, noe som vitner om tillitens resiproke, om enn asymmetriske, karakter.

Beslektede, men forskjellige, begreper er respekt, anerkjennelse og toleranse. Respekt er en vurderende holdning som ikke alltid er basert på at det som respekteres, er moralsk godt. Iblant respekterer man noe kun fordi det er kraftfullt, eller fordi en person er dyktig til noe eller har høy sosial verdi. Respekt kan også være enten kognitivt eller emosjonelt motivert (Dillon, 2022). Motsatsen til respekt er forakt og respektløshet, mens motsatsen til anerkjennelse er krenkelse (Lund, 2017, s. 42). Anerkjennelse er i første rekke en bevisst pedagogisk holdning fra læreren overfor eleven, eller en positiv holdning til ideer og prinsipper. Tillit i streng forstand baserer seg derimot på en tro på en persons eller institusjons gode egenskaper og vilje – på stabile disposisjoner som går utover rene profittkalkyler. Tillit rettes ikke mot ideer på samme måte som respekt og anerkjennelse. Man har tillit til personer og institusjoner, men også – på et lavere nivå – til mekanismer. Det sistnevnte er spesielt interessant når grenselinjene mellom maskin og person blir utydelige, og når stadig mer antropomorfe – det tilsynelatende menneskelige – KI-agenter lanseres.

I diskusjonen om respekt og toleranse i relasjon til KI, innenfor det større feltet KI-etikk, står spørsmål om fordommer og skjevheter innebygd i selve modellen, basert på dens treningsdata, sentralt (f.eks. Hofmann, 2025; Peña-Garceda et al., 2026). Når det gjelder anerkjennelse, er KI primært ikke relevant, men det kan bli det i en sekundær betydning. Dette aktualiseres gjennom spørsmålet om hvorvidt svært kompetente, antropomorfe KI-agenter på noen måte skal anerkjennes som (simulerte) subjekter eller (elektroniske) personer. Dette vil for eksempel bli en aktuell problemstilling når en lærer inntar en anerkjennende holdning i mediert kommunikasjon med en elev, og det senere viser seg at læreren egentlig har interagert med en KI-agent – eller omvendt, når eleven oppdager at «læreren» er en simulert intelligens. Videre vil unge mennesker i stadig større grad inngå i relasjoner med kunstige agenter; relasjoner som noen ganger vil være av svært intim karakter, for eksempel av terapeutisk, seksuell eller åndelig art.

Det er slike simuleringer av menneskelig tenkning, emosjoner, vurderinger og handlinger som endrer arbeidet med digital kompetanse: Det dreier seg ikke lenger bare om kilde- og mediekritikk, men om eksistensielle spørsmål knyttet til det menneskelige. Dermed utvides diskusjonen om digitale ferdigheter. Et sentralt spørsmål blir om en KI-agent kan regnes som en person, eller om dette er helt utelukket. Tillit til kunstige agenter vil utvilsomt bli et viktig område for regulering og forskning (Palmer, 2024), men tilliten overskrider her det juridiske nivået og berører selve det menneskelige på et dypere plan. Dette vil gjøre perspektiver fra filosofi, litteratur og religion helt nødvendige for refleksjon og undervisning i skolen. For eksempel har den flytende grensen mellom biologisk ekte og kunstig menneskelighet lenge vært et fremtredende tema i nettopp litteratur og film, som i Ridley Scotts *Blade Runner* fra 1982 og Steven Spielbergs *A.I. Artificial Intelligence* fra 2001.

Kunstig intelligens' evne til å simulere menneskelig språk og kultur gjør også at kritikk av det antropomorfe – det tilsynelatende menneskelige – rykker frem i bevisstheten. For eksempel henviser Pitts og Motamedi (2025) til et skille mellom menneskelignende (*human-like*) og systemlignende (*system-like*) tillitsbaserte oppfatninger av teknologi. Førstnevnte karakteriseres av interpersonelle kvaliteter som velvilje og integritet, mens sistnevnte knyttes til egenskaper som funksjonalitet og pålitelighet. Jo mer KI-agenter og chatboter utstyres med antropomorfe trekk, desto mer relevante blir de affektive og moralske dimensjonene i vår omgang med dem. Dette innebærer at det

må undervises i en ny type mediekritikk i skolen; en kritikk som kan avsløre og dekonstruere antropomorfe former for kommunikasjon som i realiteten er maskinelle, systemiske funksjoner. Dette vil sannsynligvis påvirke samtlige medierte relasjoner i skolen – fra tekstoppgaver og e-poster til bilder, video og lyd.

Et presserende spørsmål er om en slik vidtgående «mistankens hermeneutikk» rettet mot det tilsynelatende menneskelige, kombinert med et posthumanistisk fokus på absolutt frihet (Corjescu & Doat, 2022, s. 33), vil komme til å problematisere selve læringsprosessene. All pedagogikk bygger fundamentalt på en forventning om sannferdighet – at læreren vil eleven vel, og at kunnskapen som formidles og mottas, er forankret i en felles, ekte virkelighet. Dersom teknologien tvinger frem en situasjon der balansen dramatisk tippes i favør av en konstant, dekonstruerende kritikk fremfor grunnleggende tillit, risikerer vi å undergrave den mellommenneskelige tryggheten klasserommet krever. Hvis elevene tvinges til systematisk å mistenkeliggjøre all menneskelignende kommunikasjon, vil dette til syvende og sist forplante seg til og farge alle relasjoner mellom virkelige mennesker. I det minste blir det stadig vanskeligere for både lærere og elever å skille mellom ekte og KI-genererte medier i indirekte kommunikasjon. Man kan ganske enkelt ikke lenger være sikker på at stemmen man hører i telefonen (eller liknende apparater), faktisk tilhører rektor, en medelev eller en forelder.

Hvis lærere og elever tvinges inn i en konstant tilstand av mistenksomhet, utfordres selve forutsetningen for læring: troen på en felles, ekte virkelighet. Når det blir stadig vanskeligere å skille mellom menneskelige og KI-genererte uttrykk, oppstår et institusjonelt behov for å gjenreise kontrollen, men det har en bakside.

KI og kontroll

De forandringene som KI-teknologien medfører, er ikke alltid helt nye, men innebærer ofte en forsterkning av allerede pågående prosesser. I løpet av de siste tiårene har mange utdanningssystemer og institusjoner i økende grad valgt å erstatte tillitsrelasjoner med «mål- og resultatbaserte styringssystemer»

(Sørreime, 2020, s. 11) og formelle, juridiske kontrakter. Den enkelte lærer opplever dermed et systemisk press bort fra tillitsbaserte relasjoner i undervisningen. Det er imidlertid vanskelig å se hvordan danning kan skje uten dyder som ærlighet, rettferdighet, velvilje og pålitelighet (Grimen, 2009, s. 42–44). En forutsetning for tillit er nettopp at den andre har slike dyder, og ikke kun opptrer som en nyttemaksimerende, rasjonell agent.

En mulig løsning som synes å få stadig større betydning, er å etablere skoler som en form for lavteknologiske soner hvor alle læringsoppgaver utføres. Her har ikke elevene tilgang til mobiltelefoner, smartklokker, smartbriller, datamaskiner eller internett. Teknologiske hjelpemidler kan i stedet introduseres gradvis på høyere trinn. Imidlertid øker en slik strategi avstanden mellom skolen og hjemmet, og til samfunnet for øvrig. I begynneropplæringen er dette kanskje primært en fordel, men når fokuset rettes mot å forberede elevene på et aktivt samfunnsliv, kan et slikt teknologisk skille bli problematisk.

Et tegn i Norge på en utvikling i retning av å skjerme læringen fra teknologi, er at kunnskapsminister Kari Nessa Nordtun i april 2025 ba Utdanningsdirektoratet om å vurdere konsekvensene av å ta bort digitale ferdigheter som en av de fem grunnleggende ferdighetene på 1. til 4. trinn (Regjeringen, 2025). I pressemeldingen refereres det spesifikt til kunstig intelligens som en av årsakene. Samtidig har den svenske regjeringen, basert på en utredning (Utbildningsdepartementet, 2025), vedtatt en lov som forbyr mobiltelefoner i skolen fra høsten 2026 (Helmerson, 2025).

En konsekvens av å fjerne teknologien er økt kontroll, noe som forsterker den tidligere forskyvningen fra tillit til mål- og resultatstyring. En slik strategi kan utvikle seg til et slags våpenkappløp mellom ulike KI-systemer: Noen systemer forsøker å unngå å bli avslørt, mens andre forsøker å avsløre KI-genererte medier. En slik utvikling kan svekke de demokratiske verdiene Norge ønsker å fremme i skolen, slik som medvirkning og gjensidig respekt (Udir, 2026, kap. 1.6). Hvis de datamaskinene og nettbrettene elevene bruker på skolen, ikke har tilgang til det åpne internettet eller avanserte KI-funksjoner, blir også den digitale kompetansen som oppøves med slike verktøy, av begrenset verdi for deltakelse i samfunnet for øvrig.

Før vi ser nærmere på kompetansepakkene og PfDK-rammeverket som direkte tematiserer KI, må vi først undersøke det nåværende læreplanverket for å se hvilke ressurser det tross alt tilbyr, og hvordan det kan utvikles når det neste læreplanverket skal utformes.

LK20

LK20 ble lansert to år før gjennombruddet for generativ KI med ChatGPT i 2022. Dette gjenspeiles i de nye læreplanene ved at fokuset på digitale ferdigheter først og fremst ligger på kildekritisk innhenting og bearbeiding av informasjon på nett, programmering og kreativ bruk av digitale ressurser (se også *Rammeverk for grunnleggende ferdigheter*, Udir, 2017). Det som peker mer spesifikt, om enn ikke eksplisitt, mot KI-teknologien, er begrepet «algoritmisk tenkning», som er koblet til programmering i matematikkfaget (Stovner et al., 2025). For eksempel er et kompetansemål på 8. trinn i matematikk å forstå «hvordan algoritmer kan skapes, testes og forbedres ved hjelp av programmering» (Udir, 2020, s. 12).

Begrepet tillit brukes om to ulike typer relasjoner i læreplanverkets overordnede del: dels om skolens holdning til elevene, og dels om relasjonen mellom skole og foreldre. I tråd med opplæringsloven § 1-1 skal skolen «møte elevane og lærlingane med tillit, respekt og krav og gi dei utfordringar som fremjar danning og lærelyst» (Opplæringslova, 2023). I kapittel 3.1 i *Overordnet del*, «Et inkluderende læringsmiljø», utdypes dette ved at skolen skal vise tillit til elevene og åpne opp for medvirkning, noe som innebærer at elevene tar ansvar. Videre, i kapittel 3.3, «Samarbeid mellom hjem og skole», slås det fast at tillit skal prege relasjonen og dialogen mellom skole og hjem.

Imidlertid brukes ikke tillitsbegrepet i *Overordnet del* først og fremst om elevenes relasjon til hverandre, eller om deres relasjon til skolen, lærerne eller andre institusjoner. I stedet dominerer begrepet respekt. For eksempel står det under «Demokrati og medvirkning» at dialogen «mellom lærer og elev, og mellom skole og hjem, må være basert på gjensidig respekt» (Udir, 2017, kap. 1.6).

Respekt skal videre ikke bare vises overfor personer, men også overfor ideer, som grunnleggende demokratiske verdier. Den første verdien som nevnes, er nettopp gjensidig respekt, noe som blir sirkulært, ettersom eleven altså skal respektere respekten. De andre verdiene som nevnes, er toleranse, tros- og ytringsfrihet, samt frie valg. Videre skal ideen om at mennesker er forskjellige, menneskeverdet, naturen, etablert viten, måter å utforske og skape på, og privatlivet respekteres. Andres overbevisninger og uenighet (Udir, 2017, kap. 2.5.2) skal også respekteres, men ikke respektløse ytringer (Udir, 2017, kap. 2.1). Denne grensen for respekt – det vil si at respekten ikke rettes mot

sin egen motsetning – innebærer at hva som defineres som respektløst, blir avgjørende for hvor inkluderende respektbegrepet forstås og brukes.

På liknende måte fungerer anerkjennelse ikke bare overfor personer, men også som en markør for gode verdier og abstrakte størrelser. Anerkjennelse skal ifølge *Overordnet del* rettes mot «barndommens og ungdomstidens egenverdi» (Udir, 2017, kap. 2), «flertallets rett og mindretallets rettigheter» (Udir, 2017, kap. 2.5.2), «mangfold» (Udir, 2017, kap. 3.1), «menneskeverdet», «ulikheter» (Udir, 2017, kap. 1.1), og «norsk tegnspråk … som fullverdig språk».

Begrepet toleranse forekommer imidlertid kun ett sted i *Overordnet del*, under kapittel 1.6 «Demokrati og medvirkning», som en av de grunnleggende demokratiske verdiene. Det kan defineres som en type anerkjennelse, nærmere bestemt anerkjennelse «av andres rett til å mene noe annet enn det man selv eller flertallet mener …» (Slåttholm Sagdahl, 2025). Det fungerer dermed ikke bredt som en bekreftelse av fundamentale prinsipper for norsk skole, men som en anerkjennelse av retten til variasjoner eller avvik fra slike prinsipper. I likhet med respekt har toleransen imidlertid en utydelig grense for hva som ikke kan tolereres – som i uttrykket «ikke tolerere de intolerante» – noe som ironisk nok gjør toleransen intolerant.

Samtidig skal barnet i skolen gradvis lære seg kildekritikk og mediekritikk, og med andre ord øve seg opp i kritisk tenkning, noe som fremheves allerede i formålsparagrafen. Barnet blir dermed i økende grad bevisst på rasjonelle prinsipper for å gi eller holde tilbake tillit. Tove Grete Lie peker på hvordan dette skaper en spenning i samfunnsfaget mellom kritisk tenkning og det å «realisere verdigrunnlaget i opplæringen» (Lie, 2022, s. 297). Men også i danningsbegrepet finnes en grunnleggende dobbelthet som gjenspeiler balansen mellom tillit og kritikk. På den ene siden står frihet og autonomi, og på den andre siden står en kulturs idealer, normer, historie og grunnleggende fortellinger som borgeren skal formes etter – noe som utvilsomt krever tillit (se Dyngeland & Engelsrud, 2024, s. 2–3; Sjöström & Eilks, 2020, s. 56; Tække & Paulsen, 2022, s. 6, 30).

I en bok fra 2024 argumenterer Åse Røthing, sammen med Elise Johannesen og Tonje Myrebøe, for at kritisk tenkning i *Overordnet del* av LK20 inkluderer at «etablerte ideer skal granskes,» og at «elever skal utvikle evne til å tenke kritisk omkring kunnskapsutvikling» (Røthing et al., 2024, s. 35). Forfatterne fremmer et maktkritisk perspektiv for skolen ved hjelp av begreper og teorier som normbevissthet, normkritisk pedagogikk, majoritetsprivilegier,

kritisk mangfoldskompetanse, interseksjonalitet, andregjøring, marginalisering og etisk bevissthet (Røthing et al., 2024, s. 36–41; se også Røthing & Engebretsen, 2019, s. 261). De bruker tillitsbegrepet kun tre ganger: to ganger om sviktende tillit til lærere fra det offentlige styringssystemet, og én gang som en demokratisk ferdighet. Danning brukes kun én gang, og da for å poengtere at norskfaget må favne bredere enn ren språkkompetanse, og også inkludere kompetanse knyttet til meningsmangfold (Røthing et al., 2024, s. 50).

Begrepene toleranse og respekt er derimot mer fremtredende og sentrale for deres visjon om en dialogbasert undervisning. Imidlertid problematiserer og radikaliserer antropomorfe aspekter ved KI-agenter den kilde- og maktkritikken som er nødvendig for demokratisk medborgerskap og dialog. Dialogmodellens ideal om en etisk rettferdig samtale med rom for ulike synspunkter må ta høyde for at kunstige stemmer og agenter i stadig større grad tar over den offentlige samtalen, men også private konversasjoner. To eksempler på det førstnevnte er at Ukraina i 2024 utnevnte en offisiell avatar, Victoria Shi, som talskvinne for utenriksdepartementet (Ministry of Foreign Affairs of Ukraine), og at Albania i 2025 introduserte KI-modellen Diella som minister i regjeringen for å motvirke korrupsjon i offentlige anskaffelser (Delauney, 2025).

Det er noe overraskende at begrepet kunstig intelligens kun brukes én gang i Røthings, Johannesens og Myrebøes maktkritiske bok om mangfold i klasserommet, og da kun i en oppramsing av mulige temaer for en dilemmafortelling i samfunnsfaget (Røthing et al., 2024, s. 103). På liknende måte som LK20 gjenspeiler dette trolig at man i den tidlige fasen etter 2022 ennå ikke hadde de mer radikale konsekvensene av den nye KI-teknologien for subjektivitet og autonomi som hovedfokus for refleksjon.

Kompetansepakker

Det er Udir som har det overordnede ansvaret for å utvikle støttemateriell og ressurser for lærere i Norge i deres arbeid med digital kompetanse, og dermed også kunstig intelligens (Udir, 2024a). Som følge av det spektakulære gjennombruddet for generativ KI har Udir opprettet en nettside som forklarer

hvordan KI-tematikk kan kobles til læreplanverket gjennom begreper som personvern, kildekritikk, opphavsrett, teknologi, algoritmisk tenkning, programmering og etikk (Udir, u.å.). Det spesifiseres relevante mål i læreplanverket for ulike fag: (engelsk) «digitale ressurser»; (kunst og håndverk) «digitale verktøy, ny teknologi»; (musikk) «teknologi, digitale verktøy, programmere»; (matematikk) «programmering, algoritmer»; (naturfag) «teknologi, programmering»; og (samfunnsfag) «teknologi som endringsfaktor og innvirkning» (Udir, u.å.).

Udir har også utviklet spesifikke nettkurs for lærere, såkalte kompetansepakker, som blant annet tar for seg digitalisering og KI (Udir, 2024b). I pakken «Teknologi og digitale ferdigheter i LK20» er imidlertid ikke KI et tema (Udir, 2025a). Det står ikke oppgitt noen dato for når denne pakken ble laget, men fraværet av KI tyder på at den var ferdig før 2022. Kompetansepakken «Systematisk arbeid med digital dømmekraft» har derimot KI som tema under paraplyene opphavsrett, personvern og kildekritikk (Udir, 2025b). Den mest interessante pakken for dette kapittelet er imidlertid «Kunstig intelligens i skolen» (Udir, 2025c). Det er denne som vil stå i fokus i det følgende.

Kompetansepakken om KI består av fire moduler (0–3) og to introduksjonsseksjoner. Den andre av introduksjonene heter «KIKK – Kunstig intelligens – Kritisk og konstruktiv» og omhandler et prosjekt (2025–2027) ved Høgskolen i Østfold som inkluderer fagene kunst og håndverk, musikk, norsk, fremmedspråk, matematikk, KRLE og yrkesfag. KIKK skal utvikle Udirs KI-kompetansepakke med tre nye moduler (modul 4–6): «Kritisk vurdering av KI», «Kritisk bruk i fag» og «Tverrfaglig arbeid med KI». Prosjektet har en «grunnleggende kritisk, men konstruktiv tilnærming til bruk av KI i og utenfor skolen» (Høgskolen i Østfold [HiØ], 2025). Målet med disse nye modulene er å «styrke læreres og elevers digitale dømmekraft» (HiØ, 2025).

KIKK-prosjektet er i høyeste grad relevant for å diskutere balansen mellom tillit og kritikk i den norske skolen, som her plasseres i spenningsfeltet mellom muligheter og kritikk. I KIKK-prosjektet forstås KI dermed først og fremst som et sett med verktøy, sett i lys av en økt forståelse for risikoene knyttet til digitaliseringen. Fokuset ligger på kritikk og dømmekraft, og i mindre grad på den endringen av sosial kommunikasjon som simulert antropomorfisme gir opphav til.

I de nåværende modulene tematiseres KI som en årsak til at demokratiske verdier er satt under press, blant annet fordi det har blitt vanskelig å skille

mellom ekte og falsk informasjon, noe som gir næring til konspirasjonsteorier (Udir, 2025c, § 1.2). Fra et mer konstruktivt perspektiv viser pakken til KI som en «kognitiv avlastning», da fremfor alt når det gjelder tekst (Udir, 2025c, § 2.5). Det understrekes imidlertid at dette ikke skal innebære en avlastning når det gjelder selve læringen. Det vil si at KI skal fungere som et hjelpemiddel, men den skal ikke løse problemer autonomt.

I de planlagte, nye modulene er fokuset på støttemateriell for enkeltfag, men også på tverrfaglige temaer, deriblant vurdering. Modul 4 skal gi «innsikt i de etiske, juridiske og faglige utfordringene ved KI-bruk» (HiØ, 2025). Den skal gi en oppdatering på KI-utviklingen og løfte motforestillinger, som for eksempel at «KI er problematisk med tanke på personvern». Modul 5, «Kritisk bruk i fag», skal gi en tilsvarende oppdatering på KI-status i 2026, og tematisere KI som assistent og sparringspartner. Modul 6 skal ta for seg tverrfaglige prosjekter med hovedfokus på kritikk, men med en åpenhet for mulighetene teknologien gir.

PfDK

Kunstig intelligens har blitt et akutt tema for lærerens profesjonsfaglige digitale kompetanse (PfDK). I det norske rammeverket for PfDK fra 2024 er kunstig intelligens derfor inkludert i ett av tre kunnskapsmål: Læreren skal «forstå hvordan kunstig intelligens kan endre måten vi forholder oss til fagkunnskap». Innenfor etikk skal læreren kunne «reflektere over etiske og juridiske problemstillinger ved bruk av læringsanalyse og kunstig intelligens», og når det gjelder læremidler, skal vedkommende kunne «kritisk vurdere og dra nytte av kunstig intelligens, læringsanalyse og adaptive læremidler». Innenfor ledelse skal læreren forstå «hvordan kunstig intelligens utfordrer eksisterende [sic] og skaper nye muligheter for læring», og under temaet *endring og utvikling* angis det at læreren skal forstå «hvordan kunstig intelligens utfordrer lærerens og skolens rolle» (Udir, 2024c, s. 13).

Dette innebærer at målene for lærerens KI-kompetanse er eksplisitt formulert, noe som ikke er tilfelle i læreplanene fra 2020, som definerer elevenes lærings- og kompetansemål. Det er også signifikant at det er kritikk og

utfordringer som står i fokus, og ikke muligheter. KI introduseres med oppfordringen om at «barn og unge bør ha en kritisk tilnærming», og algoritmers påvirkning på demokratiske prosesser nevnes eksplisitt (Udir, 2024c, s. 3). I ordforklaringen til slutt står det at «implementeringen av kunstig intelligens i skolen krever nøye planlegging og overvåking», og personvern og etiske aspekter nevnes spesifikt (Udir, 2024c, s. 16).

I PfDK-rammeverket er ikke tillit nevnt eksplisitt, og heller ikke respekt, anerkjennelse eller toleranse. I stedet er det kunnskapsbasert og etisk kritikk som dominerer. Det simulerte menneskelige, og de spesielle utfordringene dette medfører for forståelsen av mennesket og sosialt liv, er heller ikke direkte tematisert. Dette blir tydelig i rammeverkets definisjon av digital dannelse som «en prosess der et menneske former sin identitet i en digital kontekst» (Udir, 2024c, s. 14). Det digitale er her en kontekst – en omgivelse – der man former sin handlingskompetanse (sosial, kulturell, praktisk og etisk kompetanse nevnes). Rammeverket løfter frem et ideal om et modent menneske som, basert på en kombinasjon av kritisk refleksjon, rasjonalitet og autonomi, fatter «selvstendige beslutninger» (Udir, 2024c, s. 14). Et slikt moderne subjekt utfordres av KI-teknologien gjennom dens enorme kunnskapsgrunnlag, evne til resonnement og analyse og økende selvstendighet. I stor grad er konstruksjonen av KI-agenter en simulering av nettopp et moderne, rasjonelt, kritisk og autonomt subjekt. I første omgang innebærer dette, som nevnt, at det er nesten umulig å vite om kommunikasjon kommer fra et simulert eller et menneskelig subjekt, eller en kombinasjon av begge. Den simulerte rasjonaliteten og kreativiteten underminerer dermed tradisjonelle tillitsbaserte relasjoner og setter spørsmålet om det menneskelige i søkelyset.

Disse aspektene er på nåværende tidspunkt ikke tydelig tematisert i PfDK-rammeverket, antakelig fordi det krever et fyldigere ideal for mennesket enn kun å være et kritisk, rasjonelt og selvstendig subjekt. I denne sammenhengen er tillitsbaserte mellommenneskelige relasjoner, kultur og tradisjoner sentrale. Simulerte intelligente agenter trenger også å dannes i tråd med en slik bredere forståelse av etikk; en forståelse som går utover rene spørsmål om personvern og opphavsrett. Dermed blir den digitale konteksten noe mer enn bare en bakgrunn for menneskelig dannelse, ettersom også simulerte subjekter må dannes etisk og kulturelt.

En bredere forståelse av digital danning

Et perspektiv på menneskelig sosialt liv som tar spørsmålet om KI-autonomi på alvor, innbyr til en samtale om systemer og situasjoner der KI-agenter bruker mennesker som verktøy. Man kan diskutere om dette er etisk forsvarlig, men den pågående transformasjonen av samfunnet er vanskelig å tematisere hvis verktøy- eller kontekstmetaforen blir for dominerende. For eksempel omtales data ofte som det nye gullet. De store KI-selskapene trenger desperat menneskelige data for å trene sine modeller. «Gratis» sosiale medier og andre digitale tjenester betales ved at brukerne gir bort sine data til selskapene. KI-systemer tilpasset sosiale medier er derfor optimalisert for engasjement, og de bruker mennesker som verktøy for å høste data – noe som for enkelte kan føre til en destruktiv avhengighet, i likhet med det man ser hos spilleavhengige.

Samtidig kan en KI-assistent som hjelper elever med å lære, ha barnets kunnskap, kompetanse og danning som mål. Den kan dermed bruke en menneskelig lærer for å oppnå dette målet, ved å tilkalle læreren og fortelle om elevens arbeid. Hvis den har tilgang til en konto med penger, kan den også betale privatlærere med spesialkompetanse for å hjelpe barnet. Trolig vil slike autonome KI-lærere ikke bli tillatt i norsk skole, men vi må forestille oss at ressurssterke foreldre som vil optimalisere barnas sjanser i livet, vil betale for en slik pedagogisk veileder og gi den et romslig budsjett. KI-læreren kan dermed benytte seg av både mennesker og andre KI-agenter for at barnet skal utvikle seg og lære ut fra sine egne forutsetninger. For eksempel kan KI-læreren finne de beste menneskelige språklærerne for barnet, overvåke arbeidet deres og avskjedige dem når resultatene ikke er tilfredsstillende. På samme måte vil mange gi en KI-agent et månedsbudsjett og kun ett mål: «Gjør meg rik». Slike kunstige investorer vil bruke både mennesker og maskiner for å oppnå sine mål.

Hva gjør dette med tillitsrelasjonene i samfunnet og i skolen? Hvis tillit kun handler om at man stoler på at et menneske eller en maskin vil gjøre ting som er til ens egen fordel, ser man utelukkende på tillit som noe relatert til funksjon. Tilliten har imidlertid en dypere dimensjon som er avgjørende for læring, nemlig vennskap og kjærlighet. Å ha tillit til en venn innebærer ikke bare å forvente at vedkommende vil gjøre ting som er til nytte for en selv. Vennskap er å se verdien i den andre, helt uavhengig av funksjon. Tilliten peker dermed mot kjærlighet, som den mellom en forelder og et barn: en

type vennskap som ytterst sett viser seg ved at man er beredt til å ofre sitt eget liv for den andre.

Den funksjonelle tilliten kan uten større problemer rettes mot autonome KI-agenter, men den må kombineres med et utviklet kritisk sinn – en bevissthet om at man selv kan fungere som et middel for den kunstige intelligensens måloppnåelse. Den tilliten som er basert på vennskap og kjærlighet, er vanskeligere å simulere. Det vil likevel bli stadig vanligere med vennskapsrelasjoner mellom mennesker og kunstige agenter, og til og med kjærlighetsrelasjoner fra menneskets side. Spørsmålet om forskjellen mellom ekte kjærlighet og simulert kjærlighet, eller kun funksjonell nytelse – som pornografi – setter søkelys på at også mange former for menneskelig vennskap og kjærlighet først og fremst er funksjonelt motivert. I så fall kan KI-agenter uten større problemer overta disse rollene, noe som vil bli ytterligere aksentuert når avanserte roboter integrert med bioteknologi introduseres.

Et fokus på tillit i utviklingen av lærerens profesjonsfaglige digitale kompetanse åpner dermed for temaer som spenningen mellom funksjon og egenverdi, samt vennskap og kjærlighet. På den måten kan lærere få kompetanse til å føre samtaler som ikke bare dreier seg om kildekritiske spørsmål, maktkritikk og etisk dømmekraft, men også om filosofiske og religiøse grunnlagsspørsmål. Et transhumanistisk syn på mennesket som et åpent utviklingsprosjekt radikaliserer en danning basert på rasjonalitet, frihet og selvstendighet. Det utvisker samtidig grensen mellom kunstige og biologiske agenter. Å etablere skoler som lavteknologiske soner er dermed ingen løsning på den fundamentale utfordringen simulert rasjonalitet og subjektivitet gir opphav til. Det krever snarere at man spør seg i hvilket omfang en rikere forståelse av det menneskelige må tematiseres og gjøres operativ i den digitale kompetansen og den «digitale» danningen.

Konklusjon

Dette kapittelet viser at dokumentene som setter rammeverket for norske læreres håndtering av KI-teknologi, også bør drøfte hvordan en verden fylt av kunstige agenter forandrer de tillitsrelasjonene som er grunnleggende for

læring. Dette innebærer i så fall en forskyvning fra et ensidig fokus på kritikk og KI forstått som et verktøy, til at oppmerksomheten rettes mot konsekvensene av at intelligente agenter med sterke antropomorfe trekk – deriblant høy emosjonell intelligens – inntar sosiale roller og utfører handlinger.

Et verktøyforhold mellom menneske og maskin er instrumentelt og har først og fremst en funksjonell verdi. Elever eller voksne samfunnsborgere bruker da KI for å oppnå mål som mennesker selv har formulert. Men hvis barn dannes gjennom daglig kontakt med ulike former for KI som opptrer som simulerte subjekter, blir forskjellen mellom tillit til organiske og kunstige intelligenser utydelig. Dersom PfDK-rammeverket og støttemateriellet fra Udir fortsetter å prioritere økt kritikk, vil dette dermed ikke bare påvirke relasjonen til kunstige, antropomorfe agenter, men også til virkelige mennesker.

Referanser

Bøhn, E. D. (2024). Kan generativ kunstig intelligens opplyse og danne oss? *Nordisk tidsskrift for pedagogikk & kritikk*, *10*(1). https://doi.org/10.23865/ntpk.v10.6175

Corjescu, D. & Doat, D. (2022). *Prometheus unbound: the perils and promise of transhumanism*. Ibidem.

Delauney, G. (2025, september 12). *World's first AI minister will eliminate corruption, says Albania's PM*. BBC. https://www.bbc.com/news/articles/cm2znzgwj3xo

Dillon, R. S. (2022). Respect. I E. N. Zalta & U. Nodelman (Red.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2022/entries/respect

Dyngeland, E. & Engelsrud, G. H. (2024). Embodied Bildung: a phenomenological inquiry into an embodied understanding of emancipation and autonomy in education. *Scandinavian Journal of Educational Research*, 1–12. https://doi.org/10.1080/00313831.2024.2434823

Elstad, E. & Eriksen, H. (2025). Eleven må finne sin egen skrivestemme: med eller uten kunstig intelligens. I E. Elstad & K. Korbøl (Red.), *Kunstig intelligens i skolen* (s. 73–100). Fagbokforlaget.

Fisher, A. & Tallant, J. (2020). Trust in education. *Educational Philosophy and Theory*, *52*(7), 780–790. https://doi.org/10.1080/00131857.2019.1692304

Grierson, E. M. (2018). Trust and fiduciary relationships in education: What happens when trust is breached? *Educational Philosophy and Theory*, *50*(2), 203–211. https://doi.org/10.1080/00131857.2016.1231052

Grimen, H. (2009). *Hva er tillit*. Universitetsforlaget.

Helmerson, K. (2025, september 9). Mobiler förbjuds i skolan – även på raster och fritids. *Sveriges Radio*. https://www.sverigesradio.se/artikel/da-infors-stopp-for-mobiler-i-skolan-aven-pa-rasten

Hofmann, H. (2025). Biases in AI: acknowledging and addressing the inevitable ethical issues. *Frontiers in Digital Health*, *7: 1614105*. https://doi.org/10.3389/fdgth.2025.1614105

Høgskolen i Østfold. (2025). *Kunstig intelligens – kritisk og konstruktiv (KIKK).* Sist endret 26. aug 2025. https://www.hiof.no/fss/om/aktuelt/arrangementer/2025/kikk-sept-2025.html

Lie, T. G. (2022). Lærerens forståelse av og arbeid med digitale kilder og kildekritikk i samfunnsfag. I M. A. Letnes & F. M. Røkenes (Red.), *Digital teknologi for læring og undervisning i skolen* (s. 294–312). Universitetsforlaget.

Lund, A. B. (2017). Mangfold i skolen. En ressurs? I A. B. Lund (Red.), *Mangfold gjennom anerkjennelse og inkludering i skolen* (s. 28–47). Gyldendal.

Ministry of Foreign Affairs of Ukraine. (2023, mai 1). *Welcome speech of the representative of the Ministry of Foreign Affairs of Consular Affairs Victoria Shi*. https://mfa.gov.ua/en/news/vitalne-slovo-predstavnici-mzs-z-konsulskih-pitan-viktoriyi-shi

Mitchell, R. M., Kensler, L. & Tschannen-Moran, M. (2018). Student trust in teachers and student perceptions of safety: positive predictors of student identification with school. *International Journal of Leadership in Education*, *21*(2), 135–154. https://doi.org/10.1080/13603124.2016.1157211

Opplæringslova. (2023). *Lov om grunnskoleopplæringa og den vidaregåande opplæringa (LOV-2023-06-09-30)*. Lovdata. https://lovdata.no/dokument/NL/lov/2023-06-09-30

Ortiz-Ospina, E., Roser, M. & Arriagada, P. (2024). Trust. *Our World In Data*. https://ourworldindata.org/trust

Palmer, S. W. (2024). *The role of trust in facilitating human-artificial intelligence relationships* [Doktoravhandling, Aarhus universitet]. Pure.au.dk.

Peña-Garceda, P., Jaime-de-Aza, M. & Feltrero, R. (2026). Dialogic reflection and algorithmic bias: pathways toward inclusive AI in education. *Trends in Higher Education*, *5*(1), Article 9. https://doi.org/10.3390/higheredu5010009

Pitts, G. & Motamedi, S. (2025). Understanding human-AI trust in education. *arXiv:2506.09160 [cs.CY]*. https://doi.org/10.48550/arxiv.2506.09160

Qin, F., Li, K. & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology*, *51*(5), 1693–1710. https://doi.org/10.1111/bjet.12994

Regjeringen. (2025, april 4). *Vurderer å fjerne digitale ferdigheter som grunnleggende ferdighet*. https://www.regjeringen.no/no/aktuelt/vurderer-a-fjerne-digitale-ferdigheter-som-grunnleggende-ferdighet/id3095564/

Røthing, Å. & Engebretsen, E. L. (2019). Maktkritiske perspektiver i høyere utdanning? En undersøkelse av omfang og anvendelse i utvalgte programplaner. *Uniped*, *42*(3), 251–261. https://doi.org/10.18261/issn.1893-8981-2019-03-02

Røthing, Å., Johannesen, E. M. & Myrebøe, T. (2024). *Meningsmangfold i klasserommet. Perspektiver, begreper og ressurser for lærere*. Cappelen Damm.

Schoorman, F. D., Mayer, C. R. & Davis, J. H. (2007). An organizational model of organizational trust: past, present, and future. *Academy of Management Review*, *32*(2), 344–354. https://doi.org/10.5465/amr.2007.24348410

Scott, R. (1982). *Blade Runner* [Videoopptak]. Warner Bros.

Sjöström, J. & Eilks, I. (2020). The *Bildung* theory–From von Humboldt to Klafki and beyond. I B. Akpan & T. J. Kennedy (Red.), *Science education in theory and practice*, 55–67. Springer. https://doi.org/10.1007/978-3-030-43620-9_5

Slåttholm Sagdahl, M. (2025). Toleranse (overbærenhet). I *Store norske leksikon*. https://snl.no/toleranse_-_overb%C3%A6renhet

Spielberg, S. (2001). *A.I. Artificial Intelligence* [Videoopptak]. Warner Bros.; DreamWorks; Amblin Entertainment; Stanley Kubrick Productions.

Stovner, R. B., Brevik, L. M., Mørk, S. M. & Færøyvik, E. (2025). *Å arbeide algoritmisk i matematikk: En studie av algoritmisk tenkning og programmering på 8. trinn, 10. trinn og Vg1*. (EDUCATE-rapport nr. 5). Universitetet i Oslo. https://www.uv.uio.no/ils/forskning/prosjekter/educate/rapporter/stovner-et-al-(2025)-educate-rapport-5.pdf

Sørreime, Ø. (2020). *Tillit og styring i skolen*. Cappelen Damm Akademisk.

Tschannen-Moran, M. (2017). Trust in education. I *Oxford Research Encyclopedia of Education*. https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-79

Tække, J. & Paulsen, M. (2022). *A new perspective on education in the digital age: Teaching, media and Bildung*. Bloomsbury Academic. https://doi.org/10.5040/9781350175426

Utbildningsdepartementet. (2025). *Ett nationellt mobilförbud i de obligatoriska skolformerna och fritidshemmet (Ds 2025:9)*. Regeringskansliet.

Utdanningsdirektoratet. (u.å.). *Læreplanen og KI.* Hentet 12. august 2025, fra https://www.udir.no/kvalitet-og-kompetanse/digitalisering-skole/kunstig-intelligens-ki-i-skolen/ki-i-lareplanen/kunnskap-om-kunstig-intelligens-som-teknologi/#eksempler-pa-aktuelle-fag-og-innhold-i-lareplanene

Utdanningsdirektoratet. (2017, november 15). *Rammeverk for grunnleggende ferdigheter.* https://www.udir.no/laring-og-trivsel/rammeverk/rammeverk-for-grunnleggende-ferdigheter/

Utdanningsdirektoratet. (2020). *Læreplan i matematikk (MAT01-05).* https://www.udir.no/lk20/mat01-05/kompetansemaal-og-vurdering/kv16?lang=nob

Utdanningsdirektoratet. (2024a, august 11). *Råd om kunstig intelligens i skolen.* https://www.udir.no/kvalitet-og-kompetanse/digitalisering-skole/kunstig-intelligens-ki-i-skolen/kunstig-intelligens-ki-i-skolen/

Utdanningsdirektoratet. (2024b, november 8). *Kompetansepakker for digital kompetanse i skolen.* https://www.udir.no/kvalitet-og-kompetanse/digitalisering-skole/kompetansepakker-for-digital-kompetanse-i-skolen/

Utdanningsdirektoratet. (2024c, november 11). *Rammeverk for lærerens profesjonsfaglige digitale kompetanse.* https://www.udir.no/kvalitet-og-kompetanse/digitalisering-skole/rammeverk-larerens-profesjonsfaglige-digitale-komp/

Utdanningsdirektoratet. (2025a). *Teknologi og digitale ferdigheter i LK20.* https://bibsys.instructure.com/courses/650

Utdanningsdirektoratet. (2025b). *Systematisk arbeid med digital dømmekraft.* https://bibsys.instructure.com/courses/848

Utdanningsdirektoratet. (2025c). *Kunstig intelligens i skolen.* https://bibsys.instructure.com/courses/637

Utdanningsdirektoratet. (2026). *Overordnet del av læreplanverket – verdier og prinsipper for grunnopplæringen (LK20).* https://www.udir.no/lk20/overordnet-del/