



LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

Stora språkmodeller som resurser för lärare: möjligheter och utmaningar inom högre utbildning

En studie om lärares användning av stora språkmodeller

Kandidatuppsats 15 hp, kurs SYSK16 i Informatik

Författare: Malin Tennander
Hanna Håkansson

Handledare: Nicklas Holmberg
Markus Lahtinen

Rättande lärare: Benjamin Weaver
Niki Chatzipanagiotou

Stora språkmodeller som resurser för lärare: möjligheter och utmaningar inom högre utbildning

ENGELSK TITEL: Large Language Models as Resources for Educators: Opportunities and Challenges in Higher Education

FÖRFATTARE: Hanna Håkansson och Malin Tennander

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Osama Mansour, Docent

FRAMLAGD: Maj, 2024

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 135

NYCKELORD: Large Language Models, stora språkmodeller, digital kompetens, prompt engineering, lärare, högre utbildning, universitet

SAMMANFATTNING:

Denna kandidatuppsats undersöker hur universitetslärare använder stora språkmodeller (LLM) i sitt arbete. Genom en kombination av litteraturstudier och kvalitativa intervjuer med lärare vid svenska universitet utforskar studien hur dessa teknologier integreras i akademiska arbetsuppgifter. Resultaten visar att användningen varierar stort och inkluderar allt från undervisningsstöd till administrativa uppgifter. Vidare är användningen också starkt influerad av individuella behov snarare än institutionella direktiv. Studien identifierar även en rad tekniska och etiska begränsningar, som dataskyddsfrågor och risken för missvisande information. Trots dessa hinder framhäver lärarna potentialen hos LLM att främja pedagogisk innovation och effektivisera arbetsprocesser. Dock uppmärksammas ett behov av ytterligare stöd och riktlinjer från universiteten för att hantera teknologin säkert och etiskt. Studien bidrar med insikter om den komplexa relationen mellan teknologi och pedagogik och understryker behovet av att universiteten skapar stödjande strukturer som kan hjälpa lärare att navigera i de tekniska och etiska utmaningar som stora språkmodeller medför.

Innehåll

Figurer	3
Tabeller	3
1 Introduktion	1
1.1 Bakgrund.....	1
1.2 Problemområde och forskningsfråga.....	2
1.3 Syfte.....	3
1.4 Avgränsningar.....	3
1.5 Definitioner.....	3
2 Litteraturgenomgång	4
2.1 Stora språkmodeller.....	4
2.1.1 Arkitektur.....	4
2.1.2 Tekniska begränsningar och etiska dilemman.....	5
2.2 Prompt engineering.....	6
2.2.1 Prompt engineering för effektiv användning av stora språkmodeller.....	6
2.2.2 Prompting-metoder.....	6
2.3 Digital kompetens.....	8
2.3.1 Digital kompetens hos lärare.....	9
2.3.2 Prompt engineering som digital kompetens.....	10
2.4 Stora språkmodellens potential inom utbildning.....	11
2.4.1 Nuvarande kunskapsläge och relaterad forskning.....	11
2.4.2 Lärares roll vad gäller AI inom utbildning.....	13
2.4.3 Universitetens roll.....	13
2.5 Teoretiskt resultat.....	13
3 Metod	16
3.1 Val av metod.....	16
3.2 Urval av respondenter.....	17
3.3 Datainsamling.....	18
3.3.1 Litteraturinsamling.....	18
3.3.2 Utformning och genomförande av intervjuer.....	18
3.3.3 Intervjuguide.....	20
3.4 Bearbetning av data.....	20
3.4.1 Förberedelse av data.....	20
3.4.2 Kodning och tematisering.....	21
3.5 Validitet och reliabilitet.....	22
3.6 Etiska aspekter.....	23
4 Empiriska resultat	24

4.1 Användningen av stora språkmodeller.....	24
4.1.1 Verktyg och användningsområden.....	24
4.1.2 Upplevda förbättringar.....	25
4.1.3 Upplevda begränsningar.....	27
4.2 Prompt engineering.....	29
4.3 Digital kompetens.....	31
4.3.1 Lärares syn på digital kompetens.....	31
4.3.2 Utbildning.....	32
4.3.3 Kritisk förhållning.....	33
4.4 Stora språkmodellens potential inom utbildning.....	35
5 Diskussion.....	38
5.1 Användningen av stora språkmodeller.....	38
5.1.1 Användningsområden.....	38
5.1.2 Upplevda förbättringar.....	38
5.1.3 Upplevda begränsningar.....	40
5.2 Prompt engineering.....	41
5.3 Digital kompetens.....	42
5.3.1 Lärares syn på digital kompetens.....	42
5.3.2 Utbildning.....	43
5.3.3 Kritisk förhållning.....	44
5.3.4 Applicering av DigComp-ramverket.....	45
5.4 Stora språkmodellens potential inom utbildning.....	46
6 Slutsats.....	48
6.1 Förslag på vidare forskning.....	49
Bilagor.....	50
Bilaga 8 - Samtyckesformulär.....	50
Bilaga 9 - AI-bidragsredogörelse.....	52
Referenser.....	53

Figurer

Figur 2:1 - The DigComp conceptual reference model.....	9
---	---

Tabeller

Tabell 2:1 - Rekommendationer för text-till-text prompting.....	7
Tabell 2:2 - AI-prompting som en digital kompetens enligt DigComp 2.2.....	11
Tabell 2:3 - Teoretisk resultatmodell.....	15
Tabell 3:1 - Respondentöversikt.....	17
Tabell 3:2 - Intervjuguide.....	20
Tabell 3:3 - Kodningsschema.....	21
Tabell 4:1 - Kartläggning av användningsområden.....	24
Tabell 4:2 - Kartläggning av upplevda förbättringar.....	26
Tabell 5:1 - Applicering av DigComp-ramverket på stora språkmodeller.....	45

1 Introduktion

1.1 Bakgrund

År 2016 introducerade Klaus Schwab, grundare och styrelseordförande för World Economic Forum, begreppet *den fjärde industriella revolutionen* (Schwab, 2016). I hans bok “The Fourth Industrial Revolution” identifieras de främsta drivkrafterna bakom revolutionen som tekniska genombrott inom områden som artificiell intelligens (AI), Internet of Things (IoT), 3D-utskrifter, nanoteknik och bioteknik. Schwab (2016) framhåller att den fjärde revolutionen utmärker sig genom den hastighet, omfattning och kraft med vilka förändringar drivs, där teknologisk utveckling inte längre sker linjärt utan snarare exponentiellt. Dessa framsteg beskrivs ha en påverkan på nästan alla industrier och samhällsskikt och omforma hur vi lever och arbetar, vilket ställer krav på anpassningar på individer såväl som samhället i stort.

I slutet av 2022 lanserades ChatGPT-3.5, en så kallad Large Language Model (LLM), på svenska benämnd stor språkmodell, som behärskar att både tolka och generera naturligt språk och således är förmögen att simulera mänskliga konversationer (OpenAI, 2022). Porter (2023) beskriver att verktyget snabbt blev en global sensation och att användandet av ChatGPT ökade explosionsartat under loppet av bara ett år. Att ChatGPT har genererat så stor uppmärksamhet förklarar Tate et al. (2023) med att ChatGPT var unik i det avseende att verktyget kostnadsfritt tillgängliggjordes för bruk till allmänheten. Med ett lättillgängligt gränssnitt som inte kräver någon form av teknisk expertis erbjuder ChatGPT en omedelbar och bred användbarhet som är möjlig för vem som helst att nyttja. OpenAI har således tillgängliggjort kraftfull AI-teknik för en bredare publik vilket ger potential för en mångfald av yrkesgrupper att kunna utnyttja verktyget och tillämpa tekniken över ett brett spektrum av arbetsfält.

Sedan ChatGPT lanserades har flera liknande stora språkmodeller utvecklats, där ibland OpenAI:s ChatGPT-4, Googles Gemini och Microsofts Copilot (OpenAI, 2023; Pichai & Hassabis, 2023; Spataro, 2023). Detta har bidragit till att språkmodeller har blivit ett allt mer framträdande begrepp och avsevärt ökat allmänhetens uppmärksamhet för tekniken. Flera författare ser potentialen för att integrera denna typ av AI-verktyg i olika arbetsmiljöer kan medföra positiv förändring mot mer effektiva tillika produktiva yrkesmetoder och på så vis underlätta ett bredare omfång av kreativt och analytiskt arbete (Alshurafat, 2023; Liu et al., 2023). David Rotman (2023) belyser hur ChatGPT och liknande modeller snabbt blivit en integrerad del av många branscher och skapat en intensifierad konkurrens där företag nu tävlar om att finna de mest innovativa och effektiva sätten att utnyttja verktygen för att revolutionera arbetsprocesser och skapa nya affärsmöjligheter. Rotman skildrar fenomenet som en slags modern “artificiell intelligens-guldrush”, där jakten på att nyttja AI:s potential skapar en ny era av innovation och ekonomiska möjligheter.

Uppståndelsen kring generativa språkmodeller har inte minst påverkat utbildningssektorn, där det gett upphov till en intensiv debatt kring AI och dess eventuella plats inom utbildning. I en analys från 2023 fastställer Sullivan et al. (2023) att nyhetsrapporteringar om ChatGPT:s påverkan på högre utbildning till övervägande del har varit fokuserat kring negativa aspekter.

En oro har bland annat varit elevers möjligheter att kringgå lärande som ett resultat av att använda ChatGPT (Brodie, 2023). Det finns dock de som är mer positivt inställda och argumenterar för att språkmodellerna öppnar upp för nya innovativa undervisnings- och inlärningsmetoder. Argumentet stöds av flertalet forskningsresultat som visar att användandet av denna typ av chatbotar i inlärningsssammanhang har potential att förbättra pedagogiska upplevelser och inläring (Kuhail et al., 2022; Chen et al., 2022). Topprankade svenska universitet såsom Lunds universitet, Uppsala universitet och Kungliga Tekniska Högskolan (KTH) tillhör några av de som tagit fasta på den nya tekniken och inom viss mån inte bara tillåter utan därtill även uppmuntrar eleverna att nyttja denna typ av AI som ett pedagogiskt stöd, förutsatt att vissa försiktighetsåtgärder vidhålls (Esaiasson, 2023; Stålberg, 2023; van der Wijngaart, 2023).

Debatten om AI inom utbildning har varit mycket fokuserad på elevers användning av verktygen, men det finns även lärare som har visat intresse för att nyttja dessa verktyg. Thomas Edwall, lärare och författare av boken "ChatGPT för lärare" är en av de som ser positivt på och därtill fullt anammat användandet av ChatGPT i skolmiljön (Wasell, 2023). Edwall anser att fördelarna sträcker sig bortom elevernas lärande i klassrummet till att även omfatta hans eget arbete och välbefinnande. Han använder ChatGPT dagligen som både som ett pedagogiskt verktyg, vilket han hävdar har höjt kvaliteten på hans undervisning, men även för eget bruk till mer administrativa ändamål. Han liknar ChatGPT med en ny digital lärarassistent som kraftigt minskat hans stressnivå som lärare. Edwall anser att allmänheten är ovetandes om vilka möjligheter det finns för lärare att effektivisera det egna arbetet med hjälp av ChatGPT och vill lyfta en diskussion kring hur lärare bättre kan nyttja denna typ av teknik för att underlätta det egna arbetet och höja kvaliteten på undervisningen.

1.2 Problemområde och forskningsfråga

I takt med den växande debatten kring användningen av stora språkmodeller har även forskningen inom området tilltagit. Sammantaget finns en tendens bland befintlig språkmodellforskning att fokusera på studenters användning av verktygen. I samband med detta har flera fördelar identifierats för stora språkmodellens potential att effektivisera lärandet, såsom att nyttja språkmodellerna för språkstöd eller att ge förklaringar till specifika frågor, men även som stöd för att hitta källor, material och data (Dwivedi et al., 2023; Kasneci et al., 2023; Baidoo-Anu & Ansah, 2023). Det framhävs emellertid även vissa risker med studenters användning av stora språkmodeller. Dwivedi et al. (2023) framhäver risker såsom minskad motivation till eget skrivande och självständigt tänkande och Neumann et al. (2023) bygger därtill på med risken för oetisk användning i form av exempelvis fusk. Befintlig forsknings tendens att fokusera på studenters användning ger upphov till en nyfikenhet för att undersöka hur lärare använder sig av dessa verktyg i sitt arbete.

Studier som specifikt fokuserar på lärares användning av stora språkmodeller har huvudsakligen studerats inom lägre utbildningsnivåer, där användningsområden såsom att avlasta arbetsbörda, stödja innehållsgenerering eller att nyttja språkmodeller som en lärarassistent har identifierats (Hashem et al., 2024, Jeon & Lee, 2023). Vad gäller lärares användning inom specifikt högre utbildning har fördelar som pedagogisk innovation och utveckling av läroplaner bevitnas (Neumann et al., 2023). Cerratto et al. (2024) framhäver dock vissa utmaningar, som att stora språkmodeller inte är utvecklade för att linjera med akademiska värderingar vilket utgör ett problem för etisk användning. Van Wyk (2024) argumenterar för att fler studier krävs inom området för hur akademiker inom just högre

utbildning ser på stora språkmodeller som ett verktyg i deras yrkesroll vilket denna studie har tagit fasta på. Befintlig litteratur tenderar att huvudsakligen fokusera på användning kopplad till undervisningsmoment, vilket denna uppsats delvis kommer att undersöka. Därtill finns även ett intresse för att undersöka hur dessa teknologier integreras i ett bredare spektrum av arbetsuppgifter, vilket delvis sträcker sig bortom lärarnas pedagogiska arbete.

Slutligen har majoriteten av dessa forskningsbidrag genomförts utanför Sverige, vilket väckt intresset för att utforska ämnet ur en svensk kontext. Den sammantagna bilden av det nuvarande kunskapsläget understryker relevansen av att undersöka användningen av stora språkmodeller bland lärare vid högre utbildningsinstitutioner i Sverige. Ett behov av att undersöka detta fenomen på en mer lokal nivå har identifierats, något som skulle bidra till att fylla en kunskapslucka i den lokala förståelsen för hur stora språkmodeller används inom akademien. Det är viktigt att förstå inte bara hur universitetslärare kan använda dessa teknologier i sitt arbete, utan även vilka utmaningar som kan uppstå i samband med en sådan integrering samt vilka attityder som finns bland lärarna kring deras integration.

Mot bakgrund av detta ämnar studien att besvara följande forskningsfråga:

- *Hur använder universitetslärare stora språkmodeller i sitt professionella arbete och vilka attityder finns kopplade till denna teknologi?*

1.3 Syfte

Studiens syfte är att utifrån enskilda universitetslärares faktiska användning av och attityder till stora språkmodeller undersöka möjligheterna för en mer enhetlig integrering av dessa teknologier inom läraryrket.

1.4 Avgränsningar

Denna studie har avgränsats till att endast undersöka användningen av stora språkmodeller bland universitetslärare i Sverige. Dessutom begränsas de språkmodeller som behandlas i uppsatsen till de verktyg som respondenterna aktivt nämner, med motiveringen att det inte tjänar ett syfte att analysera användningen av verktyg som inte används.

1.5 Definitioner

I uppsatsen används begreppet *lärare* genomgående, benämningen används för att innefatta individer anställda vid en institution för högre utbildning vars arbetsuppgifter till någon del omfattar undervisning.

Inom ämnet *prompt engineering* kommer termerna *prompt engineering* och *systematisk prompting* att syfta till samma koncept i denna uppsats. Vidare ger detta upphov till verbet *att prompta* samt substantiven *prompt* och *prompting*. Dessa ord kommer att användas och behandlas enligt det svenska språkets grammatiska regler.

2 Litteraturgenomgång

I följande kapitel presenteras den litteratur och de relevanta ramverk som undersökningen tar stöd ifrån. Inledningsvis beskrivs stora språkmodeller tillsammans med relaterade begrepp med syfte att lägga en solid grund för att förstå hur tekniken är uppbyggt, både vad gäller möjligheter för verktygen likväl som dess svagheter och medföljande risker. Därefter beskrivs digital kompetens i syfte att sätta kontexten för litteraturen, med specifikt fokus på lärare. Vidare redogörs det för det rådande kunskapsläget vad gäller forskning om användningen av stora språkmodeller i en yrkeskontext för lärare. Avslutningsvis presenteras det teoretiska resultatet.

2.1 Stora språkmodeller

2.1.1 Arkitektur

Stora språkmodeller, även kända som Large Language Models på engelska, är en form av generativ AI som har tränats på omfattande datamängder. Denna träning möjliggör för modellerna att hantera ett brett spektrum av uppgifter, från grundläggande språköversättningar till mer avancerade funktioner som textgenerering, sammanfattningar och dataanalys. Stora språkmodeller används även som drivkraft i olika chattbotar (Khurana et al., 2022; Rathore, 2023; Naveed et al., 2024).

Khurana et al. (2022) beskriver att stora språkmodeller bygger på så kallad Natural Language Processing (NLP), en form av lingvistisk AI-forskning som möjliggör för dator och människa att kommunicera med varandra genom användandet av naturligt (mänskligt) språk. Författarna delar in NLP-tekniken i två kategorier - Natural Language Understanding (NLU) och Natural Language Generation (NLG). NLU omfattar datorns förmåga att förstå och tolka inmatad text, medan NLG istället omfattar datorns förmåga att själv kunna generera text som är meningsfull och sammanhängande. Tillsammans beskrivs dessa tekniker möjliggöra för stora språkmodeller att simulera personliga, människoliknande konversationer som är sammanhängande och relevanta för kontexten. Khurana et al. (2022) attribuerar denna förmåga till den omfattande korpus av lingvistiska data som modellerna har tränats på, vilken är sammansatt av text och transkriberat tal.

År 2017 presenterade Vaswani et al. (2017) den så kallade transformermodellen. Transformermodellen beskrivs som en arkitektur designad för att effektivt kunna hantera sekvens-till-sekvens (sequence-to-sequence) uppgifter utan behovet av traditionella nätverkslager. Författarna framhåller att centralt för modellen är att träna teknik som liknar artificiella nervceller till att väga betydelsen av olika delar av en inmatningssekvens i samband med att den genererar en utmatningssekvens. Den förutspår nästkommande ord i en sekvens baserad på orden som föregick den, vilket resulterar i en slutprodukt i textform som liknar mänskligt skriven text med ett naturligt flöde. Sammantaget beskriver författarna att denna teknik således möjliggör parallell bearbetning av data för AI-modellerna och ger dem en djupare förståelse av sekvensrelationer som är särskilt fördelaktig i samband med uppgifter som berör just naturlig språkbehandling. Genom att använda positionella kodningar kan transformermodellen således även bevara ordning och kontext i data. Vidare beskriver

Bommasani et al. (2022) att så kallad transfer learning är en nyckeldel i transformermodellens funktion, där lärdomar från en uppgift tas med och appliceras på en annan uppgift. Detta genomförs genom pre-training, vilket är vad ChatGPT och andra stora språkmodeller som ses som dominanta idag, bygger på. Författarna beskriver att detta uppnås genom att träna modellen på stora mängder textdata med syfte att den ska lära sig känna igen och ta efter språkliga-, kontextuella- och grammatiska mönster i det mänskliga språket. Transformerarkitekturen och språkmodellering utgör således en viktig teknisk förklaring till den underliggande teknologin som möjliggör för avancerade förmågor inom NLP i hänseende att kunna generera sammanhängande, kontextuella och personliga svar.

2.1.2 Tekniska begränsningar och etiska dilemman

Litteraturen framhäver vissa begränsningar och etiska aspekter som bör beaktas i samband med användandet av stora språkmodeller. Weidinger et al. (2022) framlägger en taxonomi för dessa problem, där diskriminering, felaktig information och så kallad Human Computer Interaction Harms anses vara mest relevanta för denna studie.

Ray (2023) framhåller att stora språkmodeller kan ge upphov till snedvridna vinklingar i sina svar till användaren, där fördomar om exempelvis ras, kön, etnicitet och kultur kan framträda. Dessa snedvridningar beskrivs härstamma från att modellerna är tränade på mänskligt genererade data, som av naturliga skäl tenderar att bära subjektiva perspektiv.

Vidare beskriver Ji et al. (2023) att så kallade hallucinationer kan uppstå. Hua et al. (2023) beskriver hallucinationer i kontexten som text genererat av verktyget som kan vara både semantiskt och grammatiskt korrekt, men sakna betydelse eller består av osanningar. Fenomenet kategoriseras som antingen inre hallucinationer eller yttre hallucinationer. Inre hallucinationer refererar till när de inmatade instruktionerna från användaren inte speglas i svaret som genereras från modellen, medan yttre hallucinationer istället innebär att verktyget skapar innehåll som är direkt felaktigt eller saknar grund. I båda dessa fall anses hallucinationer bero dels på bristande kvalitet av den data som språkmodellen är tränad på, dels på hur algoritmerna är tränade i slutledningsförmåga och slutligen hur modellen har blivit "promptad" av användaren. Kasai et al. (2024) beskriver vidare att problematik med felaktigt innehåll i vissa fall kan vara relaterat till giltigheten av den information som språkmodellen tränats på. Vid de fall att information från träningsdata blir utdaterad uppstår en risk för att felaktig information genereras.

Tillsammans utgör dessa grunden i problematiken kring den sista aspekten, Human-Computer Interaction Harms. Weidinger et al. (2022) beskriver att betydande framsteg inom stora språkmodeller har förbättrat deras förmåga avsevärt, till den grad att vissa av dem nu förmår att efterlikna mänskligt språk mer än någonsin. Enligt författarna kan detta resultera i att användare utvecklar ett orimligt stort förtroende för språkmodellerna och omedvetet tilldelar en modell mänskliga kvaliteter. Weidinger et al. (2022) framhåller att detta kan påverka hur användaren interagerar med modellen och att det kan resultera i att tidigare nämnda tillkortakommanden, såsom fördomar och hallucinationer, befäster sig i användarens egna världsbild.

Slutligen beskrivs problem med integritet som ett återkommande dilemma för stora språkmodeller. Carlini et al. (2023) påpekar att modellerna uppvisar tendenser att memorera delar av träningsdata, vilket kan innebära att personligt identifierbar information oavsiktligt läcker ut om modellerna får specifika frågor från användaren. Staab et al. (2024) kompletterar

detta genom att visa hur modellerna också har förmågan att från träningsdata dra slutsatser om känslig information som härleds från träningsdata, även från till synes anonymiserad information. Detta kan leda till indirekt identifiering och profilering, och därmed potentiella integritetsintrång och avslöjanden av känsliga uppgifter om modellerna inte används på ett etiskt korrekt sätt.

2.2 Prompt engineering

2.2.1 Prompt engineering för effektiv användning av stora språkmodeller

Prompt engineering beskrivs av många vara en avgörande teknik för effektiv användning och fullt utnyttjande av potentialen hos stora språkmodeller. Denna åsikt delas av flera författare, där bland White et al., (2023), Chen et al., (2023), Ekin, (2023) och Giray (2023) som alla argumenterar för att prompt engineering är nyckeln till att optimera användningen av stora språkmodeller. Författarna beskriver sammantaget att prompt engineering i sammanhanget motsvarar en process där användaren skapar en uppsättning instruktioner som tilldelas språkmodellen genom input prompts för att anpassa språkmodellens generativa förmåga utefter konversationens specifika kontext och få den att genomföra en viss uppgift. Genom användandet av prompts beskrivs användaren kunna sätta riktlinjer eller regler för hur konversationen ska struktureras, vilken information som är viktig samt vilken form eller innehåll den önskade outputen bör ha och på så vis aktivt manipulera språkmodellen till att filtrera information och generera mer specifik och målinriktad output.

White et al. (2023) ger som exempel att en prompt skulle kunna användas för att instruera en språkmodell till att enbart generera kod som följer en specifik kodningsstil, medan Ekin (2023) framhåller att en prompt kan användas för att anpassa format, längd och omfattning av dess textsvar utifrån önskade begränsningar. White et al. (2023) beskriver vidare att kvaliteten på outputen är direkt beroende av kvaliteten på de prompts som användaren förser språkmodellen med. Genom strategiskt formulerade prompts uttrycker författarna att det därmed är möjligt att programmera en språkmodell till att utföra uppgifter som sträcker sig förbi enkel text- eller kodgenerering till att frambringa mer avancerad output, exempelvis i form av att generera ett quiz relaterat till ett visst ämne. Vidare beskriver författarna att det finns potential hos språkmodellerna för viss självanpassning i form av att själv kunna föreslå nya prompts som kan användas för att strategiskt samla ytterligare information eller generera annan relaterad output.

2.2.2 Prompting-metoder

Flertalet studier som gjorts i koppling till prompt engineering lyfter fram riktlinjer för användandet av prompt engineering. Två av dessa är *“The CLEAR path: A framework for enhancing information literacy through prompt engineering”* (Lo, 2023) och *“Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT”* (Korzyński et al. 2023). Studierna redogör för rekommendationer och bästa praxis för prompt engineering och presenterar därtill varsitt teoretiskt ramverk bestående av riktlinjer för optimal AI-prompting. Dessa bekräftar tidigare nämnd forskning i att påvisa hur effektiv utformning av prompts kan förbättra prestandan för stora språkmodeller och således generera mer exakta och kontextuellt relevanta resultat.

Korzyński et al. (2023) forskning riktar sig bland annat till utbildare och organisationer som integrerar AI i sin verksamhet och betonar behovet för dessa att investera i prompt engineering. Enligt författarna bör en effektiv prompt definiera följande element:

1. Context
2. Instruction
3. Input data
4. Expected output format.

Med grund i detta sammanställer författarna ett ramverk med rekommendationer för text-till-text prompting. Ramverket är specifikt utformat utifrån ChatGPT och Bard (numera Gemini).

Tabell 2:1 Rekommendationer för text-till-text prompting (Korzyński et al. 2023)

Acronym	Recommendation	Description
A	Articulate the instruction	Clearly state the task to be performed, such as 'write,' 'classify,' 'summarize,' or 'translate,' and specify how the output should look (table, list, Python code).
I	Indicate the prompt elements	Show the model where the instructions and input data are and what the expected output format should be.
P	Provide ending cues and context	Offer the model clear ending cues, such as three dots for continuation or a colon, dot, or placeholder like 'answer:' for indicating a response is needed. Furthermore, ground the model by providing a context for the task (e.g. 'You are a manager of a tech team').
R	Refine instructions to avoid ambiguity	Give the model-specific instructions and a detailed description of the task to avoid any confusion or imprecision.
O	Offer feedback and examples	For conversational models, such as ChatGPT, feedback on the model's responses can help it better understand the desired output. Moreover, providing the model with a few examples of expected responses (few-shot learning) can help it adapt its style and way of responding.
M	Manage interaction	Treat the model as your sparring partner, asking it to provide counterarguments or point out flaws in your ideas.
T	Track token length and task complexity	Break complex tasks into smaller steps for better performance. Remember to control the token length, keeping the prompt and response under the token limits of the model (usually 4096 tokens for commercially available LLMs). The token length of a text can be checked here: https://platform.openai.com/tokenizer .

Lo (2023) presenterar ytterligare ett teoretiskt ramverk för prompt engineering, designat i syfte att optimera interaktioner med stora språkmodeller. Ramverket framhåller fem huvudprinciper för framgångsrik prompting:

1. Concise
2. Logical
3. Explicit
4. Adaptive
5. Reflective

Genom att integrera dessa fem principer i utformningen av prompts uttrycker Lo (2023) att användare kan förbättra precisionen och relevansen i svaren från stora språkmodeller. Författaren betonar att ett metodiskt tillvägagångssätt i skapandet av prompts enligt dessa principer inte bara förbättrar interaktionens kvalitet utan också bidrar till en djupare förståelse för hur språkmodeller fungerar och svarar.

Trots att de två ramverken inte är identiska kan det sammantaget sägas att de båda lyfter fram liknande rekommendationer för effektiv prompting. Båda ramverken betonar vikten av ett strategiskt tillvägagångssätt och betydelsen för en användare av att skapa tydliga, kontextuella och specifika uppmaningar för att effektivt kommunicera med språkmodeller. Vidare är de överens om behovet av kortfattade och logiska prompts som beskriver den aktuella uppgiften, vikten av att ge tydliga instruktioner för att undvika oklarheter, samt vikten av anpassningsförmåga och reflektion för att förbättra interaktionen med språkmodellen. Sammantaget kan det därmed sägas att prompting enligt författarna kännetecknas av en balans mellan detaljerad vägledning och flexibiliteten att förfina, vilket förbättrar både effektiviteten i kommunikationen och kvaliteten på språkmodellens svar. Tillsammans lägger ramverken en gemensamma grund som beskriver ett rekommenderat tillvägagångssätt för prompting som ses avgörande för att utnyttja språkmodellens fulla potential.

Avslutningsvis framhåller litteraturen ett fokus på vissa specifika tekniker för prompting. Nedan beskrivs några centrala sådana tekniker:

- *Zero-shot prompting*: språkmodellen besvarar frågor den aldrig sett förut utan att ha getts exempel på liknande kategorier eller koncept i prompten (Naveed et al., 2024).
- *Few-shot prompting*: Användare förser en språkmodell med flera input-output demonstrationspar i syfte att få språkmodellen att generera output i önskat format (Brown et al., 2020).
- *Chain-of-thought prompting*: Innebär att modellen förklarar sitt svar för användaren i flera logiska steg som följer varandra. Det ger modellen en struktur för att bryta ner och lösa problemet steg för steg, vilket gör lösningen och resonemanget bakom svaret mer förståeligt för användaren (Zhou et al., 2022).
- *Role-prompting*: Användaren tilldelar språkmodellen en specifik roll att spela, exempelvis en hjälpsam assistent eller en kunnig expert. Detta görs i syfte att guida språkmodellen till att förse användaren med svar som bättre stämmer överens med förutsättningarna för den önskade outputen (Chen et al., 2023).
- *Iterative prompting*: Innebär att användaren successivt finjusterar och bygger på språkmodellens svar genom att ställa uppföljningsfrågor eller ge kompletterande instruktioner baserat på tidigare svar (Wang et al., 2022).

2.3 Digital kompetens

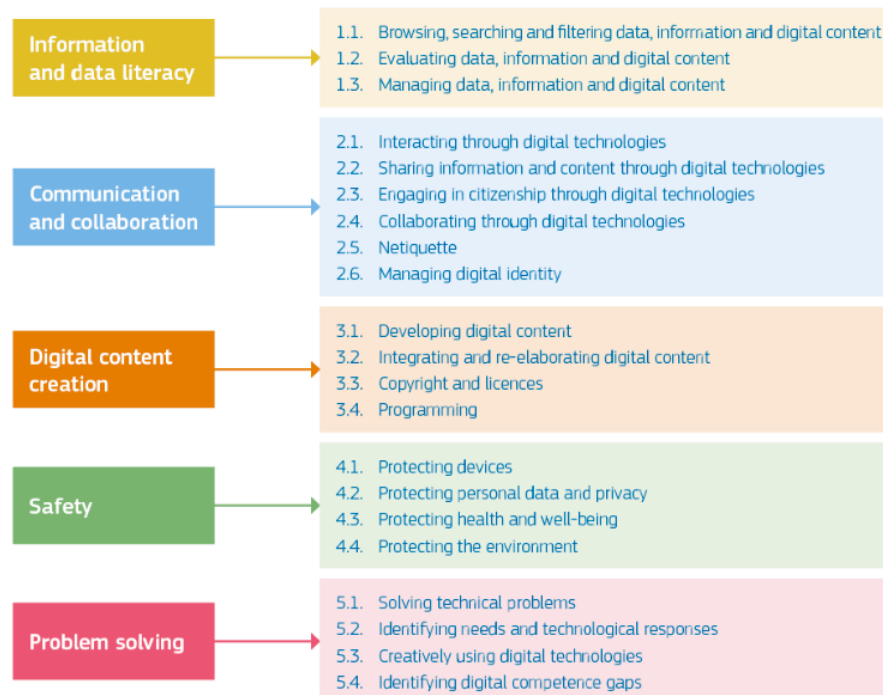
I maj 2018 framställde Europeiska Unionen (EU) ett ramverk med rekommendationer på åtta nyckelkompetenser för livslångt lärande. Ramverket är menat att fungera som ett referensverktyg och syftar att skapa en gemensam förståelse för kompetenser som anses behövas nu och i framtiden för att medborgare ska nå bland annat anställningsbarhet, aktivt medborgarskap och social inkludering. Ramverket betonar lärarens centrala roll i att främja kompetensutveckling och uppmanar medlemsstaterna att stödja utbildningspersonal för att säkerställa kvalitativ utbildning och lärarkompetens för att möta samhällets och individens föränderliga behov (European Commission, 2019).

En av nyckelkompetenserna som framhålls av EU:s rekommendationer är digital kompetens. Sammanfattningsvis kan digital kompetens enligt EU-kommissionens definition ses som en omfattande förmåga som berör hur individer använder digitala tekniker på ett säkert, kritiskt och ansvarsfullt sätt inom lärande, arbetsliv samt samhällsdeltagande i stort. Digital kompetens omfattar hur individer hanterar information och data samt hur de kommunicerar och samarbetar via digitala plattformar. Detta inkluderar en förståelse för medier, skapande av digitalt innehåll och att kunna navigera i frågor kring säkerhet och digitalt välmående (European Commission, 2019).

2.3.1 Digital kompetens hos lärare

Med grund i EU:s betoning på digital kompetens har EU kommissionen framställt *The Digital Competence Framework for Citizens (DigComp)*, vilket är ett ramverk för att definiera och beskriva nyckelområden av digital kompetens som anses avgörande för att medborgare ska kunna agera säkert, kritiskt och ansvarsfullt i en digitaliserad värld, samt stödja livslångt lärande och anpassning till snabbt föränderliga teknologiska miljöer. DigComp 2.2 understryker lärarnas centrala roll i utvecklingen av digital kompetens där lärare inte bara förväntas undervisa i tekniska färdigheter utan även främja kritiskt tänkande och etisk förståelse i samband med digital teknik. Ramverket framhåller vikten av att lärare anpassar sin undervisning till framväxande teknologier, säkerställer inkluderande tillgång till digital utbildning för alla elever samt uppmuntrar till kontinuerligt lärande. Utifrån DigComp-ramverket utgörs digital kompetens av 5 kompetensområden; Information och datakompetens, Kommunikation och samarbete, Skapande av digitalt innehåll, Säkerhet och Problemlösning (Vuorikari et al., 2022).

Figur 2:1 The DigComp conceptual reference model (Vuorikari et al., 2022)



Flera studier inom området digital kompetens är relaterade till utbildningssektorn och framhäver, i enlighet med EU-kommissionens ståndpunkter, vikten av digital kompetens hos lärare. Dessa studier framställer digital kompetens som en av de största utmaningarna för lärare i dagens digitala era och betonar behovet av kvalificerade, digitalt kompetenta lärare som är kapabla att använda teknik för att understödja studenters utveckling. Detta anses specifikt avgörande för lärare och studenter inom högre utbildning, särskilt i en tid där digital teknik får en allt större roll i undervisningen (Fernández-Batanero et al., 2020; Bojukrapan et al., 2023; Babaev et al., 2020).

Vidare är Pozo-Sánchez et al. (2020) och Babaev et al. (2020) överens om att det finns en direkt koppling mellan digital kompetens och pedagogisk innovation. Författarna belyser inte bara vikten av att lärare behärskar digitala verktyg utan även deras förmåga att kunna integrera dessa verktyg på ett sätt som förbättrar kvaliteten och tillgängligheten i undervisningen. Genom att utveckla och stärka lärarnas digitala kompetens anses pedagogisk innovation och kvalitetsförbättring i undervisningen främjas. Författarna anser att digital kompetens således sträcker sig bortom grundläggande tekniska färdigheter till att omfatta en djupare förståelse för hur digitala resurser kan användas för att stödja och berika lärandeprocessen. Digital kompetens kan således ses som en katalysator för att integrera nya undervisningsmetoder och teknologier som kan berika lärandeupplevelser och göra utbildningen mer tillgänglig och inkluderande.

Slutligen råder konsensus gällande att det finns ett omfattande och växande behov av att kontinuerligt utbilda och stödja lärare i deras professionella utveckling av digital kompetens. I detta inkluderas även fortbildning och metodiskt stöd för att lärare regelbundet ska kunna uppdatera och fördjupa sina färdigheter. Detta anses avgörande för att kunna möta nuvarande såväl som framtida utmaningar inom utbildningssektorn och för att säkerställa att lärare är väl förberedda för att navigera och utnyttja de möjligheter som digital teknik erbjuder (Pozo-Sánchez et al., 2020; Babaev et al., 2020; Fernández-Batanero et al., 2020; Bojukrapan et al., 2023). OECD (2021) använder Covid-19-pandemin som ett exempel för att belysa vikten av kontinuerlig kunskapsutveckling för lärare som en nödvändighet för framtidens undervisning. OECD framhåller att snabba omställningar, liknande de som pandemin medförde, förväntas bli ett allt vanligare fenomen i framtida undervisningssammanhang.

2.3.2 Prompt engineering som digital kompetens

Korzyński et al. (2023) undersökte AI-prompt engineering som en digital kompetens, med syfte att belysa utmaningar och framtidsutsikter som kan förknippas med metoden. I sin analys använde författarna sig bland annat av DigComp 2.2 ramverket med syfte att analysera AI-prompt engineering som en digital kompetens.

Tabell 2:2 AI-prompting som en digital kompetens enligt DigComp 2.2. (Korzyński et al., 2023)

Area	Competence	AI prompt engineering
1. Information and data literacy	1.1. Browsing, searching, and filtering data, information, and digital content	✓
	1.2. Evaluating data, information, and digital content	✓
	1.3. Managing data, information, and digital content	✓
2. Communication and collaboration	2.1. Interacting through digital technologies	✓
	2.2. Sharing through digital technologies	✓
	2.3. Engaging in citizenship through digital technologies	
	2.4. Collaborating through digital technologies	✓
	2.5. Netiquette	
	2.6. Managing digital identity	
3. Digital content creation	3.1. Developing digital content	✓
	3.2. Integrating and re-elaborating digital content	✓
	3.3. Copyright and licenses	
	3.4. Programming	✓
4. Safety	4.1. Protecting devices	
	4.2. Protecting personal data and privacy	
	4.3. Protecting health and well-being	
	4.4. Protecting the environment	
5. Problem solving	5.1. Solving technical problems	✓
	5.2. Identifying needs and technological responses	✓
	5.3. Creatively using digital technologies	✓
	5.4. Identifying digital competence gaps	✓
	5.5. Learning in digital environments	✓

Analysen mappar prompt engineering mot de 5 kompetensområden som definieras i DigComp 2.2 ramverket. I den högra kolumnen har markeringar gjorts där respektive kompetens anses nyttjas. Baserat på resultatet kan det konstateras att det tycks finnas en stark koppling mellan prompt engineering och digital kompetens inom 4 av de 5 listade kompetensområdena, med särskilt stark koppling till områdena *information och datakompetens*, samt *problemlösning*.

2.4 Stora språkmodellers potential inom utbildning

2.4.1 Nuvarande kunskapsläge och relaterad forskning

Bates (2020) framhäver att AI har blivit alltmer relevant i samband med utformningen av framtidens högre utbildning, där det specifikt för just stora språkmodeller har det visat sig att dessa verktyg kan användas för lärare att exempelvis planera lektioner, genomföra rättning samt vara behjälplig i skapandet av lektionsmaterial (Naveed et al., 2024; Trust et al., 2023; Montenegro-Rueda et al., 2023). Författarna får medhåll av Kasneci et al. (2023) som framhåller att stora språkmodeller har potentialen att revolutionera undervisningen och samtidigt fungera som ett assisterande verktyg till läraren i dennes yrkesroll. Författarna belyser flertalet områden där stora språkmodeller med fördel kan användas inom utbildning:

- Individanpassat inläring
- Lektionsplanering
- Akademiskt skrivande och forskningsrelaterat arbete

- Rättning och utvärdering
- Förse studenter med relevanta kompetenser för framtidens samhälle
- Egen utveckling i sin professionella roll

Kasneci et al. (2023) förklarar dessa punkter närmare. Genom att ta hjälp från språkmodeller kan lärare skapa *individ Anpassad inläring*. Språkmodellerna kan användas för att analysera elevtexter och sedan förse dem med individanpassad feedback och materialförslag som går i linje med studentens specifika behov. Genom detta kan studentens lärande förbättras samtidigt som det sparar tid för läraren i fråga. Detta medför att lärare istället kan fokusera på andra områden inom utbildningen, såsom engagemang och interaktiv undervisning. Vidare kan de användas till *lektionsplanering*. Lärare kan förse dessa AI-modeller med kurslitteratur och annat material som de vill basera en kurs på och som svar få en kursplan som går i linje med lärarens mål med kursen. På samma sätt kan denna typ av AI skapa relevant övningsmaterial som säkerställer att studenterna lär vad som krävs för att klara kursen. Stora språkmodeller beskrivs även kunna användas för *akademiskt skrivande*, både för lärares eget bruk och som stöd för att ge feedback till studenters arbete. Språkmodellen kan assistera i att hitta grammatiska felaktigheter, hitta förbättringsområden i texterna och agera som medel för att sammanfatta källor samt förklara det väsentliga innehållet, vilket sammantaget ökar lärares förståelse av en källa. Språkmodellerna beskrivs även kunna användas för att *automatisera rättning och feedback* genom att identifiera styrkor och svagheter i studenters arbete. Detta innebär både en tidsbesparing för lärare och en möjlighet att ägna tid åt mer individuell feedback. Chang et al. (2024) instämmer i dessa påståenden, men framhåller även att verktygen kan sakna de insiktsfulla perspektiv som en mänsklig lärare har angående hur detta kan översättas till vidare lärande hos studenten, och därmed alltså inte fullt ut kan ersätta läraren. Ytterligare en användningspotential som framhävs av Kasneci et al. (2023) är att *förse studenter med framtidens kompetenser* genom att låta dem bekanta sig med en transformativ teknologi, såsom stora språkmodeller. Detta möjliggör för studenter att lära sig etisk och korrekt användning av AI, vilket förväntas vara en nyckelkompetens i framtidens samhälle. Slutligen kan användandet av dessa verktyg även bidra till att *utveckla lärares egen kompetens* i deras yrkesroll, genom att de förblir uppdaterade med den senaste teknologin som i sin tur säkerställer deras digitala kompetens. Chang et al. (2024) nämner också informationssökning som en av de delar stora språkmodeller kan användas till. Det har visat sig att med samma resultat, har de som använt stora språkmodeller hittat likvärdig information snabbare jämfört med de som använt sig av Googles sökmotor. En reservation gällande detta fynd verkar däremot vara att detta inte verkar fungera lika väl på domänspecifika ämnen.

Det råder dock ännu en viss osäkerhet kring hur stora språkmodeller bäst implementeras i en undervisningskontext, särskilt med hänsyn till etiska aspekter. EU:s (2016) dataskyddsförordning (GDPR) fastställer ett lagligt ansvar för lärosäten och dess medarbetare att säkerställa skyddet av personuppgifter, bland annat från obehörig åtkomst. Detta inkluderar alla former av personlig information som samlas in om elever, såsom namn, adresser, och utbildningsrelaterad information. Enligt författarna Nguyen et al. (2022) saknas konsensus om vilka etiska principer som bör följas för att en sådan implementering ska anses lämplig inom specifikt utbildningssammanhang, detta trots goda intentioner hos såväl utvecklare av AI-verktygen som lärare. De betonar att digital kompetens vad gäller just AI krävs från både lärare och samhälle för att på ett säkert sätt se till att dessa vävs in på ett säkert sätt i utbildningen. Enligt Nguyen et al. (2022) skulle detta inte enbart resultera i säkrare implementering av AI-verktyg, utan även förbättrade undervisningsmoment. På liknande spår betonar Cerratto et al. (2024) att språkmodeller är en produkt av medarbetare hos privata företag vars kommersiella intressen mycket väl kan styras av ambitioner om

teknisk innovation och vetenskapliga framsteg, men desto mindre för vad dessa verktyg får för implikationer på andra viktiga sektorer inom samhället som högre utbildning.

2.4.2 Lärares roll vad gäller AI inom utbildning

Isemonger (2023) argumenterar för att den mest centrala frågan i koppling till generativa språkmodeller och utbildning är att förstå vilken roll läraren kommer att ta i detta. Författaren uttrycker att det har funnits en allmän oro i att lärarrollen kommer att ersättas av denna typ av verktyg, men författarens själv argumenterar för att de snarare kommer att behöva förbereda sig för att anta en ytterligare roll i form av expertanvändare av verktygen för att kunna utbilda elever till att bli detsamma. Författaren tror därmed att ChatGPT och liknande verktyg kommer att medföra en omkonfigurering av lärares roll snarare än en ersättning av dem. Påståendet kompletteras av Ausat et al. (2023) som framhäver att lärare måste ha en grundlig förståelse av den teknik de använder för att kunna nyttja tekniken till sin fördel i arbetet, samt behöver utveckla ny kompetens och färdigheter när det kommer till att integrera tekniken i sina utbildningsmoment. Montenegro-Rueda et al. (2023) framhåller att implementeringen av stora språkmodeller inom utbildningsmiljön kan ha en positiv påverkan på lärandeprocesser. Resultaten framhäver dock vikten av att lärare utbildas i att använda verktyget samt att de är bekanta med hur dessa fungerar. Även Eager & Brunton (2023) argumenterar för att lärares roll i framtiden kommer att skifta i takt med att AI tillåts ta större plats inom utbildningen, i form av att högre krav kommer att ställas på lärares kompetens och anpassning. Författarna lyfter särskilt prompt engineering som ett medel för att uppnå ett effektivt och givande samarbete med AI-modeller, där slutprodukten i form av engagerande, träffsäkert och pedagogiskt material för undervisning är av högsta kvalitet. Slutligen har även Jeon & Lee (2023) dragit slutsatsen att samarbetet mellan mänskliga lärare och stora språkmodeller bör stärkas för att förbättra utbildningskvaliteten.

2.4.3 Universitetens roll

Aoun (2017) argumenterar istället för att institutioner för högre utbildning, snarare än lärarna, har ett unikt samhällsansvar i att säkerställa att dagens medborgare är väl rustade för ett framtida samhälle genom att främja livslångt lärande. Författaren bygger sitt resonemang på faktumet att AI och annan teknik kommer att revolutionera vår ekonomi och samhälle som gör att utbildningsväsen, specifikt högre utbildning, har ett ansvar i att fastställa en sådan miljö för såväl studenter som medarbetare. Vanthournout et al. (2014) argumenterar på liknande vis för att för just kunskapsarbetare, en kategori där läraryrket ingår, är det organisationens roll att underlätta deras professionella utveckling och inläring genom att kontinuerligt erbjuda relevanta möjligheter. Anoun (2017) understryker att i takt med att AI eliminerar behovet av mänskligt arbete i flertalet sektorer, kommer yrkesverksamma individer i högre grad att stå inför kravet att kontinuerligt se över och utveckla sina kompetenser. Universitet anses ha en central roll i ett sådant utformande.

2.5 Teoretiskt resultat

Den sammantagna litteraturen som granskats pekar på att en eventuell integration av stora språkmodeller inom utbildningssektorn innebär både betydande möjligheter och omfattande utmaningar. Litteraturen betonar att stora språkmodeller kan effektivisera och förbättra pedagogiska processer genom verktyg för individanpassat lärande, lektionsplanering och

akademiskt stöd. För studenter kan detta innebära individanpassat inlärningsmaterial och feedback vilket resulterar i en mer dynamisk lärandemiljö. För lärare kan det istället handla om avlastning från repetitiva och tidskrävande uppgifter, vilket frigör tid för mer engagerande och interaktiva undervisningsmetoder.

Oro över att AI skulle kunna ersätta lärarnas roll har förvisso uttryckts, men litteraturen visar snarare på att stora språkmodeller kommer att transformera lärarrollen där lärare förväntas anta en mer vägledande och faciliterande roll och besitta kunskaper om hur man använder AI-verktyg på ett pedagogiskt och etiskt sätt. Detta innebär en förskjutning från traditionell kunskapsöverföring till att bli mentorer och facilitatorer som använder teknik för att berika lärandet. Forskningen framhåller således att lärarens roll är central, dels i att rusta elever för framtida utmaningar i ett allt mer digitaliserat samhälle, men även för att själva hålla sig uppdaterade på den senaste tekniken. Digital kompetens, som inkluderar både teknisk kunskap och ett kritiskt förhållningssätt till digitala verktyg, ses därför som avgörande för lärare att utveckla, både för egen del men också för att kunna överföra denna kompetens till sina elever. Prompt engineering framträder som en särskilt relevant färdighet för att effektivt hantera dessa utmaningar, vilket betonar behovet av teknisk kompetens i kombination med pedagogiska färdigheter.

Den litteratur som presenteras understryker vidare universitetens och andra utbildningsinstitutioners roll i att säkerställa att såväl lärare som studenter är väl förberedda för att hantera och dra nytta av dessa nya teknologier. Institutionerna beskrivs ha ett ansvar att erbjuda fortbildning och stöd för att lärare ska kunna integrera och använda dessa teknologier effektivt i sin undervisning. Det är avgörande att dessa institutioner möjliggör en miljö där kontinuerlig utbildning och utveckling av digital kompetens kan ske. Detta säkerställer att lärare och studenter inte bara håller jämna steg med tekniska framsteg utan också kan navigera i dess användning på ett kritiskt och etiskt sätt.

Litteraturen identifierar även viktiga etiska utmaningar som stora språkmodeller medför, såsom risk för diskriminering, felaktig information och etiska implikationer, vilka kräver en noggrann hantering. Dessa problem uppstår ofta på grund av fördomar i träningsdata och brister i AI-modellernas förmåga att förstå komplexa mänskliga kontexter. Lärare och utbildningsinstitutioner måste vara medvetna om dessa begränsningar och aktivt arbeta för att implementera teknologier på ett ansvarsfullt och inkluderande sätt.

Sammanfattningsvis landar den teoretiska argumentationen i att stora språkmodeller erbjuder betydande potential för att förbättra och individualisera lärandet, men att de samtidigt medför betydande ansvar och behov av omsorgsfull hantering när det gäller utbildningens kvalitet och etik. Denna dubbla dynamik mellan potential och risk gör det nödvändigt för utbildningssektorn att noggrant navigera i hur dessa verktyg implementeras och används.

Tabell 2:3 Teoretisk resultatmodell

Område	Relevanta faktorer	Litteratur
Stora språkmodeller	<ul style="list-style-type: none"> - Tekniska begränsningar - Etiska hinder 	Khurana et al. (2022) Rathore (2023) Naveed et al. (2024) Vaswani et al. (2017) Bommasani et al. (2022) Weidinger et al. (2022) Ray (2023) Ji et al. (2023) Hua et al. (2023) Kasai et al. (2024) Carlini et al. (2023) Staab et al. (2024)
Prompt engineering	<ul style="list-style-type: none"> - Betydelse för effektiv användning av stora språkmodeller - Strategier för att utforma effektiva prompts 	White et al. (2023) Chen et al. (2023) Ekin (2023) Giray (2023) Lo (2023) Korzyński et al. (2023) Naveed et al. (2024) Brown et al. (2020) Zhou et al. (2022) Wang et al. (2022)
Digital kompetens	<ul style="list-style-type: none"> - DigComp ramverket - Utbildning - Kritisk förhållning 	European Commission (2019) Vuorikari et al. (2022) Fernández-Batanero et al. (2020) Bojukrapan et al. (2023) Babaev et al. (2020) Pozo-Sánchez et al. (2020) OECD (2021) Korzyński et al. (2023)
Språkmodellens potential inom utbildning	<ul style="list-style-type: none"> - Tillämpningsområden - Etiska överväganden - Lärares roll och kompetens 	Bates (2020) Naveed et al. (2024) Trust et al. (2023) Montenegro-Rueda et al. (2023) Kasneci et al. (2023) Chang et al. (2024) Nguyen et al. (2022) Cerratto et al. (2024) Isemonger (2023) Ausat et al. (2023) Eager & Brunton (2023) Jeon & Lee (2023) Aoun (2017) Vanthournout et al. (2014)

3 Metod

I följande kapitel presenteras den metod som använts för litteratur- och datainsamling samt för analys av den empiriska data som ligger till grund för studien. Inledningsvis motiveras valet av metod och därefter följer en diskussion rörande urval samt insamling och bearbetning av data. Därefter diskuteras uppsatsens validitet och reliabilitet, följt av några avslutande etiska aspekter som beaktats för att upprätthålla forskningens integritet.

3.1 Val av metod

Denna studie har influerats av en kvalitativ metod. Enligt Recker (2013) används en kvalitativ metod med fördel som forskningsmetod för icke-numerisk datainsamling som är fokuserad på att förstå sociala fenomen inom en viss kontext. Författaren beskriver den kvalitativa metoden som lämplig vid djupgående undersökning av ämnen som är otillräckligt utforskade eller som fortfarande är under framväxt, vilket gör den till en lämplig inspiration av metod för denna studie. Vidare grundas valet av metod i insikten om att fenomenet att använda stora språkmodeller inte kan ses som isolerat från sin kontext. Universitetslärares användning av denna typ av verktyg bedöms nära integrerad med och påverkad av en rad kontextuella faktorer. En kvalitativ metod erkänner och tar hänsyn till sambandet mellan det fenomen som undersöks och dess kontext, och anses därmed ge en mer heltäckande förståelse för hur stora språkmodeller integreras i universitetslärares yrke.

Vidare söker undersökningen att skapa en omfattande och detaljerad bild av universitetslärares inställning till och erfarenheter av användandet av stora språkmodeller som ett arbetsverktyg. Eftersom studien inte enbart undersöker lärares användning av verktyget utan även faktorer som påverkar användandet, stämmer en kvalitativ metod väl överens i detta avseende med målet för undersökningen. Recker (2013) beskriver att kvalitativ metod möjliggör för en mer nyanserad datainsamling och således kan lärarnas erfarenheter, känslor och åsikter tillåtas utgöra den empiriska grunden för studien. Med detta tillåts således en grundlig analys som kan bidra till att uppnå en mer omfattande och mångfacetterad förståelse för ämnet.

Datainsamlingen beskrivs av Recker (2013) vara fokuserad kring att undersöka beteenden och åsikter kring ett särskilt fenomen och sker i regel via personliga interaktioner, såsom intervjuer, observationer eller genom att studera dokumentation. För insamlingen av kvalitativ data har personliga intervjuer med sju stycken universitetslärare genomförts. Syftet med intervjuerna var att skapa en grundlig förståelse för hur dessa lärare använder sig av stora språkmodeller i det dagliga arbetet och fånga upp deras personliga reflektioner och attityder kopplade till användningen.

Recker (2013) framhäver även vissa svagheter som medföljer den kvalitativa metoden, däribland svårigheten att generalisera respondenternas åsikter till en större population samt de reliabilitetsproblem som uppstår till följd av metodens djupt kontextualiserade natur. Även om dessa svårigheter har beaktats vid val av metod, har det emellertid bedömts ha mindre relevans för den specifika studien. Studien syftar inte till att skapa en allmängiltig bild av hur samtliga universitetslärare använder stora språkmodeller, utan istället utforska och analysera specifika erfarenheter och perspektiv från utvalda lärare. Denna inriktning ger en fördjupad

förståelse för hur stora språkmodeller används och uppfattas i respondenternas specifika kontext och belyser nyanser som är kritiska för att förstå de underliggande dynamikerna i användningen av stora språkmodeller. Genom att fokusera på individuella erfarenheter erkänns och utnyttjas de kontextualiserade insikterna som en styrka, vilket går hand i hand med studiens fokus på att erbjuda detaljerade insikter snarare än breda generaliseringar. Trots att kvalitativa metoder kan vara svåra att replikera i andra sammanhang anses denna egenskap nödvändig för att skapa en mer djupgående förståelse.

3.2 Urval av respondenter

Urvalet för intervjuerna har utgjorts av sju respondenter. Undersökningen har förlitat sig på ett ändamålsenligt urval, vilket Oates et al. (2022) och Recker (2013) beskriver som en urvalsprocess där respondenter medvetet väljs ut till följd av att de besitter vissa intressanta egenskaper eller baserat på deras potential att bidra med värdefull data. En förutsättning för att respondenterna skulle ses som lämpliga intervjuobjekt var att dessa hade testat att använda sig av någon form av stor språkmodell för arbetsrelaterade uppgifter. Med stöd i detta genomfördes ett ändamålsenligt urval, där inklusionskriterierna begränsade urvalet till lärare som bekräftat att de använt sig av ett sådant verktyg i sitt arbete. För att säkerställa att inklusionskriterierna var uppfyllda kontaktades initialt potentiella respondenter från populationen universitetslärare via mejl. Syftet med mailet var att ge mottagarna en övergripande beskrivning av studien och en förståelse för vad intervjun kommer att behandla. För att avgöra ifall de kontaktade var relevanta intervjuobjekt frågades de i samband med detta ifall de använder sig av någon variant av en stor språkmodell i sitt dagliga arbete samt huruvida de ställde sig intresserade till att delta i studien eller inte. Baserat på deras svar valdes sedan sju lämpliga kandidater ut som respondenter. I första hand prioriterades alltså urvalet utifrån bekräftad användning av stora språkmodeller. Därefter gjordes ett ytterligare urval utifrån förutsättningen att försöka generera en mångfald i urvalsgruppen. Detta baserades på kriterierna kön och universitets- samt institutionell tillhörighet och gjordes med förhoppning att inkludera bredare perspektiv och erfarenheter i studien.

Tabell 3:1 Respondentöversikt

Respondent	Universitet	Ämne	Plats	Datum	Tid	Bilaga
R1	U1-Sydsverige	Affärsrelaterat	Fysisk	17/4	43 minuter	1
R2	U1-Sydsverige	Teknikrelaterat	Fysisk	18/4	42 minuter	2
R3	U1-Sydsverige	Teknikrelaterat	Fysisk	18/4	43 minuter	3
R4	U2-Mellansverige	Utbildningsvetenskapligt	Google Meet	22/4	55 minuter	4
R5	U2-Mellansverige	Samhällsvetenskapligt	Zoom	23/4	29 minuter	5
R6	U2-Mellansverige	Samhällsvetenskapligt	Google Meet	23/4	30 minuter	6
R7	U3-Mellansverige	Finansrelaterat	Google Meet	23/4	29 minuter	7

3.3 Datainsamling

3.3.1 Litteraturinsamling

Inledningsvis har en noggrann granskning av befintlig forskning inom studieområdet genomförts, där vetenskapliga sökmotorer som Google Scholar och LUBSearch använts för att hitta akademiska artiklar, journaler, e-böcker och andra källor som anses vara relevanta för denna studie. Sökningar har kombinerat sökord som:

- Large Language Models/LLM
- Higher education
- University
- Faculty
- Teachers
- Prompt Engineering
- Digital competence

Sökningarna har i stor mån filterats till publikationer från 2022 och framåt. Filtreringen motiveras av den snabba utveckling och växande intresset för stora språkmodeller som ChatGPT medförde. Sökningen begränsades därmed i syfte att fånga relevant forskning och säkerställa att analysen speglar den nuvarande diskursen kring teknologin.

Därefter granskades dessa publikationer för att filtrera ut de som ansågs relevanta för studien. Utöver detta har även fysiska metodikböcker använts. Denna systematiska översikt av litteraturen var avgörande för att etablera en heltäckande bild av det nuvarande kunskapsläget, identifiera kunskapsluckor och säkerställa relevansen för forskningsbidraget mot att förstärka den befintliga kunskapsbasen. Studien bygger således på befintliga insikter och strävar efter att utvidga förståelsen inom det valda ämnesområdet.

3.3.2 Utformning och genomförande av intervjuer

Personliga intervjuer genomfördes för datainsamling med syfte att generera kvalitativ data. Intervjuer beskrivs av Recker (2013) som den mest framträdande formen av kvalitativ datainsamlingsteknik då de har fördelen att vara målinriktade och, till följd av deras detaljrika karaktär, även mycket insiktsfulla. Oates et al. (2022) argumenterar för att intervjuer är en passande form av datainsamling i de fall att en undersökning ämnar att utforska upplevelser, vilket går i linje med denna studie. Vidare beskriver Recker (2013) att intervjuer tenderar att anta mer eller mindre formella strukturer beroende på vilken typ av intervju det avser. Intervjuerna som genomförts i samband med denna studie har antagit en semistrukturerad struktur. Detta beskriver Recker (2013) innebära att intervjun antar en något mer flexibel struktur. Respondenterna ställs inför en rad förutbestämda frågor, samtidigt som utrymme lämnas för följdfrågor och dubbelriktade diskussioner kopplade till ämnet. Intervjun beskrivs på så vis följa en mer avslappnad konversationsstil och är endast delvis vägleda i den mening att en intervjuguide innehållande förbestämda frågor används som ett ramverk för intervjun. Med grund i detta utgick de genomförda intervjuerna från de frågor som definierades på förhand men därtill tilläts viss styrning av samtalet i form av mer specifika följdfrågor som

uppstod ur diskussionen och med relevans till ämnet. Recker (2013) belyser en rad fördelar med semistrukturerade intervjuer, vilka bland annat inkluderar att intervjuer tenderar att upplevas som mindre påträngande för den som intervjuas tack vare att intervjustrukturen uppmuntrar tvåvägskommunikation, samt att intervjuerna tenderar att inte bara belysa svaren utan även skälen bakom svaren.

Samtliga intervjuer genomfördes personligen och i enrum med en respondent i taget för att undvika att respondenterna färgades av varandras svar. Fysiska intervjuer förespråkades framför digitala intervjuer i den mån det var möjligt. Recker (2013) påstår att individer som intervjuas personligen och i ett mer avslappnat samtalssätt ofta är mer benägna att diskutera känsliga frågor. Intervjun följde strukturen av de intervjufrågor som definierats på förhand i intervjuguiden (tabell 3:2, alt. bilaga 8). Dessa har valts med utgångspunkt i litteraturen och med målet att kunna besvara forskningsfrågan. För att säkerställa att viktig information inte gick förlorad spelades samtliga intervjuer in.

Recker (2013) framhäver vissa svagheter med intervjuer. Där bland att respondenten inte svarar helt ärligt på frågorna till följd av att den vill framställa sig själv i ett positivt ljus eller att de svarar utifrån vad den tror att intervjuaren vill höra. Dåligt formulerade eller ledande intervjufrågor beskrivs också kunna orsaka en snedvriden bild av verkligheten. För att hantera dessa risker har frågorna utformats enligt principen om öppna frågor. Att intervjuerna genomfördes personligen främjar vidare en mer personlig relation till intervjuobjektet, något som enligt Recker (2013) ofta resulterar i att intervjuobjekten är mer benägna att diskutera känslig information, vilket ökar möjligheten att svaren förbli sanningsenliga. Slutligen lovades respondenterna av samma anledning även full anonymitet. Värt att notera är att respondenterna inte tillhandahölls frågorna i förväg och därför inte hade möjlighet att förbereda sina svar. Följaktligen kan vissa relevanta aspekter som potentiellt hade kunnat påverka studiens resultat ha uteblivit från diskussionen. Detta bör beaktas vid tolkning av resultaten.

3.3.3 Intervjuguide

Tabell 3:2 - Intervjuguide

Nr	Fråga	Tema
1	Kan du kort beskriva dina huvudsakliga arbetsuppgifter?	Bakgrund
2	Vilka stora språkmodeller använder du dig av i ditt arbete?	Kartläggning av användning
3	Kan du beskriva hur du och till vilka arbetsuppgifter du använder stora språkmodeller?	
4	Anser du att användningen av stora språkmodeller på något vis har förbättrat eller förenklat ditt arbete?	
5	Har du fått någon utbildning, riktlinjer eller vägledning i hur du kan använda stora språkmodeller i ditt arbete?	Digital kompetens
6	Arbetar du aktivt med prompting av språkmodeller, och i sådant fall, på vilket sätt?	Prompting
7	Hur ser du på vikten av att lärare utvecklar sin digitala kompetens? - Ser du något samband mellan digital kompetens och användandet av stora språkmodeller?	Digital kompetens
8	Har du stött på några utmaningar eller begränsningar med att använda stora språkmodeller och hur har du i sådant fall hanterat dessa?	Begränsningar
9	Hur förhåller du dig kritisk till det stora språkmodeller genererar?	Digital kompetens
10	Hur ser du på frågor om etik och integritet gällande användningen av stora språkmodeller i undervisningen och administrativa syften? - Påverkar detta hur du använder stora språkmodeller?	
11	Hur ser du på stora språkmodeller:s framtida roll inom utbildning och i ditt yrke?	Framtid
12	Finns det någon ytterligare aspekt där du använder stora språkmodeller som vi inte frågat om under denna intervju eller något mer du själv vill tillägga inom ämnet efter att ha genomfört denna intervju?	Avslutande

3.4 Bearbetning av data

3.4.1 Förberedelse av data

För att kunna dra samband och slutsatser från undersökningen var effektiv bearbetning och tolkning av den insamlade rådatan nödvändig. Inledningsvis transkriberades samtliga intervjuer med hjälp av transkriberingsverktyget Whisper. Detta underlättade och tidseffektiverade transkriberingsprocessen men krävde ändå en del handpåläggning för att säkerställa korrektheten i transkriberingarna och korrigera vissa felaktigheter. I samband med

detta omformaterades även transkriberingarna om från fri text till tabellform för att underlätta läsbarheten, i enlighet med vad Oates (2022) beskriver som en slags förberedelse av data. Därefter påbörjades dataanalysen, vilken följde en induktiv analysmetod. Recker (2013) framhäver denna analysmetod som en så kallad “bottom-up” analys där detaljerad data gradvis omvandlas till bredare kategorier och teman som kan förklara det studerade fenomenet på en övergripande nivå. Analysen inleddes således med noggrann granskning av transkriberingarna och i samband med detta inleddes en kodningsprocess av datan för att identifiera delar av texten som ansågs extra viktiga. Med detta påbörjades således processen av att sortera och reducera ner datan till meningsfull information. I kodningsprocessen var syftet således att låta eventuella återkommande mönster och teman växa fram för att utifrån dessa sedan grunda relevanta diskussioner och slutsatser.

3.4.2 Kodning och tematisering

Alvehus (2019) betonar en trestegsprocess för dataanalys enligt faserna *sortering*, *reduktion* och *argumentation*. Denna struktur har lagt grunden för hur den insamlade datan har bearbetats och strukturerats, vilket möjliggjorde en fokuserad och genomtänkt presentation av de empiriska resultaten.

I en förberedande fas organiserades varje transkription i en tabell bestående av fyra kolumner. Den första kolumnen innehåller radnummer, vilket underlättade referenshantering och återfinnelse av specifika datafragment. Den andra kolumnen härleder vem av respondenten och intervjuarna det är som talar. Kolumn tre innehåller själva transkriberingen och den tredje kolumnen användes för att anteckna kommentarer och koder, enligt metodiken beskriven av Oates (2022). Efter en första genomläsning av datan där syftet var att skapa en övergripande förståelse av materialet, initierades en inledande kodningsfas. Under denna fas tillämpades kodningskategorier föreslagna av Oates (2022). Dessa inkluderade *data utan relevans för forskningen*, *data som ansågs tillhandahålla nödvändig kontext för läsaren* och *data som ansågs direkt relevant för vår forskning*. Därtill infördes en ytterligare kategori för att markera relevanta *exempel* som givits av respondenterna. Dessa kategorier färgkodades för att tydliggöra deras skiljelinjer enligt nedanstående färgschema.

Tabell 3:3 - Kodningsschema

Kategori	Färg
Irrelevant data	Röd
Kontextuell data	Gul
Relevant data	Grön
Exempeldata	Blå

För den data som bedömdes vara relevant för forskningen utfördes en ytterligare kodningsprocess där data kopplades mot uppsatsens huvudområden, såsom *användningsområden*, *digital kompetens* och *kritisk förhållning*. Detta genomfördes i enlighet med Oates (2022) beskrivningar genom notering i den fjärde kolumnen, vilket underlättade en strukturering av data för den följande analysfasen i form av tematisering av datan. Under

denna fas grupperades och analyserades de kodade datasegmenten för att identifiera och formulera distinkta teman som återspeglades i återkommande mönster och insikter från respondenternas svar. Tematiseringen var avgörande för att syntetisera och integrera de individuella koderna till ett sammanhängande empiriavsnitt som väl återspeglade den insamlade forskningsdatan.

Genom att tillämpa Alvehus (2019) och Oates (2022) metodologiska ramar för kodning och tematisering kunde datan effektivt sorteras och reduceras till att fånga den information som var relevant för forskningen, vilket la grunden för att utveckla en argumentation baserad på det empiriska materialet.

3.5 Validitet och reliabilitet

Recker (2013) beskriver validitet som utsträckningen till vilken en mätning eller undersökning faktiskt mäter eller undersöker det som den är avsedd att mäta eller undersöka. En studie uppnår hög validitet vid det att dess mätinstrument korrekt mäter de variabler de är avsedda att undersöka, vilket säkerställer trovärdiga resultat som är relevanta för studien. Reliabilitet beskrivs istället som utsträckning till vilken en mätning konsekvent mäter det den är avsedd att mäta, vilket indikerar mätningens tillförlitlighet. För att en studie ska anses ha hög reliabilitet bör mätningarna kunna reproducera konsekventa resultat under likvärdiga förhållanden.

Att uppnå hög validitet och reliabilitet anses avgörande för forskningens kvalitet och trovärdighet. Inom ramen för den kvalitativa datainsamlingen beskriver Recker (2013) att reliabilitetsproblem vanligtvis härstammar från dåligt formulerade frågor som är oprecisa eller tvetydiga, eller från situationer där respondenterna ställs inför frågor som de inte är bekanta med, alternativt känner sig okvalificerade eller obekväma med att svara på. För att motverka detta har intervjufrågorna noggrant utformats och initialt testats på en kvalificerad utomstående part för att säkerställa frågornas tydlighet, precision och relevans för studien. Genom detta kunde potentiella problem med frågeformuleringar identifieras och justeras innan den fullständiga datainsamlingen inleddes.

Både Recker (2013) och Oates et al. (2022) framhäver att en begränsning med den kvalitativa metoden är att den bygger på en tolkningsanalys, vilket är en högst subjektiv aktivitet som är starkt beroende av forskaren själv. Ord har olika betydelser för olika personer vilket är viktigt för kvalitativa forskare att bära med sig. Författarna är också ense om att forskningen är beroende av forskarens förmåga att identifiera mönster och teman i data. I syfte att minimera subjektiva biaser i största möjliga mån har en aktiv medvetenhet om risken för subjektivitet upprätthållits under tolkningen av data. Forskningsprocessen har även dokumenterats enligt beskrivningen i kapitel 3. Detta inkluderar dokumentation av hur data samlats in, hur intervjuer genomförs samt hur kodnings- och tematiseringsprocessen utformats. Denna redogörelse syftar att öka studiens transparens och göra den efterprovbar. Vidare har en aktiv medvetenhet om risken för subjektivitet upprätthållits under tolkningen av data, i syfte att försöka motverka en sådan påverkan i största möjliga mån. I samband med databearbetningen skickades därför transkriberingarna till respondenterna för att säkerställa att inga felaktiga tolkningar gjorts och därmed kunde det valideras att studien bygger på korrekt information. Genom dessa initiativ strävar denna forskning efter att höja studiens validitet och reliabilitet och därigenom stärka forskningens övergripande tillförlitlighet och giltighet.

3.6 Etiska aspekter

Oates (2022) framhäver vissa rättigheter som deltagare i en studie innehar, vilka har beaktats i denna studie. Dessa inkluderar rätten att inte delta, rätten att dra sig ur studien, rätten att ge informerat samtycke, rätten till anonymitet och rätten till konfidentialitet. I samband med studien har alltså åtgärder tagits för att säkerställa att dessa etiska aspekter respekteras. Alla potentiella deltagares beslut att tacka nej till studien har respekterats. De deltagare som tackade ja till studien har på förhand blivit informerade om studiens syfte och vad deras deltagande kommer att spela för roll för studien, vilket gjordes skriftligen i samband med den initiala mailkontakten. Vid intervjuens start upprepades denna information och respondenterna informerades om hur intervjun kommer att struktureras och även om deras rätt att när som helst avbryta sin medverkan och därmed dra tillbaka sitt samtycke vid det att respondenten skulle ångra sig. Därtill signerade samtliga respondenter ett samtyckesavtal för att säkerställa att samtycke hade erhållits, vilket återigen upprepade studiens syfte samt vad deras deltagande innebär för studien.

Respondenternas identiteter har vidare skyddats för att säkerställa deras anonymitet. Därmed kommer personlig information som kan knytas till respondenternas identitet inte att nämnas i uppsatsen. Istället har respondenterna blivit representerade av och refererats till som en siffra mellan 1-7. All data som samlats in under intervjuerna har hanterats med försiktighet och endast använts för forskningens syften, med tillgång begränsad till endast de som ingick i forskningsgruppen. Efter att transkriberingarna var färdigställda och godkända av respondenten raderades inspelningarna av intervjuerna.

Respondenterna har också i samband med transkriberingen aidentifierats, vilket tydligt kommuniceras till vederbörande inför och i samband med intervjun för att öka tryggheten för att delta i intervjun och svara sanningsenligt på frågorna. För att säkerställa att intervjuerna genomfördes på ett konsekvent sätt följde samtliga intervjuer en på förhand fastställd intervjuguide.

4 Empiriska resultat

I följande kapitel presenteras de empiriska resultaten. Kapitlet inleds med en översikt av användningen, vilket följs av en detaljerad presentation av de viktigaste fynden från intervjuerna, vilka belyser hur stora språkmodeller integreras i lärarnas arbete samt de attityder som framkommit gentemot tekniken.

4.1 Användningen av stora språkmodeller

4.1.1 Verktyg och användningsområden

Samtliga respondenter bekräftar att de använder sig av stora språkmodeller, eller verktyg som bygger på stora språkmodeller, i sina yrkesroller. De verktyg som respondenterna beskrivit sig använda i sitt arbete inkluderar Gemini, Microsoft Copilot, GitHub Copilot, Whisper, ChatGPT-3.5, ChatGPT-4, DALL-E, Mistral och Perplexity, där ChatGPT-4 var det verktyg som var mest förekommande. Användningsområdena och graden av användning skiljer sig dock åt mellan respondenterna. I samband med genomgången av intervjudatan identifierades fler än femtio stycken olika användningsfall. Dessa har kategoriserats enligt framträdande teman för användning vilket resulterade i definitionen av sex stycken övergripande användningsområden. Nedanstående tabell presenterar en kartläggning av respondenternas användning utifrån de definierade användningsområdena. Kartläggningen inkluderar endast aktiv och avsiktlig användning och således har sporadiska tillämpningar uteslutits.

Tabell 4:1 Kartläggning av användningsområden

	R1	R2	R3	R4	R5	R6	R7
Kunskapsinhämtning	x	x	x	x	x		x
Skapa undervisningsmaterial	x	x	x	x		x	x
Analysarbete						x	x
Kreativt bollplank	x	x	x			x	
Språkhjälp & generativ förmåga		x	x	x	x	x	x
Skapa examinerande material		x	x		x	x	

Kunskapsinhämtning. Kategorin omfattar de användningstillfällen där respondenterna använder stora språkmodeller i syfte att inhämta kunskap. Exempel inkluderar beskrivningar som att respondenterna använder en språkmodell för att lära sig något, friska upp minnet eller sätta sig in i ett nytt ämnesområde.

Skapa undervisningsmaterial. Här används verktygen för att assistera i utformningen av pedagogiskt material. Exempel inkluderar allt från att be språkmodellen skapa diskussionsfrågor eller föreläsningsexempel till att skapa powerpoints eller kursplansmål.

Analysarbete. Denna kategori täcker de användningsfall där respondenterna använder språkmodeller för att assistera i analytiskt arbete. Kategorin täcker beskrivna fall av analys av både text och statistik, där språkmodellen används för att identifiera mönster eller förekomsten av specifika element.

Kreativt bollplank. I dessa fall har en språkmodell använts som en slags dialogpartner eller coach för att generera idéer och överbrygga kreativa blockeringar. Exempel inkluderar beskrivningar där stora språkmodeller uttryckligen använts som ett bollplank eller inspiration men även exempel som att modellerna använts som assistans för att starta igång en uppgift.

Språkhjälp och generativ förmåga. Kategorin innefattar assistans med språkrelaterade uppgifter såsom översättning, transkribering eller renskrivning, men även generella generativa uppgifter som att generera kod, tabeller eller diagram.

Skapa examinerande material. Slutligen har stora språkmodeller använts för att skapa examinerande material för studenterna. Respondenterna återger exempel som att de har använt språkmodeller för att skapa tentafrågor, essäfrågor, quiz eller muntliga diskussionsfrågor i examinerande syfte, men även för att skriva uppgiftsinstruktioner som case eller exempel till tentor.

Ett ytterligare användningsområde som diskuterades under flera av intervjuerna var att ta hjälp av språkmodeller för att genomföra *rättning*. Ingen av lärarna hade dock ännu implementerat detta i praktiken och därav har användningsområdet utelämnats från den övergripande kartläggningen av respondenternas användning. R3 och R7 tillhör de som har experimenterat med rättning via språkmodeller och är ense om att språkmodellen erhåller kompetens i detta avseende. De båda framhäver dock lärarens fortsatta betydelse för rättningsprocessen. R3 jämför användningen av en språkmodell för rättning med att anlita en lärarassistent och att man som examinator i slutändan oavsett har en skyldighet att attestera rättningen (bilaga 3, rad 17). På samma spår betonar R7 att tanken inte är att helt överlämna rättningen till språkmodellen utan att läraren kommer bära fortsatt ansvar för att validera och godkänna rättningen (bilaga 7, rad 24 & 56). Trots den uppmärksammade potentialen för rättning i deras preliminära försök har ingen av respondenterna ännu satt systemet i bruk.

R7: “Nej, vi måste akta oss och följa det som universitet säger så vi har bara testat om det går.” (bilaga 7, rad 71)

R3: “Jag får inte göra detta på riktigt för att det kan vara så att studentens svar uppnår verkshöjd. I de fall har studenten som upphovsrättsinnehavare rätt att kontrollera spridningen av det här verket. [...] Så det har jag experimenterat en del med och det ser mycket lovande ut, men det här kan vi inte riktigt produktionssätta förrän vi har rätt ut de legala aspekterna.” (bilaga 3, rad 17)

4.1.2 Upplevda förbättringar

Fortsättningsvis ombads samtliga respondenter beskriva ifall de upplevt några förbättringar eller förenklingar i arbetet sedan de började använda språkmodeller för arbetsrelaterade uppgifter. Respondenternas svar har enligt samma metod som ovan kategoriserats enligt framträdande teman och grupperats därefter. Nedanstående tabell presenterar en kartläggning av respondenternas upplevda yrkessamma förbättringar.

Tabell 4:2 Kartläggning av upplevda förbättringar

	R1	R2	R3	R4	R5	R6	R7
Tidsbesparing	x	x	x	x	x	x	x
Kognitiv avlastning		x	x	x	x	x	x
Ökad produktivitet		x	x	x		x	x
Kvalitetshöjning		x	x	x	x	x	
Motivationshöjning		x		x		x	

Gemensamt för respondenterna var att samtliga identifierade *tidsbesparing* som en fördel med användningen av stora språkmodeller, vilket tyder på att språkmodellerna har en förmåga att tidseffektivisera lärarnas arbetsprocesser och reducera den tid de ägnar till specifika arbetsuppgifter. *Kognitiv avlastning* var ytterligare en fördel som var starkt framträdande. Majoriteten av respondenterna beskriver att de använde språkmodellerna för att underlätta arbetsuppgifter som kräver kreativt eller analytiskt tänkande, vilket i flera fall ansågs leda till minskad mental belastning.

R2: “[...] då avlastar den en del av det här kognitiva arbetet och gör det mycket roligare och snabbare [...]” (Bilaga 2, rad 25)

R3: “Istället för att göra det så kan [ChatGPT] göra grovjobbet åt mig och det ligger ju massa tankekraft bakom det, massa kognitiva processer som jag slapp köra i min hjärna då. [...]. Du kan bespara dig själv mental energi.” (Bilaga 3, rad 23)

Ökningen av *produktivitet* innebär att respondenterna anser sig ha blivit mer produktiva i sitt arbete, det vill säga att respondenterna hinner med mer arbete jämfört med innan. R3 belyser att hen upplever ökad produktivitet som en direkt konsekvens av den kognitiva avlastning som ChatGPT och motsvarande verktyg medfört, alltså att hen orkar arbeta mer. Hen exemplifierar detta i följande citat:

R3: “Ni som studenter vet ju själva att om ni [...] arbetat med er kandidatuppsats under en väldigt lång tid så blir ni mentalt trötta. Ja, men det blir ni ju inte om ChatGPT gör allt detta arbete åt er. Så du kan spara [den mentala energin], vilket gör att du kan ha ett fyra timmar långt arbetspass där du gör massa grejer, och sedan kan du jobba i två timmar till. Det hade du inte kunnat annars, om du hade behövt sitta och vara superkoncentrerad i fyra timmar för att få saker gjorda.” (Bilaga 3, rad 23)

Vidare uttrycker en klar majoritet av respondenterna att de upplevt en *kvalitetsförbättring* i det egna arbetet till följd av användandet av stora språkmodeller. R2 beskriver att användandet av ChatGPT har bidragit positivt till kvaliteten på bland annat gruppuppgifter och quizfrågor då chatten kommer med förslag på saker som respondenten själv inte tänkt på innan (bilaga 2, rad 23-25). R3 framhåller på liknande spår att ChatGPT bidragit till en kvalitetsförbättring då det har gjort det enklare att arbeta utanför den egna kompetenszonen. R3 uttrycker vidare att hen kan vara smartare och erhålla mer kunskap än vad hen själv egentligen besitter och att hen till följd av detta då kan producera saker som hen inte kunde tidigare (bilaga 3, rad 25-27). R3 nämner också att hen tack vare ChatGPT-4 snabbare kunnat göra mer omfattande

kursutvecklingsarbete vilket blir en kvalitetsökning i sig (bilaga 3, rad 14). R4 ger som exempel att kvaliteten på undervisningen har höjts tack vare att ChatGPT hjälper till att generera mer pedagogiska exempel som hen kan använda i undervisningsmaterialet (bilaga 4, rad 20). R2 beskriver också att skapandet av tentafrågor och quiz tidigare innebar en stor och tidskrävande arbetsinsats och att hen därför tenderade att behålla samma frågor på testmoment då det krävdes för mycket arbete för att ändra dem. Sedan hen började använda sig av ChatGPT-4 beskriver respondenten att hen kunnat variera dessa testmoment betydligt mer och att hen framöver kommer att inkorporera allt fler quiz-moment i sin utbildning till följd av detta (bilaga 2, rad 23).

Slutligen framgår det av tabellen att några av respondenterna upplevt ökad *motivation* för det egna arbetet. Denna fördel är mindre framträdande än tidigare nämnda fördelar men kan vara värd att uppmärksamma som en potentiell sekundär effekt av användningen av språkmodellerna. R2 och R6 beskrev att användandet av stora språkmodeller har gjort att arbetet upplevs roligare och att de till följd av detta tenderar att fastna eller prokrastinera mindre. Därtill anser R4 att hen höjt sin egen ambitionsnivå och ger sig på arbetsuppgifter i en ny omfattning än tidigare.

R4: "Jag ger mig på arbetsuppgifter i en helt ny omfattning, jag höjer min egen ambitionsnivå för jag vet att jag kan använda den." (Bilaga 4, rad 20)

R2: "Vi hade mycket att göra, det kändes jobbigt men med detta så blev det kul." (Bilaga 2, rad 33)

R6: "På många sätt har det blivit lite roligare för jag kan fokusera på sånt som jag tycker är kul istället för... Jag fastnar mindre nu." (Bilaga 6, rad 27)

R2: "Men om jag hade känt mig lite så 'Nej, jag orkar inte idag, jag har ingen inspiration', så hade jag kanske inte gjort den grejen och då hade det kanske dröjt ytterligare en dag och sen kanske en annan grej hade kommit emellan. Men här fick ju studenterna feedback väldigt snabbt för att jag började med detta och kände att det var kul och sen så blev det klart snabbt." (Bilaga 2, rad 36).

Sammanfattningsvis visualiserar tabellen en översikt över de fördelar eller förbättringar som respondenterna anser att stora språkmodeller har medfört. Resultaten ger en tydlig indikation på att dessa verktyg har potential att bidra till framförallt tidseffektivisering och kognitiv avlastning för lärare i deras arbete.

4.1.3 Upplevda begränsningar

Samtliga respondenter beskriver sig ha stött på någon slags begränsning kopplad till användningen av stora språkmodeller. De problem som lyfts under intervjuerna kunde huvudsakligen delas in i tre övergripande kategorier: *prestandaproblematik*, *snedvridning av information* samt *etiska dilemman*.

Majoriteten av respondenterna benämner *hallucinationer* som en begränsning hos språkmodellerna (bilaga 2, rad 63; bilaga 4, rad 49; bilaga 6, rad 47; bilaga 7, rad 58). De uppger att de flertalet gånger har erhållit icke-tillfredsställande resultat av modellerna, vilket har manifesterat sig på ett antal olika sätt. R2 berättar om ett tillfälle där hen testade att ladda upp en PDF och bad om en sammanfattning, men upplevde en osäkerhet över om sammanfattningen som tillhandahölls av verktyget faktiskt blev korrekt och relevant (bilaga 2,

rad 29). R4 beskriver att modellerna inom vissa snäva områden har begränsad kunskap och att hen som expert inom det diskuterade kunskapsämnet kunnat identifiera en slags rundgång i kunskapsbanken (bilaga 4, rad 14). R4 upplever även att verktygen ibland “tappar bort sig” i den mening att verktyget gradvis tappar bort de instruktioner som gavs i början av chattkonversationen vilket resulterar i allt mer försämrade kvalitet på svaren (bilaga 4, rad 26). Vidare identifierar R4 en tydlig koppling mellan de områden där hen anser sig vara kunnig och verktygens begränsningar (bilaga 4, rad 43). Hen beskriver att inom hens expertisområde som professor, som kan anses vara relativt snävt, märker R4 att det blir en slags rundgång i verktygets kunskapsbank (bilaga 4, rad 14).

Respondent R2 och R4 uppmärksammar vidare en varierande grad av tillfredsställelse avseende prestanda som tenderar att vara *språkberoende*. R2 observerar att kvaliteten försämras avsevärt när användandet utgår ifrån det svenska språket, jämfört med det engelska.

R2: “Och jag skulle säga att jag använder den aldrig på svenska i stort sett för att den är ju inte gjord för svenska så där blir ju kvaliteten tio eller hundra gånger sämre, alltid.”
(Bilaga 2, rad 51)

R2 underbygger sitt argument ytterligare genom att framhålla att de olika språkversionerna arbetar med olika kunskapsbaser. Hen resonerar vidare om att en bred uppfattning om språkmodeller verkar vara att dessa har tillgång till all information på alla språk och sedan översätter denna information internt, men att de i själva verket utgår ifrån olika träningsbaser. Om en fråga ställs på svenska utgår modellerna från det svenska material som finns att tillgå och att detta är mer begränsat i omfång i jämförelse med den engelska uppsättningen (bilaga 2, rad 63). Även R4 uppvisar medvetenhet kring språkbaserad kvalitet:

R4: “Jag brukar fråga på engelska och be den ge svaret på svenska ibland. Jag använder det där rätt så mycket fram och tillbaka faktiskt. Men jag är väl medveten om att man får en helt annan kvalitet när man jobbar med den på engelska.” (Bilaga 4, rad 16)

Respondent R2 påpekar även att ChatGPT-4 saknar tillgång till data i realtid och att det därmed finns en viss *eftersläpning av information* i modellen. Hen ser en klar begränsning i att det inte är möjligt att ställa frågor om nyare fenomen eller inom ämnen där kunskapen har uppdaterats. Då R2 är aktiv inom ett teknikrelaterat ämne blir detta extra påtagligt i och med att utvecklingen går fort fram och information snabbt blir förlegad. Hen belyser dock skillnader mellan verktyg och framhåller att till exempel Microsoft Copilot har möjligheten att göra webbsökningar vilket minimerar risken för inaktuell information. Dock vittnar R2 om att hen gjort vissa tester som visat att viss dagsaktuell information inte finns att tillgå även för detta verktyg (bilaga 2, rad 63).

Under intervjun tar R6 upp begränsningar som går att urskilja som ren *prestandaproblematik*. Hen beskriver det enligt följande citat:

R6: “[...] Dels finns det några generella begränsningar kopplat till längd på texter. Ju längre uppgiften ChatGPT får, desto tröttare blir den och den presterar ju sämre. [...] Det är från dag till dag där exakt samma uppgift klarar den galant en dag och andra dagen så går det inte över huvud taget. Och sen väntar man tre dagar sen funkar det igen, alltså exakt samma prompt och exakt samma material.” (Bilaga 6, rad 45)

Vidare har R5 observerat att verktygen tenderar att generera ett svar som hen bedömer inte är i linje med det hen frågar om. Ett exempel som lyfts är när respondenten bett om språkförbättringar till ett utkast av en vetenskaplig artikel och att verktyget då genererat text

av sammanfattande karaktär i punktform vilket respondenten framhåller inte var vad som önskades (bilaga 5, rad 18; bilaga 5, rad 56).

Respondenter R1 och R7 nämner *snedvridning av information och partiskhet* som en begränsning de har upplevt i samband med användning av stora språkmodeller. R1 kopplar detta fenomen till de olika språkbaserade träningsdata som tidigare nämnts och betonar att användningen av dessa modeller inte medför något kunskapsutbyte eller nyansering. Istället bekräftar informationen som delas endast befintliga perspektiv inom respektive språk verktygen används på, vilket skapar en rundgång. Hen är av uppfattningen att detta resulterar i en slags siloeffekt över informationen där ekokammare bildas, vilket bidrar till ökad polarisering (bilaga 1, rad 84). R7 nämnde ett test som hen fått återberättat för sig där en språkmodell ombads att generera illustrationer på två motsatta koncept som kan tolkas som neutrala. Resultatet visade på en försköning av det ena konceptet medan det andra förfulats, vilket speglade en viss politisk hållning (bilaga 7, rad 56).

Flertalet av respondenterna belyser slutligen några *etiska dilemman* de stöter på i samband med användningen av stora språkmodeller. Majoriteten av respondenterna nämner att känslig data som delas med en stor språkmodell riskerar att äventyras och att detta utgör ett betydande hinder för användningen. R7 betonar denna begränsning som särskilt påtaglig i sitt arbete, där hantering av mycket känslig data är central (bilaga 7, rad 30). Både R4 och R6 reflekterar över en osäkerhet kring användningen av den data som delas med verktygen, eftersom datan sedan ägs av de storföretag som kontrollerar verktygen och fritt kan använda den för egna intressen (bilaga 4, rad 41; bilaga 6, rad 51). R3 belyser problematiken ur ett juridiskt perspektiv, särskilt i sammanhang där studentarbeten hanteras. På vissa tentor med essäfrågor kan studenters svar uppnå verkshöjd och är då skyddade av upphovsrätten. En eventuell delning av dessa texter från R3:s sida skulle därmed betraktas som en olovlig spridning av ett upphovsrättsskyddat verk (bilaga 3, rad 15). Av respondenternas svar framgår det att begränsningen inte gäller för alla verktyg som används av de som intervjuats. R2 och R5 använder Microsoft Copilot, som tillhandahåller en funktion där verktyget inte lagrar någon av den information som delas av användaren (bilaga 2, rad 38; bilaga 5, rad 66), vilket belyser att problematiken varierar beroende på det valda verktyget. R7 tar även upp illvilliga avsikter som en begränsning. Hen oroar sig för att dessa verktyg även finns tillgängliga för individer med onda agendor och att detta medför ett etiskt dilemma vid användandet av en stor språkmodell (bilaga 7, rad 48).

4.2 Prompt engineering

När tillfrågade om hur de arbetar med prompt engineering var svaren blandade från respondenterna. Samtliga respondenter visade på en medvetenhet om prompt engineering som koncept, men det fanns variationer i nivå av aktivt arbete med tekniken. Vissa respondenter beskrev sig endast använda enklare prompts, medan andra framhävde prompting som en av de viktigaste byggstenarna i samband med användning av språkmodeller. Varken R1, R2, R5 eller R6 anser sig arbeta aktivt med prompting i någon vidare större utsträckning i sitt arbete.

R1: "Jag har ju inte så stor erfarenhet men jag har ju funderat på det. [...] Sen har jag även bett den så "give me an overview..." eller "short overview..." eller "give me a quick introduction to..." , eller så för att själv snabbt komma in i ett ämne. Jag har inte grävt i det supermycket men däremot har jag lärt mig att sekvensera. [...] förbered dig på

att ge svar i flera sekvenser [...] och då kan den ju ge längre svar. Men inget jag jobbat supermycket med.” (Bilaga 1, rad 28-30).

R5: “Ja, oftast inte med någon vidare plan. Utan att jag bedömer hur bra svar jag fick och vad det brister i och vad jag vill ha mer av snarare än att jag från början tänker att nu måste jag börja med det här enkla för sen går det lite svårare.” (Bilaga 5, rad 44)

R6: “Jag har studenter som ger den en roll, det kan man ju göra men jag brukar inte skriva rollbeskrivningar, det har jag inte gjort. Jag har nog varit mer generell. Alltså det allra mesta är till proofreading. Och då är det bara korta prompts. Proofread this, translate this, make the text more fluent. Ganska konkreta, tydliga uppmaningar kopplat till text.” (Bilaga 6, rad 39)

Däremot redogör R6 att hen i sitt forskningsprojekt i motsats till ovanstående citat de facto har arbetat mycket med prompting, där de testat olika variationer av formuleringar för att undersöka vilka slutresultat som genereras av specifika typer av instruktioner. Detta beskrivs vara särskilt kopplat till textanalys (bilaga 6, rad 37). R2 arbetar inte heller med prompting i någon större utsträckning och motiverar sitt förhållningssätt med skepsis mot framtiden av prompting, då hen inte tror att det kommer att krävas i framtida modeller och att bristen på framtidsutsikter gör hen omotiverad att sätta sig in i ämnet (bilaga 2, rad 49).

I motsats arbetar R3, R4 och R7 i mycket hög utsträckning med systematisk prompting. R4 talar om prompting som en konst som utgör ett nytt kompetensområde och är en typ av skicklighet (bilaga 4, rad 20 & 26). Hen beskriver vidare sitt arbete med prompting som aktivt i den bemärkelse att hen förstår att vid de tillfällen önskat resultat inte uppnås är det en konsekvens av bristande instruktioner och inte bristande prestanda hos språkmodellen (bilaga 4, rad 26). Hen beskriver sitt arbete med prompting vidare:

R4: “Ibland använder jag AI för att be den ge mig promptar. [...] Sedan jobbar jag ju aktivt med att jag startar om en ‘kedja’, jag startar om en chatt med den för att undvika att den tappar bort sig. [...] Sedan jobbar jag ofta aktivt med att tala om för den vad jag inte vill ha.” (Bilaga 4, rad 26)

R7:s arbete med prompting är av systematiskt karaktär, vilket hen utforskar närmare i samband med konstruktionen av prompts inom ramen för sin forskning. Respondentens forskningsgrupp har utvecklat en prompt där de ber språkmodellen att analysera en transkribering i syfte att identifiera, analysera och kategorisera innehållet. De instruerar verktyget att identifiera avsnitt i transkriberingen som behandlar ett specifikt tema och sedan klassificera informationen efter typ enligt de definitioner som hens team tillhandahåller språkmodellen. I prompten framgår det även att verktyget ska förklara sin bearbetningsprocess för att på så vis möjliggöra för R7 att validera dess arbetsmetod (se bilaga 7, rad 18). Hen instämmer med R4 i att det inte är språkmodellens begränsningar som är grundorsaken till att användaren erhåller ett oönskat svar från modellerna. Hen betonar att det krävs logik i konstruktionen av en effektiv prompt och att ett misslyckande i detta avseende leder till missförstånd från verktygets sida (bilaga 7, rad 38). Hen understryker detta resonemang ytterligare:

R7: “Alla svar finns, det är bara att de som kommer långt är de som ställer rätt frågor. Så det här är ju en teknik i sig.” (Bilaga 7, rad 36)

R3 resonerar på samma sätt som R4 och R7 gällande promptens roll i att extrahera den önskade informationen från en språkmodell då hen vid upprepade gånger understryker att

utmaningar om att få fram önskad information är kopplade till användarens förmåga att kunna prompta effektivt (bilaga 3, rad 14). Hen har även ett bibliotek av längre prompts, där samtliga exemplar i samlingen är utformade efter ett specifikt syfte (bilaga 3, rad 35). Hen pekar ut promptingtekniker som *change-of-thought-prompting*, *few-shot-prompting*, *role-prompting* och att hen specificerar vilket format svaret ska utformas utifrån (bilaga 3, rad 37). R3 beskriver att det ofta är ett iterativt arbete för att komma underfund med modellen hen arbetar med, för att få insikter om hur den helst vill bli promptad (bilaga 3, rad 58). I en specifik kontext inför en uppgift beskriver hen sitt arbete med dessa komplexa prompts uppbyggd av flertalet tekniker enligt följande:

R3: "Det är en extremt lång prompt faktiskt. Men du har också ett väldigt långt kontext-window, så det funkar. Jag säger till den att den är en "grading assistant" och ger den instruktioner för hur den ska poängsätta. Det blir alltså en prompt per uppgift då. Och så kan jag köra [prompten] om och om igen. Den får även instruktioner för hur den ska betygsätta, alltså "rubrics" ungefär, och så får den ett [tenta]svar då. Det här kan ju vara en jäkla massa kod eller någon utskrift, en SQL query eller liknande."
(Bilaga 3, rad 21)

Även de mer avancerade användarna inom prompt engineering uppger däremot att de inte använder sig av avancerade prompts när det kommer till enklare typer av uppgifter, som att korrigera formatering av kod, uppfriska minnet inom ett visst ämne eller för att snabbt inhämta information (bilaga 3, rad 37; bilaga 4, rad 12; bilaga 7, rad 16).

4.3 Digital kompetens

4.3.1 Lärares syn på digital kompetens

Synen på digital kompetens var relativt enhetlig bland respondenterna. De flesta ansåg att det är viktigt för lärare att inneha digital kompetens (bilaga 2, rad 55; bilaga 4, rad 34; bilaga 7, rad 42). R3 understryker vikten av att lärare håller sig uppdaterade kring teknologiska framsteg och lyfter fram språkmodeller som exempel. R6 belyser de betydande effekter som dessa haft på examinering, undervisning och universitetsrelaterade aktiviteter, vilket gör det avgörande för lärare att förstå och anpassa sig till nya verktyg och metoder för att undvika en klyfta i förståelse mellan studenter och lärare (bilaga 6, rad 41). R3 betonar vikten av digital kompetens i relation till en ökad digitalisering i samhället:

R3: "Men sedan är ju ens digitala kompetens i allmänhet viktig att utveckla oavsett vilken institution man sitter på, i och med att allt blir mer och mer digitaliserat i högre utbildning." (Bilaga 3, rad 39)

Sammantaget är lärarna alltså överens om synen på digital kompetens som något centralt i deras yrkesroll i dagens samhälle. Digital kompetens tycks inte enbart betraktas som en teknisk färdighet, utan som en nödvändighet för att effektivt kunna navigera i en alltmer digitaliserad utbildningsmiljö. Detta innefattar inte bara utvecklingen av den egna kompetensen utan även bidraget till institutionens övergripande funktioner genom att stödja såväl kollegor som studenter. R5 nyanserar bilden något. Hen håller med om att digital kompetens kan anses viktigt för att lärare i viss mån måste kunna möta sina elever i vart de är i sina färdigheter. Däremot är hen inte övertygad om att lärare behöver vara särskilt digitalt kompetenta ur perspektivet av deras egen undervisning och att det går att genomföra bra

undervisning även utan digitala verktyg (bilaga 5, rad 46). R5 framhåller att det inte alltid bör vara ett krav att följa normen i den bemärkelsen (bilaga 5, rad 50).

Både R5, R6 och R7 exemplifierar kopplingen mellan digital kompetens och stora språkmodeller. Både R5 och R6 ser ett samband i att de kollegor som tenderar att hålla sig ajour med den digitala utvecklingen och är mest teknikpositiva också är de som utforskat språkmodellerna mest och därigenom hittat mest intressanta användningsområden (bilaga 5, rad 84; bilaga 7, rad 46), medan de som historiskt sett använt digitala verktyg i mindre utsträckning knappt är medvetna om språkmodellernas existens (bilaga 6, rad 46). Även R2 bekräftar sambandet mellan digital kompetens och användning av stora språkmodeller, men tillför en motvikt genom att betona att det är omöjligt att alltid ligga i teknisk framkant. R2 betonar att även om det är viktigt att hålla sig uppdaterad om AI så är det orealistiskt att försöka ligga i framkant med alla specifika program och tekniker på grund av det omfattande och tidskrävande arbete det skulle innebära (bilaga 2, rad 55).

4.3.2 Utbildning

Ingen av respondenterna svarade att de hade fått någon form av utbildning i hur man kan använda sig av stora språkmodeller i sitt arbete på initiativ från universitetsledningen. Vissa hade på eget initiativ fått en kortare genomgång eller ytlig presentation av andra kollegor, men ingen av de sju respondenterna säger sig ha fått ordentlig utbildning från universitetsledningens håll, mer än att de mottagit vissa riktlinjer för hur de bör förhålla sig till språkmodellerna.

R5: "Nej, jag har bara fått lite kollegiala tips." (Bilaga 5, rad 26).

R2: "I alla fall, nej, vi har inte fått något mer än hur vi ska förhålla oss till studenterna och deras användning. [...] För lärarnas användning har liksom inte varit i fokus, men det kommer säkert framöver." (Bilaga 2, rad 43-45).

R4: "Alltså jag var en super early adopter av det här. [...] Så under en period där föreläste jag på universitetet om det här, där den enda kompetens jag hade var att jag var en early adopter. Nu då så får jag istället internutbildning av de kollegor som jag visade det här för en gång i tiden." (Bilaga 4, rad 22-24).

I samband med diskussionen kring en generell inställning till utbildning om stora språkmodeller ställer sig R1 och R2 kritiska.

R1: "Vi får väldigt lite utbildning i saker och ting, formellt[...]. Jag tillhör också de som tycker att en hel del av det där är svårt att utbilda sig inom. Jag tror det handlar mycket om intressen, jag tror inte att du kan gå en kurs i hur man jobbar med [ChatGPT][...]." (Bilaga 1, rad 42).

R2: "[...]det är för låg nivå och jag tror inte det är bestående över tid, alltså ens på ett års sikt den kunskapen. Det tar jag helst inte i. Man kan diskutera lite om vad prompt engineering är och så där, konceptuellt. Men inte hantverket. Däremot i programmeringen är det ett speciellt fall. Där måste man vara med. Programmering är ju skills-baserad." (Bilaga 2, rad 40).

Även R3 argumenterar för att behovet av att använda verktyget kommer att vara individuellt, och att det därför inte finns någon klar och samlad väg framåt för detta.

R3: “[...] man kan inte sitta nere på stan och stifta massa regler som folk ska följa. Om du är vicerektor för utbildning på ett universitet kan du inte säga hur jag ska göra med min kurs.” (Bilaga 3, 29).

R5 invänder mot denna skepticism och hävdar i motsats att det är just en gemensam väg framåt vad gäller integreringen av dessa verktyg i undervisningen och arbetet som krävs.

R5: “Det är just nu upp till kurserna och det är upp till specifika lärare, vad man får göra och inte få göra. Och ganska vagt varför man får göra någonting men inte något annat. Så ja, jag tänker att vi kommer vara så illa tvungna att hitta på ett vettigare sätt att integrera det.” (Bilaga 5, rad 55).

På liknande vis argumenterar R3, R5, R7 för att lärare har ett pedagogiskt ansvar i att utbilda studenter i användning av dessa verktyg, vilket speglar en allmän positiv inställning till utbildning inom stora språkmodeller. R3 uttrycker att utbildning i tekniker som prompting och förståelse för den grundläggande arkitekturen bakom stora språkmodeller kan resultera i ansvarsfullare användning och minskad risk för hallucinationer (bilaga 3, rad 55). R7 har tillsammans med kollegor redan inkorporerat saker som prompting i kursmaterialet, vilket hen ser som en skyldighet som lärare (bilaga 7, rad 38-40). R5 har gjort en vana av att använda Copilot i klassrummet tillsammans med sina studenter. R5 använder verktyget på sin dator via storskärmen samtidigt som studenterna uppmanas göra detsamma på sina egna datorer. Därefter jämförs resultaten gemensamt för att se vilka olika svar de fått i syfte att främja en djupare förståelse och kritiskt tänkande kring verktygets kapacitet (bilaga 5, rad 12).

4.3.3 Kritisk förhållning

Samtliga respondenter uppvisade en aktiv medvetenhet om att de språkmodeller de använder har vissa begränsningar och att användandet av dem därmed inte är helt riskfritt. Respondenterna hade olika metoder för att kritiskt förhålla sig till språkmodellerna. Flera tog proaktiva åtgärder som att aktivt välja att inte använda det till ändamål där risker förelåg, medan andra antog ett mer reaktivt förhållningssätt och vidtog kritiska åtgärder under eller efter användningen. Flera av respondenterna använde en kombination av proaktiva och reaktiva kritiska förhållningssätt. De anpassningar som identifierats har placerats in under gemensamma teman *validering*, *anpassning av användningsområden* och *etiska anpassningar*.

En majoritet av respondenterna beskrev att de endast använde språkmodellerna för ämnen inom deras egna kunskapsområde och att de i och med detta själva var kapabla att *validera* riktigheten i texten. Genom denna strategi ansåg de sig kunna säkerställa att språkmodellens svar var korrekta och identifiera eventuella felaktigheter.

R1: “Jag brukar inte använda den till sånt jag inte begriper än. [...] Ja, jag skulle inte våga lite på den om jag inte kan något.” (Bilaga 1, rad 54-58)

R4: “Det bygger ju på att man har väldigt mycket innehållskunskap, att man själv kan det som man frågar om. [...] för den som inte har grundkunskaper så kommer det här att vara ett begränsat verktyg, skulle jag säga. Svaret på er fråga är alltså, genom att ha ordentliga solida domänkunskaper som jag kan relatera till den output som ChatGPT ger.” (Bilaga 4, rad 49)

R5: “Jag tycker nog att jag bara använder det på sånt som jag själv kan ganska bra. Åtminstone i jobbet. Och då har jag omdömet själv att stryka det som jag tycker är mindre bra och ersätta det med eget eller ifrågasätta den och sedan utveckla den.” (Bilaga 5, rad 50)

På liknande vis beskriver flera av respondenterna att de tar ansvar för att *validera* den information som genereras av språkmodellerna genom att själva granska och kontrollera informationen mot externa källor för att säkerställa informationens korrekthet.

R4: “Och sen gör jag stötvisa stickprovskontroller på de referenserna eller så googlar jag för att se att det här är någonting som stämmer. Vanligen får man ju någon slags aha-upplevelse, att ja, just det. Det här känner jag igen.” (Bilaga 4, rad 49)

R6: “Jag brukar tänka på den ungefär som en forskningsassistent. Någon som kan hjälpa en med saker, men som jag inte litar på till hundra procent utan att ansvaret ligger hos mig. Så jag behöver ändå validera det på något sätt eller kolla igenom svaret.” (Bilaga 6, rad 49)

R7: “Ja, alltså för det första så får den aldrig använda kreativitet och sen dubbelkollar jag. Jag menar, man måste ju validera allting. [...] Så man kan ju inte helt lita på den.” (Bilaga 7, rad 56)

R3: “Sedan är det också så att mycket av vad jag sitter och pratar med ChatGPT om är kod-, programmerings- och teknikrelaterat. [...] kodning går alltid att verifiera. Jag kan köra det, jag kan debugga den, jag kan kolla på dokumentationen, jag kan googla lite. Få en “second opinion.” (Bilaga 3, rad 60-62)

R1 beskriver att hen minimerar språkmodellens roll genom att betrakta den som en av många hjälpmedel, men inte den huvudsakliga källan som hen förlitar sig på (Bilaga 1, rad 18). R3 förhåller sig till språkmodellen på liknande vis genom att använda den som en komponent för att generera idéer som hen sorterar bland och sedan utforskar och *validerar* genom att söka ytterligare vetenskapligt stöd från andra källor. Med detta understryks språkmodellens roll som ett av många hjälpmedel snarare än som den enda primära informationskällan (Bilaga 3, rad 62).

Ett ytterligare tema som framkommer från intervjuerna gällande respondenternas kritiska förhållningssätt är en medveten *anpassning av användningsområden* för språkmodellen, vilka är direkt kopplade till de begränsningar som respondenterna upplever. Exempelvis tillåter R7 inte språkmodeller att assistera i någon form av kreativt arbete (bilaga 7, rad 56), medan R2 säger sig undvika att använda språkmodellerna för att inhämta ren fakta. Istället föredrar hen att använda Google eller liknande tjänster för den sortens ändamål (bilaga 2, rad 65). Samtidigt beskriver dock R2 att hen de facto använder språkmodeller för just detta syfte, men att detta endast görs för faktainhämtning inom just ämnet AI. Detta förklarar hen med att hen fullt förlitar sig på att språkmodellen saknar kunskapsluckor inom detta specifika ämnet och hen därmed inte ser någon risk för hallucinationer (bilaga 2, rad 27). Vidare väljer både R2 och R4 att medvetet instruera språkmodellerna på engelska till följd av de kapacitetsbrister de upplever att systemet har på svenska (bilaga 2, rad 51, 63). R2 uppvisar också en medvetenhet om att ChatGPT inte är tränad på fullt uppdaterad data och undviker därför att använda språkmodellen i uppgifter som kräver tidskänslig information (bilaga 2, rad 63).

Respondenterna belyser en rad etiska och praktiska överväganden som de hanterar vid användningen av stora språkmodeller och flera av dem vittnar därtill om att de vidtar särskilda *etiska Anpassningar* för att navigera dessa utmaningar. Flera betonar vikten av att respektera

och skydda individers integritet. R5 understryker vikten av ansvarsfull användning för att säkerställa att ingen lider skada (bilaga 7, rad 54) och R6 framhåller forskarens plikt att skydda sina forskningsobjekts integritet (bilaga 6, rad 53). R5, R6 och R7 anser att man av etiska skäl bör undvika att använda språkmodellerna för analys av material som innehåller känsliga personuppgifter (bilaga 5, rad 66; bilaga 6, rad 51; bilaga 7, rad 52). Av samma anledning avstår R4 från att skicka upp studentarbeten till ChatGPT (bilaga 4, rad 41). R4 beskriver att vissa datalagringsregleringar stoppar hen från att använda språkmodellerna även för andra uppgifter vilket hen beskriver är ett hinder mot att nyttja verktygens fulla potential (bilaga 4, rad 51).

R1 har ännu inte haft ett behov av att använda stora språkmodeller till en uppgift som hanterat denna sortens känslig information (bilaga 1, rad 66). R2 är den enda som svarar att hen inte tänker på frågor om etik och integritet i samband med användningen av stora språkmodeller (bilaga 2, rad 67) trots att hen i ett tidigare skede beskriver att hen aktivt väljer att använda sig av Microsoft Copilot framför exempelvis ChatGPT av anledningen att Copilot har en skyddad version som förhindrar att data lagras i något IOS-moln (bilaga 2, rad 38), vilket ses som en etisk anpassning. Även R5 lyfter fram fördelarna med Microsoft Copilot i sammanhanget av etik och integritet.

R5: "Och i den meningen så känns det ju bra att jag inte behöver bekymra mig så mycket om vad jag skriver där." (Bilaga 5, rad 66)

R3 avstår från att använda stora språkmodeller för tentarättning av eventuella upphovsrättsliga skäl (bilaga 3, rad 15). R3 har dock experimenterat med tentarättning via en modell som körs lokalt på datorn istället för via en webbläsare och anser att detta då inte borde bryta mot någon upphovsrätt och därmed vara en potentiell lösning på problemet. Hen understryker dock att detta inte är något hen vågar produktionssätta förens de legala aspekterna är helt utredda (bilaga 3, rad 17). R6 föreslår också lokalt körda språkmodeller som ett sätt att komma runt de etiska aspekterna av dataintegritet (bilaga 6, rad 51).

Det finns slutligen exempel då respondenterna *inte uppvisar ansvarsfull användning*. R4 nämner att hen en gång använde sig av en språkmodell för att omformatera studenters personuppgifter och uttrycker en medvetenhet om att detta eventuellt inte var etiskt korrekt. (bilaga 4, rad 12). R5 förklarar att hen ibland har frågat ChatGPT om ett ämne där hen har uppfattar sig ha viss faktakänedom men inte är helt säker. När språkmodellen svarade i linje med hens antagande, utgick R5 ifrån att svaret var korrekt men verifierade varken sin egen kunskap eller språkmodellens svar med en pålitlig källa (bilaga 5, rad 64).

4.4 Stora språkmodellens potential inom utbildning

Ingen av respondenterna ställde sig enbart negativa till att använda stora språkmodeller i deras arbete, men till vilken grad av positivism de såg på verktyget med varierade. De flesta av respondenterna var överens om att oavsett effekt och utbredning av användning så är dessa verktyg en oundviklig och självklar teknik i vårt framtida samhälle och utbildningsväsen. Både R1 och R2 likställer språkmodellerna med internet och framhäver med detta att de kommer bli självklara verktyg i framtiden (bilaga 1, rad 68; bilaga 2, rad 78).

R7: "Jag tror det kommer bli som el. Alla behöver det. Om man inte använder det så är man hopplöst bara efter." (Bilaga 7, rad 62)

R1: “Ja, det går inte att motarbeta det. Det är bara att bejaka det. Men, i vilken grad vi är rustade att göra det vet jag inte riktigt.” (bilaga 1, rad 71)

Samtliga respondenter betonade ett beroende mellan användningspotential mot den utsträckning verktygen utvecklas i framtiden. De upplevde att det var svårt att säga vad språkmodellerna kommer att ta för roll i ett framtida utbildningsväsen just med tanke på att det är omöjligt att veta vad kommande versioner kommer att vara kapabla till eller hur kraftfulla de kommer att vara.

R3: “Ingen aning. Det är väldigt svårt att säga för att man vet aldrig hur pass kraftfulla de nya modellerna kommer att vara. För det är ett sådant riktigt ‘arms race’ att utveckla de mest kraftfulla modellerna. Man vet aldrig hur GPT-5 eller GPT-6 kommer att se ut.” (Bilaga 3, rad 64)

R2: “Till exempel om GPT-5 kommer så kanske det blir så ‘Nej, nu behöver du inte prompta längre!’” (Bilaga 2, rad 49)

Flera nämner dock framtida potentiella användningsområden. R4 såg potentialen att kunna införa utbildningsmoment i större kurser som hen tidigare enbart hade kunnat genomföra i mindre skala. I en av sina kurser kategoriserar R4 elevsvar för att urskilja mönster i hur studenter resonerar och identifiera eventuella kunskapsluckor. R4 har nu börjat leka med tanken på att ta hjälp av stora språkmodeller för denna sortens uppgift och ser med det potential för ökad skalbarhet genom att kunna tillåta fler studenter i kurserna (Bilaga 4, rad 20).

R3 och R7 har istället börjat utforska möjligheterna att implementera stora språkmodeller i rättningsprocessen av tentor. R3 ser potentialen att tidseffektivisera sitt rättningsarbete genom att låta en stor språkmodell göra rättningen som hen sedan går igenom för att attestera. Läraren upplever att de tester hen gjort har gett tillfredsställande resultat, och att det är legala aspekter som först måste redas ut innan en sådan lösning kan tas i bruk, både med avseende av upphovsrättsliga skäl gentemot studenters essäsvår, men även av integritetsskäl. Hen påpekar dock att arbetet med en språkmodell lokalt hade kunnat vara en möjlig väg framåt (Bilaga 3, rad 17). R7 lyfter istället fram fördelar i form av konsekvent agerande och objektivitet, i kontrast till den mänskliga bedömarens subjektivitet och varierande prestation. Hen uttrycker att en språkmodell konsekvent kommer att tillämpa samma regler och standarder utan att påverkas av emotionella tillstånd eller fysiska svagheter såsom trötthet eller irritation, som kan leda till inkonsekvens i bedömningar. Med detta framhäver R7 en potential för språkmodeller att på sikt kunna användas för att ge mer rättvisa bedömningar (Bilaga 7, rad 24 & 73).

R5 och R6 ställer sig i motsats båda skeptiska till idén om att kunna använda stora språkmodeller till rättning.

R6: “Jag har inte använt den till att ge feedback till studenter, det har jag nog inte använt den till, för jag tycker att det oftast inte är så bra. Alltså den hittar alldeles för generella grejer. [...] Det blir tomta, meningslösa saker för det mesta tycker jag. [...] Det är väldigt specifikt och väldigt kontextbundet och kopplat till utbildningen, vad kraven och målen är i utbildningen. Då tycker jag inte att den klarar det.” (Bilaga 6, rad 47).

R5: “Alltså jag tänker så här att det där borde studenterna göra själva innan de lämnar in. Det är helt enkelt inte myndighetens uppgift att ersätta vår bedömning med någon mekaniserad process som vem som helst kan göra. Vår yrkesetik är att använda vår erfarenhet och kunskap om hur den här uppgiften löses på ett bra sätt och dela med oss

av den och investera den tid det tar att skriva ett omdöme. Eller åtminstone göra en bedömning och kortfattat ge någon... att den är baserad på vår egen tysta kunskap egentligen.” (Bilaga 5, rad 74)

R1, som uppvisat en rätt begränsad användning av stora språkmodeller, ser framtida potential att använda stora språkmodeller när det kommer till hens egen forskning.

R1: “Jag har faktiskt inte haft något forskningsprojekt sedan ChatGPT kom. På sista året har jag varit helt sänkt i undervisning. Men vi ska dra igång ett nytt forskningsprojekt här nu kommande månaden så det är möjligt att vi kommer att använda det då, men jag har inte kommit dit än.” (Bilaga 1, rad 16)

Respondenterna ger en rätt samlad bild vad gäller lärarens framtida roll inom utbildning, vilken av flera anses behöva omvärderas i takt med teknikutvecklingen och språkmodeller i synnerhet. Flertalet av respondenterna belyser att språkmodellerna är kapabla att hantera uppgifter som traditionellt legat inom lärarens arbetsbeskrivning och att det nu blir allt viktigare för lärare att kunna komplettera dessa verktyg genom att bidra med unik mänsklig kompetens och insikter som modellerna saknar. Detta betonar behovet av att lärare aktivt formar sin roll för att förstärka mänskliga värden och kritiskt tänkande inom undervisning.

R4: “Den lärare som kan ersättas av en dator bör så göras. Hans poäng var naturligtvis att om en lärare inte kan bidra med något mer än vad datorn kan göra, då bör man ersätta den.” (Bilaga 4, rad 57)

R6: “Nej, det är svårt att se att det kan ha något annat än en ganska revolutionerande roll, tänker jag. De kommer ju såklart att bli allt mer effektiva och göra allt större delar av det vi jobbar med. För man ser att det ändå klarar av komplexa analyser på liknande sätt som vi..” (Bilaga 6, rad 55)

R5 erkänner att det nog tekniskt sett snart är möjligt att ersätta lärare, men att hen inte ser någon anledning att oroa sig då hen inte tror att folk kommer vilja sluta lära sig via undervisning i klassrum med klasskamrater och lärare (bilaga 5, rad 76). Andra höll förvisso med om att verktygen förmodligen var här för att stanna, men att de trots sin användbarhet inte kommer att förändra grunden i utbildningen.

R2: “Och då kommer användandet [av ChatGPT] att öka, men det kanske inte förändrar i grunden så mycket. Mer än att det här kommer att vara ett självklart verktyg.” (Bilaga 2, rad 80)

R1: “ChatGPT för mig är ett av många av de verktyg som finns, som är del av den digitala... [uppsättningen]. Sedan råkar detta vara lite mer fancy kanske, lite mer sådär futuristiskt, men för mig är det ett verktyg som många andra.” (Bilaga 1, rad 50)

R6 betonar behovet av att etablera tydligare riktlinjer för användandet av språkmodellerna:

R6: “Det är just nu upp till kurserna och det är upp till specifika lärare, vad man får göra och inte få göra. Och ganska vagt varför man får göra någonting men inte något annat. Så ja, jag tänker att vi kommer vara så illa tvungna att hitta på ett vettigare sätt att integrera det” (Bilaga 6, rad 55)

5 Diskussion

I följande kapitel diskuteras de empiriska resultat som erhållits från intervjuerna. Diskussionen är förankrad i både befintlig litteratur och de undersökningsfaktorer som introducerades i det teoretiska resultatet, samt i egna tolkningar.

5.1 Användningen av stora språkmodeller

5.1.1 Användningsområden

Det framgår ur empirin att lärares användning av språkmodeller är högst individuell. Spridningen i användningsområden uppvisar en mångfacetterad och individbaserad användning där vissa uppvisar en viss försiktighet för användandet medan andra anammar tekniken mer öppet och aktivt experimenterar med dess möjligheter. Tillsammans bekräftade samtliga lärare vad Kasneci et al. (2023) beskriver att stora språkmodeller har potentialen att användas till. Sammantaget tycks spridningen återspegla individuella preferenser snarare än institutionella eller ämnesspecifika samband. Utifrån respondenternas svar tolkas användningen i hög grad vara beroende av personligt intresse samt vilket specifikt stöd eller avlastning lärarna känner behov av i sina yrkesroller. Vidare indikerar empirin även att det finns en utforskad potential där lärare ännu inte fullt ut har realiserat alla användningsmöjligheter som språkmodellerna erbjuder. Detta kan bero på att lärare i hög grad utgår från egna behov och förblir omedvetna om andra sätt som yrkeskollegor implementerar dessa verktyg i deras arbete. Det finns inget som indikerar att en samlad användning är nödvändig, men om en sådan skulle vilja uppnås framträder ur resultaten ett behov av att lärare får gemensam utbildning och stöttning i hur de effektivt kan utnyttja stora språkmodeller för att maximera verktygens potential i yrkessammanhang. Ett starkare kollegialt utbyte verkar också framgå som en möjlig väg att gå, vilket diskuteras mer djupgående i avsnitt 5.3.1.

Spridningen tycks, förutom personliga preferenser om användningsområden, även att variera som följd av att lärarna har skilda uppfattningar om vilka områden språkmodellerna presterar sämre och bättre i. Detta kan möjligtvis förklaras av att respondenterna har varierande nivå av skicklighet gällande att formulera prompts som resulterar i önskvärd output (vilket berörs i avsnitt 5.2), eller att lärarnas krav på vad som utgör kvalitet helt enkelt skiljer sig åt och att de därmed har olika uppfattningar om till vilken grad språkmodellerna lyckas uppfylla dessa kvalitetskrav.

5.1.2 Upplevda förbättringar

Samtliga av respondenterna vittnar om någon form av upplevd förbättring i det egna arbetet till följd av användningen av språkmodeller, vilka har kategoriserats enligt fem framträdande teman: tidsbesparing, kognitiv avlastning, ökad produktivitet, kvalitetshöjning och motivationshöjning. Det går att urskilja en röd tråd bland de upplevda förbättringarna där de nämnda fördelarna är sammanvävda och kan förstärka varandra i en slags positiv spiral. Exempelvis vittnar samtliga respondenter om att språkmodellerna medfört tidsbesparingar. Genom att minska den tid som krävs för att slutföra en viss arbetsuppgift frigörs tid för

lärarna att ta sig an flera uppgifter eller vidare fördjupa sig i befintliga uppgifter, vilket möjliggör att de kan åstadkomma mer under samma tid. Följaktligen ger detta potential för en ökning av produktivitet, alternativt en potentiell kvalitetsökning då tid frigörs för lärarna att fördjupa sig i uppgiften. Ökningen av produktivitet kan sedan på liknande vis anses sammankopplad med kognitiv avlastning, vilket R3 uttryckligen beskriver. Genom att låta språkmodellerna avlasta mentalt krävande arbetsuppgifter minskar den mentala tröttheten vilket gör att läraren klarar att upprätthålla sin koncentration längre och således producera mer material. Detta kan i sin tur även höja motivationsnivån liksom tre av respondenterna vittnar om. Motivationen kan ytterligare förbättras av att lärarna upplever kvalitativa förbättringar i det egna arbetet. Sammantaget är alltså de upplevda förbättringar som respondenterna beskriver sammanlänkade och kan agera i en förstärkande cykel.

Flera av lärarna bekräftar litteraturens fynd om att behärskning av digitala verktyg och förmågan att integrera dessa i undervisningen kan förbättra både kvaliteten och tillgängligheten i utbildningen, som beskrivet av Pozo-Sánchez et al. (2020) och Babaev et al. (2020). Lärarna har, med hjälp av språkmodeller, kunnat skapa mer kvalitativa undervisningsmoment. Exempelvis framkom det av intervjuerna att genom att använda stora språkmodeller har lärarna kunnat förbättra och framförallt diversifiera tentafrågor i större utsträckning, vilket höjer kvaliteten på examineringar och gör dem mer utmanande för eleverna. Detta ger en intressant motvikt till den pågående debatten som problematiserar användandet av stora språkmodeller i utbildningssammanhang, där bland annat Brodie (2023) rapporterar om den oro som uttrycks över att verktygen tillåter studenter att ta genvägar i lärandet. Lärarnas erfarenheter kan användas för att visa att strategisk användning av språkmodellerna har potential att förstärka både lärandet och bedömningsprocessen genom att skapa mer kvalitativa undervisnings- tillika bedömningsmoment. Flera lärare vittnar om språkmodellernas potential att genomföra rättning, där exempelvis R7 anser att detta kan bidra till en mer rättvis och konsekvent bedömningsprocess som minimerar risken för subjektiva bedömningsfel. Detta belyser teknikens potential att förbättra både undervisning och bedömning, vilket är i linje med litteraturens betoning på digital kompetens som en katalysator för pedagogisk innovation och kvalitetsförbättring i undervisningen. R4 ser därtill potential för ökad skalbarhet i sina kurser där hen ser en avlastningspotential hos språkmodellerna som skulle kunna möjliggöra för hen att tillåta fler elever i sina kurser. Dessa aspekter bekräftar litteraturens syn på att behärskning av digitala verktyg, och förmågan att integrera dessa i undervisningen, kan förbättra kvaliteten och tillgängligheten i utbildningen.

De empiriska resultaten bekräftar litteraturen inom flera områden. Potential för tidsbesparing och ökad produktivitet framhävs som en potentiell konsekvens av användningen av språkmodeller av både Bates (2020) och Naveed et al. (2024), där Naveed et al. specifikt noterar att stora språkmodeller kan effektivisera arbetsflöden som lektionsplanering och rättning. Detta stöds även av Montenegro-Rueda et al. (2023), vilka påpekar att AI kan frigöra tid för lärare genom att automatisera repetitiva och tidskrävande uppgifter.

Kognitiv avlastning och motivationshöjning framträder som nya insikter ur empirin och bidrar med värdefulla dimensioner till förståelsen av hur språkmodeller påverkar lärarnas yrkesliv i positiv mening. Kognitiv avlastning understryker hur språkmodeller inte bara sparar tid utan även minskar den kognitiva belastningen för lärarna. Genom att avlasta dem från repetitiva och mentalt krävande uppgifter gör tekniken det möjligt för lärare att upprätthålla högre koncentrationsnivåer under längre perioder vilket således förbättrar deras arbetskapacitet. Motivationshöjning indikerar att inkorporeringen av språkmodeller kan bidra till att förbättra arbetsglädjen och engagemanget hos lärare men även höja lärarnas ambitionsnivå.

5.1.3 Upplevda begränsningar

Respondenterna uppvisar en mer samlad bild av medvetenheten av tekniska begränsningar hos språkmodeller jämfört med användningsområden. De flesta verkar ha en relativt tydlig uppfattning av de begränsningar som finns. Ur den sammantagna empirin framträder inte enbart en tydlig överensstämmelse mot befintlig litteratur, utan även vissa nya dimensioner inom detta område.

Respondenterna beskriver hallucinationer och variationer i informationskvaliteten som en tydlig begränsning och bekräftar därmed Ji et al. (2023) och Hua et al. (2023) som belyser språkmodellernas begränsningar i dessa hänseenden. Hua et al. (2023) kopplar problematiken till bristande kvalitet i träningsdata och bristande slutledningsförmåga hos språkmodellerna vilket pekar på ett kritiskt behov av att verktygens pålitlighet och trovärdighet ifrågasätts och hanteras varsamt. Hua et al. (2023) belyser även att bristerna i vissa fall kan härledas till användarens bristande förmåga att prompta modellen, vilket endast vissa av respondenterna uppvisar självmedvetenhet kring, vilket diskuteras vidare i avsnitt 5.2. Inom detta område beskriver R4 att informationen kan vara bristfällig när hen söker information inom sitt specifika ämne. Detta stämmer överens med diskussionen som Chang et al. (2024) framhåller, att även om stora språkmodeller med fördel kan användas för informationssökning, presterar den sämre i vissa domänspecifika ämnen.

Kasai et al. (2024) belyser att problematik med felaktigt innehåll ofta kan härledas till att de data som språkmodellen tränats på blir blivit utdaterad. Denna problematik speglas i empirin där flera av respondenterna noterar att språkmodeller har genererat utdaterad information. Exempelvis beskriver R2 specifika tillfällen där varken ChatGPT eller Copilot kunnat tillhandahålla aktuell information inom ett visst ämne. Denna överensstämmelse med litteraturen belyser en kritisk svaghet hos språkmodellerna i form av en oförmåga att uppdatera sin kunskapsbas i realtid. Att samtliga respondenter inte framhävde problematiken indikerar på en eventuell omedvetenhet kring denna begränsning vilket medför riskfyllt användande i enlighet med det som Weidinger et al. (2022) framhåller som Human-Computer Interaction Harms.

Vidare framhåller vissa av respondenterna en medvetenhet om subjektivitetsproblem hos språkmodellerna, där en tydlig parallell framträder mellan Rays (2023) argument om snedvridningar och de faktiska upplevelser som respondenterna redovisar för. Det råder samförstånd mellan litteratur och respondenter gällande språkmodellernas tendens att reproducera, och till viss del förstärka, befintliga mänskliga fördomar i kontexter där träningsdatan innehåller subjektiva åsikter eller perspektiv. Detta styrker behovet av att utveckla strategier för att minska dessa subjektivitetsproblem hos modellerna. Vidare belyser det behovet av att öka medvetenheten om problemet bland användarna för att minska risken för att denna typ av fördomar befäster sig i användarens egna världsbild såsom beskrivet av Weidinger et al. (2022). Detta ställer krav på AI-företagen att upprätthålla en öppenhet kring hur modellerna tränas, vilket kan bidra till att användaren bättre kan förstå hur stora språkmodeller fungerar och på ett bättre sätt kritiskt kunna bedöma den information som genereras. Detta är något som flera av respondenterna tar upp, bland annat R3 som understryker behovet av förståelse av tekniken bakom språkmodellerna för att kunna uppnå ansvarsfull användning.

En observation som inte speglas i litteraturavsnittet men som lyfts av empirin är specifika subjektiva begränsningar i form av varierande prestanda som är direkt kopplade till vilket språk som används när lärarna interagerar med verktygen. Modellerna beskrivs prestera

avsevärt sämre på svenska jämfört med engelska. Trots att det inte återfinns explicit i den befintliga litteraturen kan det ses förenligt med diskussionen om språkmodellerna subjektivitetsproblematik i och med att begränsningen kan härledas till en bristande mångfald i modellernas träningsdata.

Respondenterna uttrycker slutligen en oro över hanteringen och lagringen av känslig data, vilket bekräftar den problematik Carlini et al (2023) och Staab et al (2024) berör. Denna oro begränsar deras användning av språkmodeller inom akademiska sammanhang och hindrar dem från att genomföra saker som rättning av studentarbeten med hjälp av verktygen. Inom akademiska miljöer värderas integritet högt och det kan därför ses naturligt att detta är något som skapar tveksamhet bland lärarna. Denna osäkerhet visar att språkmodellernas användning inom akademiska miljöer möjligtvis bör begränsas tills dess att integriteten för både lärare och studenter kan garanteras. En sådan aspekt ligger utanför universitetens såväl som lärares kontroll och ligger istället i händerna på de företag som utvecklar dessa teknologier. Begränsningen medför dock en bestående osäkerhet kring hur språkmodellerna får användas inom utbildningsväsendet, något som kan anses befinna sig inom universitetens ansvarsområden. Det belyser således ett behov av att lärosätena tar ett aktivt ansvar att etablera tydligare riktlinjer och regleringar kring användningen av språkmodeller i skolmiljöer.

Slutligen reflekterar empirin en osäkerhet kring användningen av den data som delas med verktygen, eftersom datan sedan ägs av de storföretag som kontrollerar verktygen och fritt kan använda den för egna intressen.

5.2 Prompt engineering

Litteraturen beskriver prompt engineering som en kritisk teknik för att maximera effektivitet och relevans i de svar som genereras av stora språkmodeller, såsom fastställt av White et al.(2023), Chen et al. (2023), Ekin (2023), Giray (2023), Korzyński et al. (2023) och Lo (2023). Även Eager & Brunton (2023) argumenterar för detta, specifikt för lärare inom en utbildningskontext. Författarna argumenterar för att kvaliteten på det genererade svaret är direkt kopplat till kvaliteten på den prompt användaren förser verktyget med. Av empirin framgår det att även om alla respondenter visar medvetenhet om prompt engineering som koncept, varierar deras engagemang och tillvägagångssätt för tekniken betydligt. Vissa respondenter betraktar, i enlighet med litteraturen, prompting som en kritisk byggsten i sin interaktion med språkmodeller medan andra huvudsakligen använde sig av enklare prompts och därtill uppvisade begränsat intresse för fenomenet.

En intressant aspekt som framkommer ur en jämförelse av empirin mot litteraturen är att Hua et al. (2023) förstärker tidigare nämnda författarnas argument genom att belysa att bristerna i prestanda i vissa fall kan härledas till användarens bristande förmåga att prompta modellen. Från empirin framkommer att flera respondenter beskriver sig ha fått otillfredsställande resultat från stora språkmodeller. I vissa fall uppvisar respondenterna en självmedvetenhet om att detta kan bero på deras egna bristande promptingförmåga medan andra ser detta som en direkt begränsning hos språkmodellen. Detta visar på en potential för att vissa av de problem som lärarna beskriver sig stött på skulle kunna förklaras med en bristande förmåga att formulera effektiva prompts. De fall där respondenterna uttrycker osäkerhet kring språkmodellernas effektivitet och tillförlitlighet kan därför vid vissa tillfällen vara direkt kopplade till användarens oförmåga att skapa välfungerande prompts.

De specifika tekniker som beskrivs i litteraturen återspeglas även i respondenternas användning. Respondenterna anger tekniker som iterative prompting (Wang et al., 2022), role-prompting (Chen et al., 2023), chain-of-thought prompting (Zhou et al., 2022) samt few-shot prompting (Brown et al., 2020). White et al. (2023) belyser även potentialen för modellerna att föreslå strategiska prompts till användaren, vilket R4 bekräftar genom sin egen användning där hen ibland använt ChatGPT för att få förslag på hur hen bäst instruerar verktyget till att nå önskat slutresultatet. R1, som möjligen uppvisat störst försiktighet till språkmodeller bland respondenterna, beskriver att hen främst använder zero-shot prompting som beskrivet av Naveed et al. (2024). De mer avancerade användarna, såsom R3, R4 och R7, uppvisar istället ett systematiskt utforskande av olika promptingtekniker. En intressant observation från undersökningen är att även de mer avancerade användare ibland föredrar att använda enklare prompts, såsom zero-shot prompting, beroende på den specifika uppgiften de står inför. Detta visar på en strategisk flexibilitet hos användarna där uppgiftens natur direkt påverkar användarens interaktion med verktyget. Detta tyder på ett digitalt kompetent användande där respondenterna förstår när och hur olika typer av prompts leder till mest tillfredsställande resultat. R2 uttrycker att hen upplever en brist på framtidsutsikter för området prompt engineering, vilket hen härleder till att sådan kunskap sannolikt snabbt blir förlegad och att detta gör att respondenten känner sig omotiverad till att sätta sig in i ämnet. Detta motsätter sig Eager & Bruntons (2023) argument om att prompt engineering kommer att specifikt för lärare vara en viktig kunskap för att uppnå kvalitativt undervisningsmaterial. Sammantaget belyser denna spridning en skillnad i hur användare med olika kunskap eller intresse kan se och utnyttja potentialen hos språkmodeller.

Avslutningsvis speglas en medvetenhet och viss beaktning, framförallt bland de mer avancerade användarna, kring de rekommendationer som Korzyński et al. (2023) och Lo (2023) beskriver i sina ramverk. Respondenterna visar en tydlig medvetenhet om behovet av tydliga och logiska instruktioner där de inkluderar element som kontext och förväntningar. Vidare uppvisas en viss lyhördhet för modellen där lärarna anpassar sina prompts utifrån tidigare erfarenheter och feedback från modellen. Även de mindre avancerade användarna uppvisar medvetenhet kring en del av de aspekter som Korzyński et al. (2023) och Lo (2023) presenterar.

5.3 Digital kompetens

5.3.1 Lärares syn på digital kompetens

Empirin visar att digital kompetens är något som de flesta av lärarna värderar som viktigt, vilket överensstämmer med litteraturen som betonar behovet av kvalificerade, digitalt kompetenta lärare i dagens teknikdrivna samhälle (European Commission, 2019; Vuorikari et al., 2022). Även om det råder enighet bland respondenterna om vikten av att hålla sig uppdaterad med digitala verktyg, framgår det av empirin en avvikelse från litteraturen i synen på lärarnas roll i att säkerställa och främja denna kompetens. Litteraturen argumenterar för lärarens aktiva roll i att både utveckla sin egen digitala kompetens och även främja och stödja sina studenters utveckling av digital kompetens (Fernández-Batanero et al., 2020). Bland respondenterna framkommer dock en generell osäkerhet kring deras ansvar i sammanhanget där flera anser att det inte är lärarnas ansvar att säkra denna utveckling bland studenterna. Denna osäkerhet pekar på att det råder olika uppfattningar om rollfördelningen, vilket framhäver en potentiell klyfta mellan EU:s förväntningar och lärarnas personliga syn på sitt

arbetsansvar. Detta kan ha direkt påverkan på hur digital kompetens faktiskt främjas i svenska utbildningsmiljöer.

5.3.2 Utbildning

Litteraturen betonar ett omfattande behov för universitetsinstitutioner att ta en aktiv roll i att kontinuerligt utbilda och stödja lärare i deras professionella utveckling av digital kompetens, för att säkerställa kvalitativ utbildning och lärarkompetens som möter samhällets och individers föränderliga behov (European Commission, 2019; Vuorikari et al., 2022). Empirin visar dock att i avseendet av just stora språkmodeller överförs inte detta helt i praktiken. Ingen av respondenterna har mottagit formell utbildning eller stöd från respektive institution för att lära sig hantera eller inkorporera stora språkmodeller i sitt arbete. Den brist på strukturerad utbildning som respondenterna vittnar om står i kontrast till litteraturens betoning på betydelsen av lärarnas digitala kompetens och utveckling, men även universitetens ansvar att säkerställa att dessa förutsättningar finns på plats. Studier av bland annat Pozo-Sánchez et al. (2020), Babaev et al. (2020), Ausat et al. (2023) och Fernández-Batanero et al. (2020) framhåller att det är kritiskt att lärare får stöd i sin professionella utveckling för att effektivt kunna navigera i den teknologiskt föränderliga utbildningsmiljön. Anoun (2017) och Vanthournout (2014) stärker detta påstående genom att argumentera för att det är just lärosätena i rollen som arbetsgivare som bär det yttersta ansvaret att säkerställa att lärare såväl som elever erbjuds möjligheter för att utveckla sådana kompetenser. Bristen på universitetsdrivna initiativ kan ses som ett potentiellt hinder för fullt utnyttjande av tekniken i lärarnas arbetssammanhang.

I samråd med empirin betonar således resultaten ett behov av fler utbildningsinitiativ från universitetets sida för att stärka den digitala kompetensen hos lärare, vilket går i enlighet med EU:s rekommendationer (European Commission, 2019) och Vuorikari et al. (2022). Dessa insatser ses avgörande för att säkerställa högkvalitativ utbildning som möter både samhällets och individernas föränderliga behov. Från intervjuerna framgår dock att flera av lärarna är tveksamma till att motta utbildning i frågan och således motsätter sig litteraturens rekommendationer. Flera respondenter är av åsikten att det inte är möjligt att effektivt utbilda sig i användningen av stora språkmodeller. R2 är exempelvis skeptisk mot formell utbildning då hen anser att den snabbt avancerade tekniska utvecklingen gör att dagens kunskap kan vara förlegad imorgon. R1 argumenterar istället för att det handlar om intressen hos individer, vilket är oberoende av huruvida man får formell utbildning eller ej. Vad som är intressant är att det, trots detta, ändå sker ett betydande kollegialt kunskapsutbyte där lärarna lär sig av varandra vilket går i motsats till dessa utbildningsskeptiska åsikter. Detta tycks dock ske sporadiskt och slumpmässigt, vilket inte lämnar någon garanti för att de framgångar eller hinder lärarna stöter på delas med andra användare. I kombination med den stora spridningen som uppvisas i användningen indikerar detta på att användningen av stora språkmodeller fortfarande befinner sig i ett experimentellt stadie, där respondenter testat sig fram för att själva hitta användningsområden som känns passande att tillämpa i sin yrkesroll. Trots att en del respondenter ställer sig kritiska till möjligheterna att formalisera utbildningar i dessa teknologier, visar det kollegiala utbytet att det finns framgångsrika exempel som demonstrerar potentialen för att lära sig av andra. Detta tyder på att lärare kan inspireras av varandra och ta del av de mångsidiga tillämpningarna av språkmodeller, vilket erbjuder potential för att lärarna kan utöka sitt användningsspektrum.

5.3.3 Kritisk förhållning

Kritisk förhållning och ansvarsfull användning av stora språkmodeller belyser ett centralt tema inom digital kompetens. Därmed utgörs viktiga delar av empirin inte enbart av respondenternas aktiva medvetenhet om språkmodellernas begränsningar, utan även deras strategier för att hantera dessa. Det framgår tydligt ur empirin att lärarna uppvisar flera exempel på god anpassning till språkmodellerna ur ett kritiskt likväl som etiskt perspektiv.

Ett genomgående tema som framträder ur empirin är medvetenhet kring valideringsprocesser. Empirin betonar delvis behovet av domänkunskaper inom de kunskapsområden där språkmodellerna används, där flera av respondenterna medvetet begränsar sitt användande av stora språkmodeller till domäner som faller inom deras egna expertis. Lärarna beskriver detta som en kvalitetssäkringsprocess som underlättar valideringen och minimerar risken för att hallucinationer passerar obemärkt förbi. Andra beskriver även att de validerar språkmodellerna mot externa källor. Dessa strategier visar på en kritisk och ansvarsfull förhållning som går väl i linje med kriterierna för digital kompetens enligt Vuorikari et al. (2022). Anpassningen av användningsområden är ett ytterligare framträdande tema för lärarnas kritiska förhållning. Vissa lärare undviker att använda språkmodeller för uppgifter där de anser att modellernas förmåga är begränsad såsom för uppgifter med kreativa eller komplexa inslag. Istället föredrar de att nyttja stora språkmodeller till uppgifter där verktygets begränsningar inte anses utgöra ett lika stort hinder för användningen.

Dessa sammantagna beskrivningar av selektiv användning visar på en djup förståelse och anpassning till språkmodellerna. Genom att aktivt begränsa användningen till specifika områden, samt noggrant validera den information som språkmodellerna genererar, minimerar lärarna aktivt risken för felaktigheter. Detta bedöms särskilt viktigt i akademiska sammanhang där korrekthet av information ses avgörande för ett tillförlitligt utbildningssystem. Eventuella felaktigheter i forsknings- eller undervisningsmaterial riskerar att både vilseleda mottagaren men också underminera förtroendet för utbildningsinstitutionen. Lärarnas metodiska begränsning och validering av stora språkmodeller är därför avgörande för att upprätthålla en hög akademisk standard. Vidare illustrerar kombinationen av proaktiva och reaktiva strategier ett dynamiskt förhållningssätt till tekniken.

De tecken på god kritisk förhållning som framkommer i empirin nyanseras av vissa inkonsekvenser i respondenternas ansvarsfulla tillvägagångssätt. Detta indikerar att även om en övergripande medvetenhet om språkmodellernas begränsningar finns, förekommer det stundtals brister i hur denna medvetenhet omsätts i praktiken. Denna aspekt blir särskilt tydlig när vissa respondenter, trots uttalad medvetenhet, ändå använder teknologin på sätt som kan ifrågasättas ur ett etiskt perspektiv. Detta belyser en viss inkonsekvens i hur lärarna förhåller sig till språkmodeller från ett etiskt perspektiv.

Sammanfattningsvis speglar lärarnas användning av stora språkmodeller en välutvecklad men varierande grad av kritisk förhållning. Respondenternas förmåga att anpassa och validera användningen av stora språkmodeller ses avgörande för att säkerställa akademisk integritet och pedagogisk kvalitet. De inkonsekvenser som belyses framhäver ett behov av att förstärka medvetenheten om de etiska dimensionerna av teknikanvändning för att säkerställa att alla lärare kan navigera i en tekniskt avancerad miljö på ett ansvarsfullt sätt.

5.3.4 Applicering av DigComp-ramverket

För att ytterligare konkretisera hur användningen av stora språkmodeller kan främja digital kompetens hos lärare används DigComp-ramverket som en lins genom vilken stora språkmodeller granskas för att analysera dess inverkan på digital kompetens. Tabell 5:1 illustrerar specifika kompetensområden som berörs vid lärarnas beskrivna användning av verktygen.

Tabell 5:1 Applicering av DigComp-ramverket på stora språkmodeller

Area	Competence	Use of LLM
Information and data literacy	1.1. Browsing, searching and filtering data, information and digital content	✓
	1.2. Evaluating data, information, and digital content	✓
	1.3. Managing data, information, and digital content	✓
Communication and collaboration	2.1. Interacting through digital technologies	
	2.2. Sharing through digital technologies	
	2.3. Engaging in citizenship through digital technologies	
	2.4. Collaborating through digital technologies	
	2.5. Netiquette	
	2.6. Managing digital identity	
Digital content creation	3.1. Developing digital content	✓
	3.2. Integrating and re-elaborating digital content	✓
	3.3. Copyright and licenses	
	3.4. Programming	✓
Safety	4.1. Protecting devices	
	4.2. Protecting personal data and privacy	
	4.3. Protecting health and well-being	✓
	4.4. Protecting the environment	
Problem solving	5.1. Solving technical problems	✓
	5.2. Identifying needs and technological responses	✓
	5.3. Creatively using digital technologies	✓
	5.4. Identifying digital competence gaps	

Tillämpningen av ramverket baseras strikt på den direkta användningen av språkmodellerna såsom beskriven av respondenterna, snarare än vad modellerna har kunskap om eller potential att assistera i. Således har exempelvis *2.1 Interacting through digital technologies* inte markerats trots att lärarna beskriver användning av verktygen för att exempelvis formulera

mail, vilket är en form av interaktion. Detta beror på att en språkmodell inte används som en direkt kommunikationsplattform i dessa fall utan som ett assisterande språkhjälpmiddel i förmån av kommunikationen. På liknande vis har inte 3.3 *Copyright and licences* markerats, trots att en stor språkmodell skulle kunna assistera i att förse användaren med kunskap i denna typ av frågor.

Genom att koppla samman undersökningsdata med DigComp-ramverket synliggörs specifika kompetenser som användningen av stora språkmodeller bidrar till att främja hos lärarna. Ramverket visar att lärarnas användning av verktyg framförallt främjar digitala kompetenser inom *Information and data literacy*, *Digital content creation* och *Problem solving*. Detta stärker synen på språkmodellernas praktiska nytta inom högre utbildningsinstitutioner sett utifrån digital kompetens, samtidigt som det även belyser de områden som användandet av modellerna inte täcker.

5.4 Stora språkmodellers potential inom utbildning

Respondenterna är eniga om att stora språkmodeller är här för att stanna. Däremot uppvisas en delad syn på potentialen för dessa framgent. Respondenterna R1 och R2 tillhör de som framhåller en viss misstro mot framtidsutsikterna för språkmodeller. De är av tron att språkmodeller inte kommer att medföra några betydande revolutionerande effekter inom utbildning, vilket sammanfaller med respondenternas tidigare uttryckta skepsis mot formell utbildning inom användningen av stora språkmodeller. Dessa resultat föreslår att potentialen för språkmodellerna inte bör övervärderas och att det inte finns någon övertygande anledning att behandla dessa modeller som något mer än ytterligare ett verktyg i det pedagogiska landskapet.

R3 och R7 ställer sig mer optimistiska till framtiden för språkmodeller inom sitt yrke och ser betydande potential för nya möjligheter inom läraryrket. De framhäver särskilt möjligheten att använda stora språkmodeller för att automatisera arbetsuppgifter såsom bedömning och feedback. R3 och R7 har båda experimenterat med att använda sådana modeller för rättning och identifierar stor potential för dessa modeller att effektivisera och standardisera denna typ av processer. De framhäver en tro om att detta på sikt skulle kunna resultera i mer objektiva och konsekventa bedömningar, vilket Chang et al. (2023) delvis motsätter sig då författarna framhåller att lärare har unika insikter som förmodligen inte går att ersätta. Huruvida rättning med hjälp av stora språkmodeller kommer att kunna sättas i bruk tycks dock, som tidigare nämnt, vara en fråga som ligger utanför lärarnas direkta kontroll eftersom en etisk implementering av denna typ av bedömningsprocess kräver förstärkt skydd från utvecklarna av språkmodellernas håll. Nguyen et al. (2022) belyser just dessa etiska aspekter och betonar vikten av att beakta dem vid integrering av stora språkmodeller i utbildningssektorn. Vidare uttrycker R4 och R6 oro över hur väl storföretagens intressen, som äger verktygen, överensstämmer med akademiska värderingar, vilket bekräftas av den problematik som Cerratto et al. (2024) belyser. Stora språkmodeller är inte specifikt byggda för utbildningsändamål, och lärarnas uppfattning är att det kommer finnas kvar etiska hinder kopplade till användningen av stora språkmodeller inom högre utbildning tills en sådan konfigurerings har fastställts.

När det kommer till språkmodellernas påverkan på respondenternas framtida roll som lärare råder det en relativt övergripande konsensus. Den allmänna åsikten är ett erkännande av att modellerna kommer att kunna hantera vissa arbetsuppgifter som traditionellt har utförts av

lärare, men att detta inte kommer att leda till en ersättning av lärarna, utan snarare en omkonfigurering av lärarrollen. Detta går i enlighet med Isemonger (2023) och Eager & Brunton (2023) som belyser lärares föränderliga roll i avseende för användning av stora språkmodeller. I takt med att tekniken accelererar förväntas nya krav ställas på lärarna att komplettera dessa teknologiska verktyg med mänskliga kvaliteter såsom kreativitet och etisk bedömning, vilka förväntas vara fortsatt avgörande för utbildningens kvalitet och relevans. Jeon & Lee (2023) talar om att det krävs ett starkare samarbete mellan mänskliga lärare och stora språkmodeller för att säkerställa en kvalitativ utbildning. R5 förhåller sig mest reserverad till stora språkmodellens påverkan för just rollen som lärare då hen är av åsikten att lärarnas yrkesetik inte går att ersätta. Detta kan kopplas till Chang et al. (2023) påstående om att stora språkmodeller inte har unika insikter på det sättet lärare har.

Sammantaget visar diskussionen att medan stora språkmodeller erbjuder betydande möjligheter för att förbättra och effektivisera undervisning, kräver det också en omvärdering och anpassning av lärarrollen för att säkerställa att kärnvärden såsom kritiskt tänkande, personlig utveckling och kreativ problemlösning fortsätter att prioriteras i en alltmer digitaliserad värld. Här råder det en oenighet hos lärarna, som indikerar att det nog kommer att dröja innan ett sådant paradigmskifte kan ske, då en andel av lärare kommer att motsätta sig denna förändring i och med att de inte ser sin roll som sådan i den utvecklingen.

6 Slutsats

I denna studie har användningen av stora språkmodeller (LLM) bland universitetslärare utforskats, vilket visar på en högst individuell användning som speglar personliga preferenser och behov snarare än institutionella direktiv eller ämnesspecifika normer. Denna variation i användning belyser en betydande men ännu inte fullt utnyttjad potential för en bredare integrering av språkmodeller inom läraryrket. Identifierade användningsområden inkluderar kunskapsinhämtning, framställning av undervisningsmaterial, analysarbete, kreativt bollplank, språkhjälp & generativ förmåga samt skapande av examinerande material. Användningen är sammankopplad med märkbara fördelar som innefattar tidsbesparing, kognitiv avlastning, ökad produktivitet, förbättrad kvalitet och höjd motivation. De identifierade förbättringarna framträder inte som isolerade vinster utan tycks interagera och förstärka varandra i en positiv spiral.

Studien har även uppmärksammat betydande begränsningar i användningen av stora språkmodeller, såsom att språkmodellerna genererar hallucinationer, föråldrad information och snedvridning av data. Vissa begränsningar i träffsäkerhet kunde kopplas till en bristande förmåga bland användarna att effektivt formulera prompts, vilket befintlig forskning framhåller som en kritisk faktor. Dessa tekniska begränsningar har lett till utvecklingen av både proaktiva och reaktiva anpassningsstrategier bland lärarna. Dessa visar på medvetna initiativ från lärarna att använda stora språkmodeller på ett säkert, kritiskt och ansvarsfullt sätt, vilket reflekterar ett digitalt kompetent förhållningssätt. Digital kompetens är något som de flesta av lärarna anser är viktigt och genom applicering av DigComp-ramverket framgår även att användningen av stora språkmodeller främjar flertalet digitala kompetenser, vilket stärker dess praktiska nytta inom högre utbildningsinstitutioner.

Ingen av respondenterna har fått formell utbildning eller stöd från sina institutioner för att integrera stora språkmodeller i sitt arbete. Även om en enhetlig användning inte anses nödvändig, indikerar studien att gemensam utbildning hade kunnat hjälpa lärare att bättre utnyttja dessa verktygs potential. Trots detta är många lärare tveksamma till att delta i formell utbildning, vilket utmanar de rådande rekommendationerna. Studien understryker därför istället vikten av att främja kollegiala kunskapsutbyten.

Sammantaget framkommer det ur studien att användningen av stora språkmodeller inom högre utbildning reflekterar en komplex blandning av möjligheter, utmaningar och reservationer. Trots den uppvisade potentialen hos språkmodeller finns det en märkbar återhållsamhet bland lärarna, rotad i både institutionella och etiska begränsningar i form av osäkerheter gällande frågor som dataskydd, integritet och ansvarsfrågor som inte kan ignoreras. Detta pekar på ett behov för lärosäten att stärka sitt stöd mot lärarna och ta ett större ansvar i att skapa tydliga riktlinjer för hur teknologin får och inte får användas, i syfte att främja en etiskt ansvarsfull och säker användning av språkmodellerna.

Studien framhäver även en bristande konsensus om språkmodellernas framtida roll inom utbildningen, vilket understryker behovet av ökad dialog och samarbete för att utveckla effektiva och etiskt hållbara användningsstrategier. Den breda mångfalden i åsikter och tillvägagångssätt bland lärarna pekar också på att ingen enhetlig 'best practice' kan fastställas, vilket gör det svårt att ge specifika rekommendationer för hur tekniken bör implementeras i den akademiska miljön. Det tycks inte vara en fråga om att hitta ett rätt sätt att använda

teknologin på, utan snarare att förstå dess potential respektive begränsningar inom olika arbetsområden. Detta ställer även krav på universiteten att främja en miljö där lärare kan utforska och experimentera med teknologin för att hitta anpassningar som bäst uppfyller deras unika behov.

6.1 Förslag på vidare forskning

Det finns flera riktningar för vidare forskning som kan bidra till en bättre förståelse av den komplexitet som denna studie har belyst. I denna studie har endast enskilda lärare intervjuats, och ett bredare perspektiv på organisatoriska faktorer har inte undersökts. Ett område som därmed kan vara av särskilt värde är att undersöka alla organisatoriska nivåer på ett enskilt universitet, då denna studie har visat att lärares attityder och användning av stora språkmodeller inte verkar vara en isolerad företeelse. Användningen styrs inte enbart av individuella uppfattningar utan påverkas också av ett samspel mellan universitet, fakultet, institution och individer. Att studera hela organisationskedjan och företagskulturen vid ett universitet kan därför ge en djupare förståelse för hur faktorer på olika nivåer utanför enskilda individer bidrar till att forma användningen av dessa verktyg inom läraryrket.

Eftersom denna studie begränsades till lärare som redan använder stora språkmodeller i sitt yrke, kan det i framtida forskning vara värdefullt att inkludera lärare som inte använder dessa verktyg och anpassa intervjufrågorna därefter. Eftersom undersökningen visade en viss återhållsamhet och skepsis bland respondenterna, kan en mer nyanserad bild uppnås genom att inkludera lärare som väljer att avstå från att använda språkmodeller i undersökningen. Detta kan ge en djupare förståelse och eventuellt en mer rättvis bild av de faktorer som påverkar detta komplexa fenomen.

Slutligen skulle det vara värdefullt att komplettera med en studie av en annan metod. En kvantitativ studie skulle kunna bidra med mer omfattande kunskap genom större datainsamling och ge en mer generaliserande bild där samband mellan faktorer som kön, institution, ämne och ålder tydligt kan ses. En sådan studie skulle ge en bredare förståelse av Sveriges lärarkår som helhet och därmed komplettera den kvalitativa undersökning som genomförts i denna uppsats.

Bilagor

Bilaga 1-7, vilka innehåller transkriberingar av de intervjuer som använts i undersökningen, har exkluderats vid denna publicering.

Bilaga 8 - Samtyckesformulär

Samtyckesformulär

1. PERSONUPPGIFTSANSVARIG (Intervjuare)

Namn	
Mail	Telefon

2. RESPONDENT

Namn	
Mail	Telefon

3. BAKGRUND

Fokus för intervjun är att samla in data till en undersökning om användning av Large Language Models i det egna arbetet. Syftet med undersökningen är att utforska upplevda möjligheter och utmaningar med att integrera Large Language Models inom den svenska högre utbildningssektorn.

4. BEHANDLING AV UPPGIFTER

Intervjun kommer att spelas in genom ljudinspelning och därtill transkriberas med hjälp av digitala verktyg. Därefter kommer ljudinspelningen att raderas. Respondenten kommer att förbli anonym genom åtgärder vidtagna för viss revidering av transkriberingen för att avidentifiera respondenten i syfte att skydda dennes identitet. Endast den reviderade transkriberingen kommer att bifogas i undersökningen. Data används endast för forskningen.

Genom att skriva under samtycker du till deltagande av intervjun och behandling av den insamlade datan enligt ovanstående krav. Deltagandet är frivilligt och du har när som helst rätt att dra tillbaka ditt samtycke.

Underskrift:

Namnförtydligande:

Ort och datum:

Bilaga 9 - AI-bidragsredogörelse

AI-bidragsredogörelse

I färdigställandet av vår kandidatuppsats har vi använt följande AI-baserade verktyg:

- ChatGPT-4
- Whisper

ChatGPT-4: Verktöget har stöttat flera delar av vårt arbete. Allra främst användes ChatGPT-4 i ett initialt skede som ett bollplank för att assistera i att forma fokuset för uppsatsen. Vidare användes verktöget till viss grad för att förbättra språket och strukturen i våra texter men även som ett kritiskt innehållsgranskande stöd.

Specifikt har ChatGPT-4 använts för att:

- Förbättra formuleringar och språklig precision genomgående i uppsatsen.
- Ge strukturförslag för kapitel 2 (Litteraturgenomgång) och kapitel 4 (Empiriska resultat).
- Sammanfatta och syntetisera vissa källor under litteraturinsamlingen. Vid det att en källa bedömdes relevant gjordes dock en egen noggrannare granskning av källan.
- Kritisk granskning av metodkapitlet för att identifiera förbättringsområden.

Whisper: Verktöget har endast använts för att transkribera ljudinspelningar från våra intervjuer, vilket direkt bidrog till datamaterialet som presenteras i kapitel 4 (Empiriska resultat). Efter transkriberingen genomfördes en noggrann genomläsning i takt med ljudinspelningen för att säkerställa att den kvalitativa datan var korrekt representerad i transkriberingarna.

Sammanfattning: Dessa verktyg har använts för att effektivisera arbetsprocesser som transkribering och skapa överblick av litteratur, men även för att höja den språkliga kvaliteten och strukturen i vår uppsats. Vid språklig assistans har ChatGPT-4 inte använts för att generera text från grunden, utan enbart för att föreslå språkliga förbättringar på redan färdigskrivna texter. Detta säkerställer att den slutliga texten behåller sin originalitet och våra egna analytiska och kritiska tankeprocesser.

Denna redogörelse bekräftar vår ansvarsfulla användning av AI-teknologier i akademiska syften, i linje med de riktlinjer och etiska överväganden som anges av vår akademiska institution.

Referenser

- Alshurafat, H. (2023). *The Usefulness and Challenges of Chatbots for Accounting Professionals: Application On ChatGPT*. [online] papers.ssrn.com. Tillgänglig via: <https://ssrn.com/abstract=4345921> [Hämtad 28 mars 2024].
- Alvehus, J. (2019). *Skriva uppsats med kvalitativ metod : en handbok*. Stockholm: Liber.
- Aoun, J.E. (2017). ROBOT-PROOF. [online] Tillgänglig via: http://repo.darmajaya.ac.id/4817/1/Robot-Proof_%20Higher%20Education%20in%20the%20Age%20of%20Artificial%20Intelligence%20%28%20PDFDrive%20%29.pdf [Hämtad 2 maj 2024].
- Ausat, A.M.A., Massang, B., Efendi, M., Nofirman, N. & Riady, Y. (2023). Can Chat GPT Replace the Role of the Teacher in the Classroom: A Fundamental Analysis. *Journal on Education*, [online] 5(4), pp.16100–16106. doi:<https://doi.org/10.31004/joe.v5i4.2745>
- Babaev, D., Saipidinova, B., Babaeva, A., Alieva, B. & Abdullaeva, Z. (2020). Information and Communicative Competence Development in Prospective Teachers. *Creative Education*, [online] 11(12), pp.2867–2875. doi:<https://doi.org/10.4236/ce.2020.1112212>
- Baidoo-Anu, D. & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, [online] 7(1), pp.52–62. doi:<https://doi.org/10.61969/jai.1337500>
- Bates, T., Cobo, C., Mariño, O. & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, [online] 17(1). doi:<https://doi.org/10.1186/s41239-020-00218-x>
- Bojukrapan, S., Laoha, R. & Jantakoon, T. (2023). Higher Education Studies. *Higher Education Studies*, [online] 13(4). doi:<https://doi.org/10.5539/hes.v13n4p227>
- Bommasani, R., Hudson, D., Adeli, E., Altman, R., Arora, S., Von Arx, S., Bernstein, M., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J., Demszky, D. & Donahue, C. (2022). On the Opportunities and Risks of Foundation Models. [online] arxiv, p.214. Tillgänglig via: <https://arxiv.org/abs/2108.07258> [Hämtad 9 maj 2024].
- Brodie, E. (2023). AI writing tools garner concern about academic integrity, education from faculty. [online] *The Vermont Cynic*. Tillgänglig via: <https://vtcynic.com/news/ai-writing-tools-garner-concern-about-academic-integrity-education-from-faculty/> [Hämtad 9 maj 2024].
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C. & Hesse, C. (2020). *Language Models are Few-Shot Learners*. [online] Tillgänglig via: <https://arxiv.org/pdf/2005.14165.pdf> [Hämtad 9 april 2024].
- Carlini, N., Ippolito, D., Jagielski, M., Lee, K., Tramèr, F. & Zhang, C. (2023). Quantifying Memorization Across Neural Language Models [online] *International Conference on Learning Representations*, p.19. Tillgänglig via: https://openreview.net/forum?id=TatRHT_1cK [Hämtad 9 maj 2024].
- Cerratto Pargman, T., Sporrang, E., Farazouli, A. & McGrath, C. (2024). Beyond the Hype: Towards a Critical Debate About AI Chatbots in Swedish Higher Education. *Högre utbildning*, [online] 14(1), pp.74–81. doi:<https://doi.org/10.23865/hu.v14.6243> [Hämtad 3 maj 2024]

- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P.S., Yang, Q. & Xie, X. (2024). A Survey on Evaluation of Large Language Models. *ACM Transactions on Intelligent Systems and Technology*, [online] 15(3), pp.1–45. doi:<https://doi.org/10.1145/3641289> [Hämtad 27 mars 2024].
- Chen, B., Zhang, Z., Langrené, N. & Zhu, S. (2023). *Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review*. [online] Tillgänglig via: <https://arxiv.org/pdf/2310.14735.pdf> [Hämtad 27 mars 2024].
- Chen, Y., Jensen, S., Albert, L.J., Gupta, S. & Lee, T. (2022). Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success. *Information Systems Frontiers*, 25(161–182). doi:<https://doi.org/10.1007/s10796-022-10291-4>
- Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koochang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D. & Carter, L. (2023). ‘So what if ChatGPT wrote it?’ Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, [online] 71(0268-4012), p.102642. doi:<https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eager, B. och Brunton, R. (2023). Prompting Higher Education Towards AI-Augmented Teaching and Learning Practice. *Journal of University Teaching & Learning Practice*, [online] 20(5). doi:<https://doi.org/10.53761/1.20.5.02>
- Ekin, S. (2023). *Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices*. [online] www.techrxiv.org. doi:<https://doi.org/10.36227/techrxiv.22683919.v2>
- Esaiasson, M. (2023). *Frågor och svar om GAI-verktyg | För lärande i en digital värld*. [online] www.education.lu.se. Tillgänglig via: <https://www.education.lu.se/aktuella-teman/generativa-ai-verktyg/fragor-och-svar-om-gai-verktyg> [Hämtad 20 mars 2024].
- European Commission, Directorate-General for Education, Youth, Sport and Culture (2019). Key competences for lifelong learning. [online] *Publications Office of the European Union*. European Commission. Tillgänglig via: <https://op.europa.eu/en/publication-detail/-/publication/297a33c8-a1f3-11e9-9d01-01aa75ed71a1/language-en> [Hämtad 8 april 2024].
- European Union (2016). *EUR-Lex - 32016R0679 - EN - EUR-Lex*. [online] EUR-Lex. Tillgänglig via: <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=celex%3A32016R0679> [Hämtad 8 maj 2024].
- Fernández-Batanero, J.M., Montenegro-Rueda, M., Fernández-Cerero, J. & García-Martínez, I. (2020). Digital competences for teacher professional development. Systematic review. *European Journal of Teacher Education*, 45(4), pp.513–531. doi:<https://doi.org/10.1080/02619768.2020.1827389>
- Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering*, [online] 51(2629–2633). doi:<https://doi.org/10.1007/s10439-023-03272-4>
- Hashem, R., Ali, N., Zein, F.E., Fidalgo, P. & Khurma, O.A. (2024). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, [online] 19(23), p.26. doi:<https://doi.org/10.58459/rptel.2024.19023>

- Hua, S., Jin, S. & Jiang, S. (2023). The Limitations and Ethical Considerations of ChatGPT. *Data intelligence*, [online] 6(1), pp.1–38. doi:https://doi.org/10.1162/dint_a_00243
- Isemonger, I. (2023). Generative Language Models in Education: Foreign Language Learning and the Teacher as Prompt Engineer. *Praxis Journal*, [online] 2(2758-5700), pp.3–17. doi:<https://doi.org/10.5281/zenodo.10402411>
- Jeon, J. & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, [online] 28(12), pp.15873–15892. doi:<https://doi.org/10.1007/s10639-023-11834-1>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y.J., Madotto, A. & Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, [online] 55(12), pp.1–38. doi:<https://doi.org/10.1145/3571730>
- Kasai, J., Sakaguchi, K., Takahashi, Y., Le Bras, R., Asai, A., Velocity, X., Dragomir, Y., Smith, N., Choi, Y. & Riken (2024). *REALTIME QA: What's the Answer Right Now?* Kentaro Inui. [online] Tillgänglig via: <https://arxiv.org/pdf/2207.13332> [Hämtad 9 maj 2024].
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T. & Stadler, M. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, [online] 103(102274). doi:<https://doi.org/10.1016/j.lindif.2023.102274>
- Khurana, D., Koli, A., Khatter, K. & Singh, S. (2022). Natural Language processing: State of the art, Current Trends and Challenges. *Multimedia Tools and Applications*, [online] 82(3), pp.3713–3744. doi:<https://doi.org/10.1007/s11042-022-13428-4>
- Korzyński, P., Mazurek, G., Krzyrkowska, P. & Kurasiński, A. (2023). Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. *Entrepreneurial Business and Economics Review*, 3(25-38). doi:<https://doi.org/10.15678/EBER.2023.110302>
- Kuhail, M.A., Alturki, N., Alramlawi, S. & Alhejori, K. (2022). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, [online] 28(1). doi:<https://doi.org/10.1007/s10639-022-11177-3>
- Liu, J., Wang, C. & Liu, S. (2023). Utility of ChatGPT in clinical practice (Preprint). *Journal of Medical Internet Research*, [online] 25(e48568). doi:<https://doi.org/10.2196/48568>
- Lo, L.S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, [online] 49(4), p.102720. doi:<https://doi.org/10.1016/j.acalib.2023.102720>
- Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J.M. & López-Meneses, E. (2023). Impact of the Implementation of ChatGPT in Education: A Systematic Review. *Computers*, [online] 12(8), p.153. doi:<https://doi.org/10.3390/computers12080153>
- Naveed, H., Ullah Khan, A., Qiu, S., Saqib, M., Anwar, S., Usman, M., Barnes, N. & Mian, A. (2024). *A Comprehensive Overview of Large Language Models*. [online] Tillgänglig via: <https://arxiv.org/pdf/2307.06435.pdf> [Hämtad 9 april 2024]
- Neumann, M., Rauschenberger, M. & Schön, E.-M. (2023). ‘We Need To Talk About ChatGPT’: The Future of AI and Higher Education. [online] *IEEE Xplore*. doi:<https://doi.org/10.1109/SEENG59157.2023.00010> [Hämtad 3 maj 2024]
- Nguyen, A., Ngo, H.N., Hong, Y., Dang, B. & Nguyen, B.-P.T. (2022). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, [online] 28. doi:<https://doi.org/10.1007/s10639-022-11316-w>

- Oates, B.J., Griffiths, M. & McLean, R. (2022). *Researching Information Systems and Computing*. 2nd ed. SAGE.
- OECD (2021). *Teaching as a Knowledge Profession STUDYING PEDAGOGICAL KNOWLEDGE ACROSS EDUCATION SYSTEMS*. [online] p.165. Tillgänglig via: https://www.oecd-ilibrary.org/education/teaching-as-a-knowledge-profession_e823ef6e-en [Hämtad 17 april 2024].
- OpenAI (2022). *Introducing ChatGPT*. [online] OpenAI. Tillgänglig via: <https://openai.com/blog/chatgpt> [Hämtad 26 april 2024].
- OpenAI (2023). *GPT-4*. [online] OpenAI. Tillgänglig via: <https://openai.com/gpt-4> [Hämtad 26 april 2024].
- Pichai, S. & Hassabis, D. (2023). *Introducing Gemini: our largest and most capable AI model*. [online] Google. Tillgänglig via: <https://blog.google/technology/ai/google-gemini-ai/> [Hämtad 26 april 2024].
- Porter, J. (2023). *ChatGPT continues to be one of the fastest-growing services ever*. [online] The Verge. Tillgänglig via: <https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference> [Hämtad 26 april 2024].
- Pozo-Sánchez, S., López-Belmonte, J., Rodríguez-García, A.-M. & López-Núñez, J.-A. (2020). Teachers' digital competence in using and analytically managing information in flipped learning (*Competencia digital docente para el uso y gestión analítica informacional del aprendizaje invertido*). *Culture and Education*, [online] 32(2), pp.213–241. doi:<https://doi.org/10.1080/11356405.2020.1741876>
- Rathore, D.B. (2023). Future of AI & Generation Alpha: ChatGPT beyond Boundaries. *Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal*, [online] 12(1), pp.63–68. Tillgänglig via: <https://www.eduzonejournal.com/index.php/eiprmj/article/view/254/214> [Hämtad 18 mars 2024].
- Ray, P.P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, [online] 3, pp.121–154. doi:<https://doi.org/10.1016/j.iotcps.2023.04.003>
- Recker, J. (2013). *Scientific Research in Information Systems A Beginner's Guide*. Berlin Springer Berlin.
- Rotman, D. (2023). *ChatGPT is about to revolutionize the economy. We need to decide what that looks like*. [online] MIT Technology Review. Tillgänglig via: <https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/> [Hämtad 28 mars 2024].
- Schwab, K. (2016). *The Fourth Industrial Revolution*. Switzerland: World Economic Forum, p.172.
- Spataro, J. (2023). *Introducing Microsoft 365 Copilot – your copilot for work*. [online] The Official Microsoft Blog. Tillgänglig via: <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/> [Hämtad 26 april 2024].
- Staab, R., Vero, M., Balunović, M. & Vechev, M. (2024). Beyond Memorizations: Violating Privacy via Inference with Large Language Models. [online] *International Conference on Learning Representations*. Tillgänglig via: <https://arxiv.org/pdf/2310.07298> [Hämtad 9 maj 2024].
- Stålberg, P. (2023). *KUF 'policy statement' om ChatGPT och annan generativ AI - Uppsala universitet*. [online] www.uu.se. Tillgänglig via: <https://www.uu.se/medarbetare/vetenskapsomrade/medicin-och-farmaci/nyheter/arkiv/>

- [2023-12-13-kuf-policy-statement-om-chatgpt-och-annan-generativ-ai](#) [Hämtad 20 mars 2024].
- Sullivan, M., Kelly, A. & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Research Gate*, [online] 6(1), p.10. doi:<https://doi.org/10.37074/jalt.2023.6.1.17>
- Tate, T., Doroudi, S., Ritchie, D., Xu, Y. & Warschauer, M. (2023). *Educational Research and AI-Generated Writing: Confronting the Coming Tsunami*. [online] *EdArXiv Preprints*, p.20. Tillgänglig via: <https://osf.io/preprints/edarxiv/4mec3> [Hämtad 9 maj 2024].
- Trust, T., Whalen, J. & Mouza, C. (2023). Editorial: ChatGPT: Challenges, Opportunities, and Implications for Teacher Education. *Contemporary Issues in Technology and Teacher Education*, [online] 23(1), pp.1–23. Tillgänglig via: <https://citejournal.org/wp-content/uploads/2023/02/v23i1editorial1.pdf> [Hämtad 2 maj 2024].
- van der Wijngaart, W.M. (2023). *KTH | ChatGPT - pragmatic guidelines for students - September 2023 | Wouter Metsola van der Wijngaart*. [online] www.kth.se. Tillgänglig via: <https://www.kth.se/profile/wouter/page/chatgpt-pragmatic-guidelines-for-students-september-2023> [Hämtad 20 mars 2024].
- Van Wyk, M.M. (2024). Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education. *Journal of applied learning and teaching*, [online] 7(1). doi:<https://doi.org/10.37074/jalt.2024.7.1.15>
- Vanhourenout, G., Noyens, D., Gijbels, D. & Van den Bossche, P. (2014). The Relationship between Workplace Climate, Motivation and Learning Approaches for Knowledge Workers. *Vocations and Learning*, [online] 7(2), pp.191–214. doi:<https://doi.org/10.1007/s12186-014-9112-1>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł. & Polosukhin, I. (2017). *Attention Is All You Need*. [online] Tillgänglig via: <https://arxiv.org/pdf/1706.03762.pdf> [Hämtad 26 mars 2024].
- Vuorikari, R., Kluzer, S. & Punie, Y. (2022). *DigComp 2.2 - The Digital Competence Framework for Citizens*. [online] JRC Publications Repository. Tillgänglig via: <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415?fbclid=IwAR2oU4AP-0aj8mp8sMfTLRgvC9ZpuaO7a942d1b8UC3YCnC0bdnu88-G5XY> [Hämtad 8 mars 2024].
- Wang, B., Deng, X. & Sun, H. (2022). *Iteratively Prompt Pre-trained Language Models for Chain of Thought*.
- Wasell, C. (2023). Så hjälper ChatGPT lärare och elever – ”En studiekompis och elevassistent på samma gång”. *Dagens Nyheter*. [online] oktober 2024. Tillgänglig via: <https://www.dn.se/ekonomi/sa-hjalper-chatgpt-larare-och-elever-en-studiekompis-och-elevassistent-pa-samma-gang/> [Hämtad 20 mars 2024].
- Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L.A., Rimell, L., Isaac, W. & Haas, J. (2022). Taxonomy of Risks posed by Language Models. *2022 ACM Conference on Fairness, Accountability, and Transparency*. [online] doi:<https://doi.org/10.1145/3531146.3533088>.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. & Schmidt, D. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT*. [online] Tillgänglig via: <https://arxiv.org/pdf/2302.11382.pdf> [Hämtad 26 mars 2024].

- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y. & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, [online] 55(1), pp.90–112. doi:<https://doi.org/10.1111/bjet.13370>.
- Zhou, H., Nova, A., Larochelle, H., Courville, A., Neyshabur, B. & Sedghi, H. (2022). *Teaching Algorithmic Reasoning via In-context Learning*. [online] Tillgänglig via: <https://arxiv.org/abs/2211.09066> [Hämtad 9 april 2024].