

Lunds universitet  
Sociologiska institutionen

## **Samhällets interaktion med digitala enheter**

*En inblick i de etiska och samhällsliga dimensionerna av artificiell  
intelligens*

**Författare:** Astrid Boström Andersson

**Kandidatuppsats:** SOCK10, 15 hp

Höstterminen 2023

**Handledare:** Axel Fredholm

**Examinator:** Olle Frödin



**Författare:** Astrid Boström

**Titel:** Samhällets interaktion med digitala enheter: en inblick i de etiska och samhällsliga dimensionerna av artificiell intelligens

**Kandidatuppsats:** SOCK10, 15 hp

**Handledare:** Axel Fredholm

**Sociologiska institutionen, höstterminen 2023**

**Antal ord:** 10 228

## **Abstrakt**

Studien fokuserar på att utforska de samhällsliga implikationerna och etiska aspekterna som är förknippade med framsteg inom artificiell intelligens, som uppfattas av ett antal studenter inom civilingenjörsprogrammet i datateknik. Studien utreder dessa studenters perspektiv på AI:s samhällsliga roller och deras egna etiska ansvar i utvecklingen av AI-modeller. Studien undersöker dessa studenters perspektiv på AI:s samhällsliga roller och deras egna etiska ansvar i utvecklingen av AI-modeller. Studien syftar till att förstå konsekvenserna av AI-framsteg genom att analysera dem utifrån symbolisk interaktionism och habitus. Vidare undersöker studien specifikt hur dessa framsteg påverkar attityder till etik och varsamhet. Slutsatsen lyfter fram betydelsen av att erkänna AI som en självständig enhet inom den sociala sfären, styrd av samma sociologiska principer som den fysiska. Tillvägagångssättet understryker detta ömsesidiga beroende när det gäller att influera samhällsprocesser. Den tillämpade metoden för studien var kvalitativ där datainsamlingen utgick från semistrukturerade intervjuer. Denna ansats ger betydande perspektiv på elevernas kontemplationer angående de samhällsliga och etiska implikationerna av AI. Studiens kvalitativa form möjliggör en fördjupad och subjektiv granskning av studenternas attityder och åsikter, vilket ger insikt om hur deras åsikter påverkas av deras utbildning och bredare sociokulturella faktorer. Att förstå de komplexa och mänskliga aspekterna av hur framtida teknologer uppfattar och påverkar ämnet AI är av stor vikt, då dess stringens är multidimensionell.

**Nyckelord:** Artificiell Intelligens (AI), Samhällspåverkan av AI, Symbolisk Interaktionism, Habitus, Etiska Implikationer inom AI, Interdisciplinära tillvägagångssätt mot AI, Etik

# Innehållsförteckning

<b>ORDLISTA .....</b>	<b>1</b>
<b>1. INLEDNING.....</b>	<b>1</b>
1.1 <i>Framsteg och utveckling av AI.....</i>	1
1.2 <i>AI:s sociala inverkan .....</i>	3
1.3 <i>Etiska aspekter vid AI.....</i>	3
1.4 <i>Civilingenjörutbildning i datateknik vid Lunds Tekniska Högskola (LTH).....</i>	4
1.5 <b>PROBLEMFÖRMULERING .....</b>	4
1.6 <b>SYFTE OCH FRÅGESTÄLLNINGAR .....</b>	5
1.6.1 <i>Frågeställningar .....</i>	6
1.7 <b>AVGRÄNSNINGAR.....</b>	6
1.8 <b>SOCIOLOGISK RELEVANS.....</b>	7
<b>2. METOD.....</b>	<b>8</b>
2.1 <b>KVALITATIVA METODER.....</b>	8
2.1.1 <i>Metodologiska ansatser .....</i>	8
2.2 <b>BEGRÄNSNINGAR VID KVALITATIV METOD .....</b>	9
2.3 <i>Tillförlitlighet och äkthet.....</i>	10
2.4 <b>INTERVJU .....</b>	12
2.4.1 <i>Förberedelse.....</i>	12
2.4.2 <i>Utformning av intervjufrågor .....</i>	12
2.4.3 <i>Val av intervjupersoner.....</i>	13
2.4.4 <i>Genomförande.....</i>	14
2.5 <b>BEARBETNING AV MATERIAL &amp; DATAANALYS .....</b>	15
2.5.1 <i>Transkribering.....</i>	15
2.5.2 <i>Dataanalys.....</i>	16
2.6 <b>ETISKA ASPEKTER .....</b>	17
<b>3. TIDIGARE FORSKNING .....</b>	<b>19</b>
3.1 <i>ALGORITHMIC INTIMACY - ANTHONY ELLIOT (2022) .....</i>	19
3.2 <i>MACHINE HABITUS - TOWARDS A SOCIOLOGY OF ALGORITHMS - MASSIMO AIROLDI (2021).....</i>	19
3.3 <i>THE 4TH REVOLUTION - HOW THE INFOSPHERE IS RESHAPING HUMAN REALITY - LUCIANO FLORIDI (2014).....</i>	19
3.4 <i>SOCIOLOGICAL INSIGHT: AN INTRODUCTION TO NON-OBVIOUS SOCIOLOGY - RANDALL COLLINS (1992).....</i>	20
3.5 <i>PÅ AI-TEKNIKENS AXLAR - OM KUNSKAPSSOCIOLOGIN OCH STARK ARTIFICIELL INTELLIGENS - PETER KÅRES (2009).....</i>	20
3.6 <i>MANAGING BIAS IN AI (2019).....</i>	21
<b>4. TEORETISK RAM.....</b>	<b>21</b>
4.1 <b>HERBERT BLUMER - SYMBOLISK INTERAKTIONISM.....</b>	21
4.1.1 <i>Samhällsanalys &amp; Premisser .....</i>	22
4.1.2 <i>Relevans .....</i>	22
4.2 <b>PIERRE BOURDIEU - HABITUS .....</b>	23
4.2.1 <i>Relevans .....</i>	23

<b>5. ANALYS OCH RESULTAT .....</b>	<b>24</b>
5.1 BIAS OCH SOCIALA STRUKTURER .....	24
5.2 INTERAKTIONER .....	27
5.3 UTBILDNINGENS PÅVERKAN .....	30
5.4 ETISKA ASPEKTER .....	32
5.4.1 Förutsättningar och insikter från studierna.....	32
5.4.2 Yttre påverkan, förutsättningar och medvetenhet .....	34
5.5 FRAMTID .....	36
<b>6. SLUTSATS OCH VIDARE FORSKNING .....</b>	<b>37</b>
6.1 ETIK.....	37
6.2 BIAS OCH HABITUS.....	38
6.3 AI:S ROLL I SAMHÄLLET .....	38
6.4 SYMBOLISK INTERAKTIONISM .....	39
6.5 TVÄRVETENSKAPLIGA LÖSNINGAR.....	39
6.6 VIDARE FORSKNING OCH SLUTREFLEKTIONER.....	39
<b>REFERENS- OCH LITTERATURLISTA.....</b>	<b>40</b>
<b>BILAGOR.....</b>	<b>42</b>
BILAGA 1: INFORMATIONSBREV .....	42
BILAGA 2: BETECKNING AV INFORMANTER.....	43
BILAGA 3: TRANSKRIBERING.....	43
BILAGA 4: INTERVJUGUIDE .....	44

## **Ordlista**

**AI** - Artificiell intelligens.

**Data** – Information, de ämnen eller element som används i processen för att utbilda AI.

**Input** – Data som skickas in till AI-modellen.

**Output** - Resultaten som produceras av AI.

**OpenAI** - Företag som specialiserar sig på att utveckla AI-modeller.

**Chat GPT** - En AI-modell som emulerar mänsklig intelligens som är speciellt byggd för att förstå och generera språk.

**Deep Learning** - Teknik som används för att träna AI för att likna mänskligt beteende.

**Machine Learning** - Del inom AI som involverar träning av en dator med hjälp av redan existerande data.

**Generativ AI** - AI-modeller som är specifikt utformade för att generera nya data, inklusive men inte begränsat till språk, text, foton, videor eller ljud.

**Sociala medier** - En grupp webbplattformar och program som specifikt vänder sig till mänskliga sociala nätverk.

## 1. Inledning

*“Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.” (Center for AI Safety, 2023)*

Ett öppet brev, undertecknat av mer än 350 företagsledare inom området artificiell intelligens, utfärdades i mitten av 2023 av den ideella gruppen "Center for AI Safety". Undertecknarna omfattar framstående chefer från de ledande företagen som har spelat en avgörande roll i utvecklingen av artificiell intelligens, i synnerhet inklusive verkställande direktören för OpenAI, organisationen som ansvarar för att etablera Chat GPT. Pionjärer inom AI, som har erkänts för sina exceptionella prestationer inom AI-utveckling, har också ställt sig bakom insamlingskampanjen. Organisationen beskriver många viktiga problem förknippade med AI, inklusive illvilligt utnyttjande, AI-konkurrens, organisatoriska risker och förlust av kontroll.

*“I want to suggest that if a real artificial intelligence (AI) is going to be built, sociologists will have to play a major part in it.” (Collins, 1992:163)*

### 1.1 Framsteg och utveckling av AI

Artificiell intelligens är ett ämne inom datavetenskapen som skapar och förbättrar programvara för att replikera mänskligt intellekt och beslutsfattande. För att AI ska efterlikna mänskligt beteende krävs maskininlärning med olika datauppsättningar, som ett resultat av detta kan en AI kopiera mönster och beteenden och utvecklas (McCarthy, 2007).

AI-system använder endast tidigare inlärd data. Eftersom syftet med AI är att spegla mänskligt intellekt, kommer modellens agerande att återspegla mänskliga ideal och fördomar. Även om AI överträffar människor i sitt avsedda syfte, betraktas det vanligtvis som spjutspetsteknologi. AI

har visat sig vara fördelaktigt för människor, dock är det viktigt att förstå att AI inte kan lösa alla problem. Då AI:s syfte, särskilt generativ AI, är att efterlikna mänskliga resonemang är den benägen att göra misstag (McCarthy, 2007).

*“The same elements and techniques that power the socio-economic benefits of AI can also bring about new risks or negative consequences for individuals or the society.” (AI ACT, 2023)*

Olika former av AI finns, där generativ AI är den vanligaste, vilket är en form av AI som genererar nya verk. Chat GPT är den vanligast förekommande generativa AI som används idag, en AI modell som skapats av företaget OpenAI, som tillkännagavs offentligt i november 2022. Den deltar i dialog och använder djupinlärning, en delmängd av maskininlärning, för att producera text eller information som liknar mänskliga svar. GPT tjänar flera syften och är inte exklusivt för AI-experter - den har förmågan att generera ett brett utbud av utdata, inklusive receptidéer, e-postskapande, textsammanfattning, kodlösning och många andra uppgifter. De potentiella applikationerna för denna Chatbot är praktiskt taget obegränsade, och dess tillgänglighet för alla användare, kostnadsfritt, förbättrar användningen ytterligare (Chat GPT release Notes, 2022)

En annan prevalent typ av AI är integrationen av AI-algoritmer på sociala medieplattformar. Systemet är tränat för att ge personligt anpassat material baserat på individens interaktioner, delad information och användning. Därför kan en människa ofta stöta på artificiell intelligens utan att medvetet känna igen dess närvaro, genom att datorn förser personer med information som den anser vara relevant för den person som använder enheten (Airoldi, 2021).

AI:s snabbt växande framkomst oroar inte endast yrkesverksamma - att experterna inom detta område aktivt varnar för det, höjer många ögonbryn. Ett stort orosområde är hur data samlas in, och om individerna vars data används, är medvetna om detta. Oron för AI har väckts genom den

snöbollseffekt som uppstår, när AI:n är tillräckligt smart kräver den inte en människa som lär den vad den ska göra, och vissa argumenterar för att AI:n själv redan blivit smartare än människor. AI:n blir således oberäknelig, både för gemene man och ingenjören som utvecklat den. En viktig fråga att överväga är metodiken för datainsamling och medvetenhetsnivån hos personer vars data används. Således skapas en situation där användaren är fångad i ett paradoxalt dilemma (ibid).

## **1.2 AI:s sociala inverkan**

Eftersom AI har blivit så pass utbredd i samhället är det givet att det har förekommit en social inverkan. Bara de senaste åren har sociologer börjat diskutera AI och hur det påverkar interpersonella relationer samt samhället som stort. Airoidi (2021) menar att AI är så pass inflytelserik att tidigare sociologiska begrepp inte räcker till för att beskriva dess ställning i den sociala världen. Elliot (2023) delar detta perspektiv och belyser liknande synsätt och argumenterar för att algoritmernas sociala inverkan måste vara ett centralt motiv i tjugohundratalets sociala teori.

## **1.3 Etiska aspekter vid AI**

AI:s framsteg har också föranlett etiska reflektioner, kontroversen kring utvecklingen och produktionen av AI härrör från den begränsade omfattningen av kunskap som dessa system besitter. Detta eftersom AI:n kan endast förbättra sin hastighet, intelligens och kapacitet genom tillhandahållen information. Europeiska kommissionen lanserade AI HELG, en autonom expertgrupp, 2018. Dess syfte är att tillhandahålla etiska principer som syftar till att främja AI:s pålitlighet. Rekommendationerna är framdragna från tre nyckelprinciper:

- 1. den bör vara laglig och garantera respekt för alla gällande lagar och förordningar,*
  - 2. den bör vara etisk och säkerställa efterlevnad av etiska principer och värden, och*
  - 3. den bör vara robust, både ur tekniskt och samhällelligt perspektiv eftersom AI-system även med de bästa intentioner kan orsaka oavsiktliga skado*
- (AI HELG, 2018)*



Organisationen betonar att etik ska vara den grundläggande principen inom mjukvaruutveckling. Strategin tjänar flera mål, som omfattar individ-, grupp- och samhällsdimensioner. Dessutom klarläggs det att grundläggande undervisning i etik för artificiell intelligens är avgörande för alla inblandade parter, särskilt för utvecklarna (*AI HELG, 2018*).

#### **1.4 Civilingenjörsutbildning i datateknik vid Lunds Tekniska Högskola (LTH)**

*“Vill du vara med och utveckla [ett] av framtidens digitala system?”/.../ (LTH, 2023)*

Civilingenjörsutbildningen inom datateknik vid LTH är en femårig högskoleutbildning som omfattar 300 högskolepoäng. Programmets tre första år ger en grundläggande behörig teknisk utbildning, som avslutas av en tvåårig specialisering där studenterna kan välja en av åtta programspecifika specialiseringar. De vanligaste jobben efter examen är programmerare, systemutvecklare och IT-specialister (LU, 2023).

#### **1.5 Problemformulering**

Studenterna på civilingenjörsutbildningen i datateknik på LTH representerar de kommande kohorterna av mjukvaruingenjörer, och det finns redan en observerbar förändring när det gäller diskussionen om etiska överväganden i mjukvaruutveckling. Även om kurser i etik är tillgängliga, är dessa frivilliga och ges endast under det sista året av programmet. Vid detta skede har studenterna redan befintlig kompetens att utveckla AI-modeller

Det är avgörande att analysera elevernas perspektiv på etiska frågor för att förstå de potentiella sociala återverkningar som kan uppstå av oaktsamt handlande och försummelse av effekterna av AI, eftersom attityder driver handlingar.

När skaparna av AI-modellerna besitter en begränsad förståelse och uppvisar en likgiltighet för dess konsekvenser kan teknologier utvecklas vars inverkan på individer verkar vara minimal och harmlös. Däremot kan dessa modeller få mer utbredda konsekvenser på samhället i det långa loppet.

Värderingar och fördomar överskrider observerbara sociala beteenden och blir alltmer genomgripande i den samtida kulturen. AI spelar en viktig roll i denna motsägelse, eftersom individer interagerar med den som om den vore en annan människa, en källa till feedback eller en personlig medhjälpare. Dessa observationer tyder på nödvändigheten av att analysera dynamiken och deras samhällsliga konsekvenser.

Den digitala miljön bör analyseras med samma sociologiska teorier som den fysiska världen, eftersom alla element i detta kontinuum är sammankopplade. Denna studie kommer självständigt att utvärdera flera områden, vilket visar den omfattande karaktären av denna kedja.

### **1.6 Syfte och frågeställningar**

Studiens syfte var att använda sociologiska teorier, specifikt symbolisk interaktionism och Habitus, för att förstå perspektiven hos kommande AI-utvecklare angående implikationerna av AI-framsteg. Detta inkluderar att undersöka deras attityder till etik och försiktighet och hur dessa attityder implementeras i samhället i stort. Studien kommer också att undersöka sambanden mellan mänskliga interaktioner och interaktioner som underlättas av datorsystem som reglerar informationsflödet. Följaktligen bildar det en omfattande sekvens av händelser, val och fenomen som påverkar samhället i dess helhet, även om alla i slutändan kan tillskrivas en enda underliggande faktor; utvecklaren som ansvarar för att skapa denna speciella typ av programvara.

### **1.6.1 Frågeställningar**

Studien har utgått från följande huvudfrågeställning:

- Vilken påverkan har AI-utvecklarens etiska attityder och fördomar, speciellt ingenjörsstudenter vid datateknik vid Lunds universitet, på utvecklingen och samhällsliga påverkan av AI som social agent?

För att uppfylla studiens syfte har även följande underfrågeställningar tillämpats:

- Hur kan begreppet symbolisk interaktionism hjälpa oss att förstå elevernas attityder till de samhällsliga konsekvenserna av AI, och hur speglar dessa föreställningar deras habitus som framtida AI-utvecklare?
- Hur påverkar AI och maskininlärning sociala strukturer, interpersonella relationer och mänskligt beteende?

### **1.7 Avgränsningar**

Studien syftade till att fastställa elevernas perspektiv på effekten av AI-användning och framsteg på samhället, samt de etiska överväganden som är förknippade med AI-utveckling.

Beslutet att välja ut datateknikstudenter vid LTH var motiverat på grund av deras omfattande kunskaper och expertis inom teknik, som förvärvas genom deras utbildningsval. Att tillämpa ett sociologiskt perspektiv på detta empiriska arbete var genomförbart då informanternas beteendemönster förblir opåverkad av samhällsvetenskapliga perspektiv. Detta möjliggjorde en mer nyanserad granskning och diskurs om samspelet mellan teknik och sociala faktorer.

Lunds universitets geografiska gränser kan tyckas vara begränsade och jämförelsen av studenter inskrivna på samma program vid andra universitet kan ha visat olika resultat. Men på grund av intervjupersonernas olika härkomst, som kommer från olika platser i Sverige och med distinkt

socioekonomisk och sociokulturell bakgrund, ansågs det lämpligt att begränsa studien till Lunds universitet.

Forskningen undersökte i första hand syn på etik och social påverkan, genom intervjufrågor, teorier och analyser. Undantagna var andra perspektiv på frågor som tekniska specifikationer eller ekonomiska aspekter.

### **1.8 Sociologisk relevans**

Det finns ingen binär motsättning mellan teknik och samhälle, primärt inriktad på artificiell intelligens och maskininlärning. Det finns en divergens av perspektiv på dess betydelse i sociologiska studier. Syftet med denna studie är att identifiera och förstå sambandet mellan maskin- och mänsklig kontroll, och dess implikationer för samhället i stort. Studien belyser vikten av att införliva ett sociologiskt perspektiv i utvecklingen av AI, vilket betonats av traditionella sociologer.

## 2. Metod

Följande avsnitt kommer att beskriva, problematisera, förklara och presentera val av metod för att sedan kartlägga studiens tillvägagångssätt.

### 2.1 Kvalitativa metoder

Denna studie tillämpade en kvalitativ metod för att undersöka attityder via ord snarare än numeriska data. Kvalitativ forskning undersöker samband, attityder och perspektiv av forskningsobjektet för fördjupad förståelse (Bryman 2016). Denna studie koncentrerade sig på immateriella åsikter, synpunkter och fenomen, vilket var viktigt för forskningsfrågorna, varför en kvalitativ teknik var motiverad (ibid; Larsson, 2011:7).

#### 2.1.1 Metodologiska ansatser

Studien conceptualiserades med hjälp av en kombination av kvalitativa metodologiska ansatser.

- *Induktiv ansats:* Intervjudatans specifika observationer användes för att göra allsidiga generaliseringar. Detta visas i studiens analysdel, där tematiseringen och analysen av informanternas svar är kopplade till forskningsfrågorna.
- *Abduktiv ansats:* Den abduktiva strategin involverade att använda en blandning av induktion och deduktion, känd som abduktion, för att formulera och införliva idéer som överensstämde med empiriska data. Detta åstadkoms genom att modifiera befintliga idéer för att passa de specifika omständigheterna i utredningen.
- *Narrativ ansats:* Även om det narrativa tillvägagångssättet inte har varit lika framträdande, har det ändå varit en viktig funktion för studien. Detta beror på att empirin presenteras i ett narrativt format.

Studien samlade in data genom att genomföra semistrukturerade intervjuer med fem informanter. Detta möjliggjorde en grundlig förståelse av potentialen, begränsningarna och omständigheterna

i det valda sociala sammanhanget. Den främsta fördelen med semistrukturerade intervjuer är deras förmåga att ge värdefulla insikter i undersökningsdeltagarnas subjektiva upplevelser. Att genomföra semistrukturerade intervjuer ger flexibilitet, eftersom informantens uppfattning och tolkning av frågor och händelser påverkas av samtalstonen (Bryman, 2016).

## **2.2 Begränsningar vid kvalitativ metod**

Trots att den kvalitativa metoden anses vara väl motiverad för denna studie har den sina begränsningar. Det är avgörande att erkänna utmaningen med att extrapolera samhällliga effekter från ett visst urval, såsom LTH-studenter, i denna studie. Även om tillämpningen av sociologiska begrepp kan vara fördelaktigt för samhället, är det viktigt att notera att resultaten från en sådan studie inte bör anses vara universellt tillämpliga på grund av det begränsade omfånget av undersökningspopulationen. Denna metod gör det möjligt för att samtidigt överväga två begrepp: attityderna är specifika för en viss gemenskap, medan en uppfattning om AI:s samhällliga påverkan kan härledas. Det är dock tillrådligt att vara försiktig när slutsatser dras baserade på en begränsad urvalsstorlek, eftersom detta också ger möjligheter till ytterligare forskning.

### **2.2.1 Objektivitet**

En framträdande begränsning för kvalitativ forskning är utmaningen att upprätthålla opartiskhet, då den grundläggande principen för denna metod är att *tolka* empiriska fakta. Detta kan beskrivas som behovet av ökad medvetenhet och omsorg för att bibehålla största möjliga opartiskhet (Bryman, 2016).

Vid användning av identiska intervjuguider är det troligt att två distinkta forskare kan ge olika resultat som ett resultat av variationer i intervjuarens personlighet och intervjumetoder. Det är avgörande att ha detta i åtanke och sträva efter att genomföra intervjun med största neutralitet (Barbour, 2008: 114).

### **2.2.2 Liten undersökningspopulation**

Kvalitativa studier har en mindre undersökningspopulation jämfört med kvantitativ forskning, vilket resulterar i en stor begränsning av generaliserbarheten. Därför är det viktigt att vara varsam vid val av respondenter, för att kunna garantera en varierad empiri (Bryman, 2016). Med tanke på den redan begränsade omfattningen av studien, som enbart fokuserade på vissa studenter, i ett specifikt program och på en specifik plats, var det avgörande att inkludera en mångsidig och varierad grupp. På grund av studiens tidsomfång var det dock inte möjligt att uppnå fullständig säkerhet för att uppnå detta mål.

### **2.2.3 Tidskrävande**

Trots att den kvalitativa metoden innehåller betydligt färre respondenter än den kvantitativa, går det att argumentera för att den är mer tidskrävande. Detta eftersom transkribering tar lång tid (Bryman, 2016). Men en tidskrävande metod innebär även en djup och informationsrik empiri, som kanske inte är möjlig med andra metoder.

### **2.2.4 Svårt att replikera**

En annan begränsning är dess svårighet att replikera resultaten, eftersom resultaten, tolkningarna och slutsatserna i sig är unika (Bryman, 2016). Däremot kan resultatens unika natur inspirera framtida forskning och bidra till nyanserade perspektiv i forskningsgemenskapen.

## **2.3 Tillförlitlighet och äkthet**

Den kvalitativa metoden är baserad på upplevelser, vilket gör empirin subjektiv och inte reella fakta. Lincoln & Guba (1985) och Guba & Lincoln (1994) hävdar att andra mätbegrepp är avgörande för att bedöma kvaliteten på en studie. I kvantitativ forskning är validitet och tillförlitlighet begrepp som är mer relevanta. Som ett resultat utgår denna studie från att garantera "tillförlitlighet" och "äkthet", eftersom det bedöms vara ett mer korrekt ramverk. Däremot är "äkthetens" genomslag och dess erkännande inte lika utbrett som tillförlitligheten, och därför kommer den att exkluderas som parametrar för denna studie (Bryman, 2016).

Tillförlitligheten går att delas in i fyra delkriterier (Bryman, 2016: 467).

### **2.3.1 Trovärdighet ("Credibility")**

Det första kriteriet är mycket likt respondentvalidering, vilket innebär att informanterna får vara delaktiga i forskningsprocessen och får ta del av den information de bidragit med för att säkerställa att allt har uppfattats korrekt. När informanterna accepterar forskarens tolkning av deras sociala verklighet upprätthålls trovärdigheten. Samtliga informanter ombads att recitera transkriberingen, samt att dessa varit tillgängliga under studiens gång för att delta i forskningsprocessen. Kriteriet uppfylls även när studien utgått i linje med reglerna och ramverk för forskningen (ibid).

### **2.3.2 Överförbarhet ("Transferability")**

Det andra kriteriet beskriver möjligheten att genomföra en jämförbar studie i en annan miljö, i en liknande grupp eller omständigheter (ibid). Denna studie kan återskapas genom att undersökningspopulationen skiftas från studenter på samma universitet i Sverige och runt om i världen. Således bevaras överförbarheten.

### **2.3.3 Pålitlighet ("Dependability")**

Det tredje delkriteriet syftar till skapandet av en fullständig och tillgänglig redogörelse för alla faser av forskningsprocessen. Transparens via potentiella förändringar, samt gentemot kollegor och kamrater bygger ett förtroende för resultaten (ibid). Genom kontinuerlig handledning under forskningens fas, med en handledare som även granskat enskilda avsnitt i studien, har det tredje kriteriet styrkts.



### **2.3.4 En möjlighet att styrka och konfirmera (“Confirmability”)**

Det fjärde och sista delkriteriet understryker att kvalitativ forskning aldrig kommer att vara helt objektiv, därför måste forskaren göra sitt allra yttersta för att undvika att subjektiva åsikter och värderingar påverkar eller äventyrar forskningsresultaten. För att undvika det ska forskaren agera i god tro och ej medvetet påverkat studiens resultat eller slutsatser (Bryman, 2016).

## **2.4 Intervju**

För att uppnå studiens syfte att undersöka elevernas etiska medvetenhet och attityder kring AI-utveckling användes ett kvalitativt tillvägagångssätt med semistrukturerade intervjuer. Genom att genomföra intervjuer med studenter på olika stadier av civilingenjörsprogrammet i datateknik vid LTH.

### **2.4.1 Förberedelse**

Ejvegård (2009: 51–52) framhåller vikten av att vara väl förberedd och ha en grundlig förståelse av studieämnet innan empirin samlas in. Med tanke på den induktiva karaktären av denna studie var det däremot avgörande att undvika överdriven förtrogenhet med ämnet. I stället låg fokus på att ställa relevanta följdfrågor som ger empiri som kunde besvara forskningsfrågorna. På grund av den tidskrävande karaktären av intervjuer som en datainsamlingsteknik, är det avgörande att vara noggrann vid val av deltagare och att förbereda sig på förhand. (ibid)

### **2.4.2 Utformning av intervjufrågor**

Inför datainsamlingen togs en semistrukturerad intervjuguide fram. Bryman (2016) föreslår flera tillvägagångssätt som bör användas vid utformning av en intervjuguide. Dessa var till underlag för samtliga intervjuer (Se Bilaga 4). Denna togs fram utifrån studiens syfte och forskningsfrågor. Den specifika utformningen av frågorna har dock varit flexibel under skapandet av guiden för att möjliggöra inkludering av alternativa synpunkter eller vinklar som kan uppstå från den insamlade informationen.

Till skillnad från enkätfrågor är intervjuguiden relativt kort och syftar till att bland annat påminna forskaren att ta upp frågor eller ämnen som annars missats i intervjun (Barbour, 2008: 115). Frågorna får inte heller vara ledande, och i stället vara raka eller direkta (Bryman, 2016; Ejvegård, 2009: 54).

### **2.4.3 Val av intervjupersoner**

När informanterna rekryterades togs hänsyn till en mängd olika egenskaper för att garantera att intervjuerna skulle ge en tillräckligt diversifierad, nyanserad och fullständig empirisk grund och teoretisk mättnad. Totalt valdes fem informanter ut, varav tre är på sitt tredje studieår och de återstående två på sista och femte studieåret. Olika urvalsmetoder användes, där användningen av den stratifierade urvalsmetoden ansågs lämplig på grund av inkluderingen av informanter som hade olika akademisk bakgrund. För att få olika synpunkter och potentiella variationer i attityder, stratiserades informanterna i distinkta grupper. Följaktligen användes även en ändamålsenlig urvalsmetod (Bryman, 2016).

Att införliva fokusgrupper i studien kan ha varit fördelaktigt på grund av den mestadels homogena undersökningspopulationen. Detta beror på att individer tenderar att känna sig mer tillfreds och är mer benägna att dela sina perspektiv och åsikter i en miljö där de är omgivna av likasinnade (Wibeck, 2011). Det skulle därför ha varit avgörande att inkludera en fokusgrupp i denna studie med tanke på att deltagarna enbart engagerade sig i en-till-en-interaktioner. För att upprätthålla denna känsla av engagemang under intervjuer är en semistrukturerad intervjumetod lämplig. Denna strategi möjliggör uppföljningsfrågor, främjar dynamiska diskussioner och skapar en samtalston i intervjun. Det finns fördelar och nackdelar med denna situation, när informanten känner sig tillfreds medger det insamling av en större mängd data, samt data som besitter en bättre grad av noggrannhet och autenticitet. Icke desto mindre är det avgörande för en intervjuare att vara försiktig och undvika att ställa ledande frågor, särskilt i semistrukturerade intervjuer när uppföljningsfrågor är naturliga för intervjuprocessen. Vidare är det avgörande att låta informanten uttrycka sig och vara uppmärksam (Barbour, 2008: 113).

Vem	Relevans	Datum	Längd
Intervjuperson 1	Student i åk 3	23/11	43 min
Intervjuperson 2	Student i åk 3	23/11	65 min
Intervjuperson 3	Student i åk 3	23/11	73 min
Intervjuperson 4	Student i åk 5	23/11	43 min
Intervjuperson 5	Student i åk 5	27/11	54 min

Figur 1, Tabell över informanter, datum för intervju samt tidsåtgång.

#### 2.4.4 Genomförande

Intervjuerna gjordes via Zoom, som varade från 40 till 70 minuter, vilket anges i tabellen ovan. På begäran fick informanterna möjlighet att antingen delta i intervjun digitalt eller personligen, samtliga valde att genomföra intervjun digitalt. Beslutet att genomföra intervjuerna online togs för deltagarnas bekvämlighet och intervjuerna gjordes med kameran aktiverad, vilket minimerar skillnaden mellan ett virtuellt möte och ett fysiskt möte (Bryman, 2016).

Möjligheten att stanna kvar i sitt eget hem kan sänka tröskeln och ge en mer bekväm atmosfär under hela intervjun. Bryman (2016) menar att det finns en skillnad på resultat av intervjuer beroende på om den genomförs digitalt eller inte. Internetintervjuer ger fördelen av att vara tidseffektiva, lämpligheten av det digitala formatet fastställdes utifrån de tillfredsställande svar som gavs under de två första intervjuerna. Under alla intervjuer togs hänsyn till noggrant övervägande och medvetenhet om potentiella utmaningar i samband med internetintervjumetoden. Dessutom uttrycktes konsekvent en beredskap och öppenhet att byta till fysiska möten under hela intervjuprocessen. Förespråkare av kvalitativa intervjuer föreslår att man gör intervjuer ansikte mot ansikte. Men med tanke på att syftet med studien var att få en djupare förståelse av attityder och erfarenheter, ansågs det fördelaktigt att låta deltagarna själva

välja plats och tidpunkt för sitt engagemang. Även om dessa faktorer inte direkt påverkade resultatet av intervjuerna, ansågs förmågan att tillgodose intervjupersonernas preferenser vara betydande (Sturges & Hanrahan, 2004; Bryman, 2016).

Innan intervjuerna påbörjades fick alla deltagare information både muntligt och skriftligt om syftet med studien och gjordes medvetna om att deras svar skulle ligga till grund för en kandidatuppsats i sociologi. Innan intervjun inleddes kommunicerades det att deras deltagande är frivilligt och kan avslutas när som helst. Dessutom informerades deltagarna om att deras svar skulle vara helt anonyma och omöjliga att länka tillbaka till dem. Informanterna visade en hög nivå av förståelse och samtycke till att intervjuerna spelades in och sedan transkriberades, eftersom de var vederbörligen informerade inför och under hela intervjun både muntligt och skriftligt (Bryman, 2016).

## **2.5 Bearbetning av material & Dataanalys**

Följande avsnitt beskriver behandlingen samt analysen av det insamlade materialet från intervjuerna.

### **2.5.1 Transkribering**

Att transkribera intervjuerna innebär att ordagrant skriva ner vad som sades i ljudinspelningen. Denna process är en viktig hörnsten i studiens arbete, vilket kräver noggrant utförande och precision. Syftet med denna process är att forskaren ska kunna granska och bekanta sig med innehållet. Dessutom ökar den förståelsen och fungerar som ett slags repetition efter att transkriberingen är klar. Efter att ha transkriberat den inspelade empirin, organiseras den systematiskt genom kodning. Denna process möjliggör identifiering av kopplingar, distinktioner och relevanta svar baserat på frågan, vilket är avgörande att införliva i analysen. Dessutom har analysdelen ordagranna citat utdragna från transkriptionen. Att transkribera innehåll hjälper inte bara författaren utan förbättrar också läsarens förståelse och underlättar sekundär analys under läsningen av analysdelen (Bryman, 2016).

Det finns olika sätt att utföra transkribering, och för denna studie tillämpades en transkribering på mellannivå. Detta innebär att alla detaljer noterades, exklusive eventuella pauser, nöjen eller liknande händelser. Omvänt lades utfyllnadsord som "ehm" och "hm" till på grund av deras upplevda betydelse för att indikera reflektion eller kontemplation kring särskilt tuffa och krävande ämnen (Skärvad & Lundahl, 2016).

Ett hinder som stöttes på vid transkriberingsprocessen är möjligheten att stöta på delar i intervjuerna där missuppfattningar uppstår, vilket leder till notering av felaktiga anteckningar. På grund av transkriberingens tidskrävande karaktär (Bryman, 2016) togs pauser konsekvent under hela processen. För att effektivisera processen och bibehålla intervjuens framträdande karaktär genomfördes transkriptioner omedelbart efter varje session. Denna funktion underlättade snabb granskning av innehållet och gav ett ytterligare sätt att förstå informationen som förmedlades under intervjun. Detta ansågs mycket värdefullt eftersom det gjorde det möjligt att behålla och tillämpa särskilda insikter i efterföljande intervjuer.

Ljudinspelningarna och medföljande pappersarbete för transkriptionen har sparats på ett säkert sätt och kommer att tas bort efter att studien har godkänts. Informanternas anonymitet garanterades, vilket säkerställde att varken ljudinspelningar eller dokumenten innehåller någon identifierande information. Informanterna fick också möjlighet att granska sin transkription i syfte att verifiera riktigheten. Ytterligare uppspelning av ljudfilen görs efter avslutad transkription för att säkerställa att ingen information förbises vid överföringen till uppsatsen.

### **2.5.2 Dataanalys**

Genom att använda en tematisk analysmetod undersöktes empirin för att identifiera och förstå distinkta teman i intervjuerna, vilket underlättade identifieringen av både divergerande och delade teman. Under hela transkriberingsprocessen var det möjligt att identifiera återkommande teman som fanns i alla intervjuer, samt områden där informanterna hade olika åsikter.

Följaktligen kategoriserades intervjuerna efter dessa ämnen. Eftersom dessa teman återfanns i alla intervjuer, om än med olika perspektiv och synsätt, utgjorde de betydande delar av

analysdelen. Teman fastställs i enlighet med forskningsfrågan och studiens syfte, i syfte att identifiera bekräftande empiri (Bryman, 2016).

Dessutom utökades tillvägagångssättet för temaanalys genom att inkludera delar av den fenomenografiska analysmetoden. Den fenomenografiska metoden används ofta i kvalitativ forskning på grund av dess betoning på informanternas subjektiva upplevelser. Genom användning av fenomenografisk analys kan informanternas upplevelser av de uppräknade händelserna beskrivas, kategoriseras och tematiseras (Larsson, 1986).

När man studerar materialet är det avgörande att ta hänsyn till sina egna värderingar och åsikter. Det är möjligt att ha önskade eller förväntade utfall eller svar i relation till vissa resultat från empiriskt arbete, och detta kan påverka hur man tolkar och reagerar på fynden. Det är viktigt att notera att detta inte bör förväxlas med att koda och uppfatta sambandet med forskningsämnet. I stället bör det ses som en försiktighetsåtgärd som måste beaktas under hela forskningsprocessen (ibid: 37).

## **2.6 Etiska aspekter**

Eftersom insamlingen av empirin skedde genom intervjuer med informanternas subjektiva åsikter är det av högsta vikt att följa forskningsetiska principer. Grunden för dessa beskrivs av Lindstedt (2019: 51) som nyttjande-, samtyckes-, informations- samt konfidentialitetskravet.

*Nyttjandekravet* hänvisar till att empirin uteslutande används för forskningssyfte, genom att både muntligt och skriftligt informera innan intervjuerna uppfylls detta krav (ibid, 2019: 55).

*Samtyckeskravet* innebär att informanterna helt och hållet har självbestämmande i deras deltagande, detta berör omständigheter gällande hur intervjun tar plats, eventuellt avbrytande under intervjun eller liknande (ibid). Även detta krav uppfylldes genom både skriftlig och muntlig information och godkännande innan intervjuens gång. Informanterna var väl medvetna om att väl vid start har de givit samtycke, men att det inte innebär att de inte får avbryta

intervjun. Genom att tillåta informanter välja intervjuplats, om den skulle vara fysisk eller digital, kan vara en annan aspekt i att se till att samtyckeskravet följs.

*Informationskravet* syftar till att informanterna är väl informerade inför deras deltagande (ibid). Vilket garanterades genom både skriftlig och muntlig information som inkluderade syftet för forskningen, dess ämne, att deltagande är frivilligt och anonymt, uppsatsen kommer att publiceras offentligt samt hur materialet kommer att användas.

Det fjärde och sista kravet, *konfidentialitetskravet*, berör principen om vem som kommer ta del av intervjumaterialets information och att obehöriga ej kommer ta del av detta (ibid: 52–53). Även detta framgick skriftligt och muntligt inför intervjuerna. För att inte kunna härleda vare sig transkriberingar eller inspelningar döptes informanterna till intervjupersoner 1–5, med randomiserad ordning så att ingen kan räkna ut vem som har varit under vilken tid eller datum. Dessutom förtydligades att materialet kommer att raderas efter studiens godkännande och transkribering kommer endast att lämnas ut till behöriga, och trots detta kommer identifiering inte kunna ske.

### **3. Tidigare forskning**

För att framställa en tydlig grund för forskning att utgå ifrån, har flertalet av olika forskning gått igenom med fokus på AI, etik och sociologi. Dessa tre områden som en kombination har inte varit självklart, därför har tidigare studier om områden som delvis går att överlappa, eller redan överlappar varandra, byggts vidare på. Samtliga efterföljande studier har dessutom citerat varandra i sina studier.

#### **3.1 Algorithmic Intimacy - Anthony Elliot (2022)**

Anthony Elliot, forskningsprofessor i sociologi vid University of South Australia, undersöker påverkan av prediktiva algoritmer på samtida interpersonella interaktioner. Elliot hävdar att dessa interaktioner äventyras av både den digitala revolutionen som inte kan förutses av automatiserad maskinintelligens.

#### **3.2 Machine Habitus - Towards a Sociology of Algorithms - Massimo Airoidi (2021)**

Massimo Airoidi, sociolog och biträdande professor vid EMLyon Business School, bygger den här studie utifrån Pierre Bourdieus teorier. Airoidi hävdar att algoritmer är sociala varelser och agenter som aktivt engagerar sig i dagliga aktiviteter. Airoidi demonstrerar samhällets inflytande på algoritmisk kod, och hur de kulturella aspekterna inbäddade i koden påverkar dess praktiska funktion i samhället och formar själva samhället. Genom att skapa en ny version av Bourdieus "Habitus", nämligen Airoidis egna "Machine Habitus", visas hur algoritmerna i vår vardag är en del av vår sociala värld.

#### **3.3 The 4th Revolution - How the Infosphere is Reshaping Human Reality - Luciano Floridi (2014)**

I den här boken visar Luciano Floridi, professor i sociologi av kultur och kommunikation vid universitetet i Bologna, att skillnaden mellan online och offline existens har upphört att existera, och vi har blivit sammanflätade med varandra. Floridi använder sig av flera begrepp, inklusive "grounding problem", för att undersöka processen genom vilken ord (symboler) får sin betydelse och karaktären av denna betydelse i sig. Enligt Floridi erbjuder AI inte ett beskrivande eller



föreskrivande förhållningssätt till den yttre världen; snarare blir det en integrerad del av den yttre världen.

### **3.4 Sociological Insight: An Introduction to Non-Obvious Sociology - Randall Collins (1992)**

Randall Collins har utövat en betydande inverkan på sociologisk forskning, med hans verk som är välkända inom sociologi och andra tvärvetenskapliga domäner. Dessutom har han haft lärartjänster vid många prestigefyllda universitet globalt. Collins arbete syftar specifikt till att belysa invecklade sociologiska processer och begrepp, vilket gör dem mer begripliga. Collins menar även att AI har en del i den icke givna sociologiska sfären. Collins anser att datorarkitektur och programmering måste likna mänsklig kognition bortom tekniska överväganden. Sociala och psykologiska frågor beaktas i hans helhetssyn. Perspektivet utmanar ramverk för AI-utveckling som prioriterar beräkningseffektivitet framför socialpsykologisk harmoni. Den här utgåvans nya kapitel om sociologiska idéer och AI-utveckling är viktigt. Collins uppmanar utvecklare att införliva mänskliga sociala egenskaper, vilket är nödvändigt för att AI-system ska efterlikna mänskliga resonemang. Bokens fokus på mänskliga sociala och psykologiska egenskaper passar denna studie för att studera mer sympatiska och socialt medvetna AI-system. Collins tvärvetenskapliga synsätt främjar ett människocentrerat AI-paradigm.

### **3.5 På AI-teknikens axlar - Om kunskapssociologin och stark artificiell intelligens - Peter Kåres (2009)**

I sin doktorsavhandling publicerad av Sociologiska institutionen och Lunds universitet inleder Peter Kåres med en pessimistisk syn på AI. Men efter att ha studerat social teori, inklusive ontologiska grunder, blir författaren mer optimistisk om sann artificiell intelligens. Denna rörelse kritiserar metodologisk individualism, som bygger sociala beteenden på individuella aktiviteter, och försöker bättre förstå sociala fenomen. Sociologin bör omfamna AI genom att fokusera på sociala fenomen utanför mänskliga aktiviteter, enligt författaren.

### **3.6 Managing Bias in AI (2019)**

"Managing Bias in AI" av Drew Roselli, Jeanna Matthews och Nisha Talagala utforskar komplexiteten av bias i AI-system, och fokuserar specifikt på riskerna förknippade med dess implementering i affärsmiljöer. Uppsatsen erkänner utmaningarna med att utrota bias som härrör från historiska data. Författarna föreslår åtgärder för att komma till rätta med fördomar kopplade till översättning av affärs mål, distribution av utbildningsprov och individuella insatsprov. Även om det inte finns någon felfri lösning, betonar författarna att implementering av föreslagna metoder avsevärt kan mildra de negativa effekterna av bias i AI-system.

## **4. Teoretisk ram**

Teorierna som valts ut till att vara grund för studien är Symbolisk interaktionism samt Habitus. Teorierna kommer att inkludera samtliga delar av det mikroskopiskt-makroskopiska kontinuumet.

### **4.1 Herbert Blumer - Symbolisk Interaktionism**

Symbolisk interaktionism, formulerad av Herbert Blumer som en förlängning av George Herbert Meads symboliska interaktionism, är en teoretisk ram som förklarar hur mänskligt beteende påverkas av människors tolkning av sociala objekt och händelser, och hur individer tilldelar dessa elementbetydelse. Blumer hävdar att symbolisk interaktionism omfattar alla aspekter av samhället (Blumer, 1969).

Teorin hävdar att den härrör från sociala interaktioner och den efterföljande tolkningen av dessa interaktioner. Betydelsen härleds från utbyte av symboler, när de inblandade parterna delar en delad uppfattning om symbolens betydelse. Detta accepteras utan någon tolkning eller tvivel. Symbolen får mening genom upprättandet av ett sammanhang och ömsesidig förståelse. Därefter deltar individen i en intern process efter interaktionen, som involverar introspektion och kontemplation. Under denna interna dialog utvärderar och assimilerar individen händelserna och innehållet i interaktionen, med hjälp av personliga erfarenheter och perspektiv för att konstruera egen tolkning av situationen. En individs handlingar är en direkt konsekvens av deras inre

kognition. Den pågående interaktionen skapar en känsla av sammanhang eftersom delade uppfattningar och beteenden visar att engagemanget fortsätter. Följaktligen ger dessa interaktioner upphov till meningsfulla symboler som tjänar funktionen att skapa betydelse. Teorier, ordval och slutsatser fungerar som avgörande symboler som har identisk betydelse för de personer som är inblandade i mötet. En känsla av enhet och enighet uppstår (Blumer, 1969).

Blumers sociala interaktionism erbjuder ett omfattande ramverk för att förstå den intrikata process genom vilken individer konstruerar och tillskriver vikt åt både materiella och immateriella element i sin omgivning (Blumer, 1969).

#### **4.1.1 Samhällsanalys & Premisser**

Blumer beskriver den symboliska interaktionismen som att den vilar på tre premisser och att det alltid finns en mening med en social interaktion. Den första premissen beskriver att människor handlar mot saker baserat på meningen som saken har för dem. Dessa saker kan vara allt ifrån immateriella till materiella ting. Det kan vara individer, institutioner och fenomen. Den andra premissen är att meningen med dessa ting kommer eller uppstår ifrån människans sociala interaktioner som denne har med sina medmänniskor. Den tredje och sista premissen är att dessa betydelser hanteras och modifieras i en tolkningsprocess hos den individ som hanterar och bemöter dessa ting (Blumer, 1986).

#### **4.1.2 Relevans**

Den symboliska interaktionismen kan därtill appliceras och tolkas olika beroende på en individs tolkning av den sociala verkligheten. Tillämpningen av symbolisk interaktionism är särskilt betydelsefull eftersom den fokuserar på den process genom vilken människor skapar betydelse genom interaktioner med både andra människor och deras miljö. Denna synvinkel är särskilt viktig när man överväger effekterna av AI på samhället. Som en teknisk enhet påverkar och påverkas artificiell intelligens av mänskliga interaktioner och uppfattningar. Att undersöka hur studenter vid Lunds universitet känner igen och lösa etiska problem och fördomar under utvecklingsprocessen kan ge insikt i den sociala konstruktionen och förståelsen av AI. Detta

tillvägagångssätt stämmer överens med min granskning av AI som en social agent, och betonar vikten av individuella och grupps attityder och handlingar i skapandet och den samhälleliga effekten av AI-teknik (Blumer, 1969).

## **4.2 Pierre Bourdieu - Habitus**

Pierre Bourdieu myntade begreppet Habitus. Bourdieu (1977) förklarar Habitus som en kollektiv enhet genom vilken och i vilken dominerande sociala och kulturella förhållanden etableras och reproduceras. Habitus hänvisar till *“ett subjektivt men inte individuellt system av internaliserade strukturer, scheman för perception, uppfattning och handling som är gemensamma för alla medlemmar av samma grupp eller klass”* (Bourdieu, 1977: 86).

Dessa internaliserade strukturer och uppfattningsscheman strukturerar subjektets (delade) världsbild och deras uppfattning av den värld i vilken de antar att de existerar (ibid).

### **4.2.1 Relevans**

Begreppet habitus är relevant för studien då det erbjuder ett ramverk för att förstå hur individen internaliserar och uttrycker social och kulturell omgivning. Begreppet habitus kan användas i samband med AI-utveckling för att analysera hur den sociala och pedagogiska bakgrunden hos AI-utvecklare, särskilt ingenjörstudenter vid Lunds universitet, påverkar deras perspektiv och metoder för att hantera etiska dilemman och fördomar inom AI. Detta koncept tillåter forskare att undersöka hur dessa internaliserade ramverk påverkar utvecklingen av AI-teknologier av utvecklare och deras efterföljande effekter på samhället. Detta leder till en djupare förståelse för den sociala dynamiken inom området AI.

## 5. Analys och resultat

I följande avsnitt kommer svaren från informanterna att presenteras, analyseras och diskuteras med bakgrund av forskningsfrågorna och syftet med uppsatsen. Empiri från informanterna kommer att åtskiljas i avsnitt, för att förstå de olika teman som är relevanta för forskningen. Svaren kommer att ställas mot varandra, tidigare presenterade teorier, tidigare forskning och mot sig själva för att ge en heltäckande analys.

### 5.1 Bias och sociala strukturer

Det efterföljande avsnittet undersöker resultaten som produceras av AI och deras potential för opartiskhet. AI är en återspeglning av data, vilket understryks av samtliga respondenter. Även om data som används för att träna AI är generisk och etiskt inhämtad, finns det ingen garanti för att resultaten skulle vara opartiska.

Airoidi (2021: 74) hävdar att maskininlärningssystem har anmärkningsvärda likheter med Bourdieus idéer om mänsklig handlande i deras sociala funktion (Airoidi, 2021: 74) (Bourdieu, 1977; 1990a). Bourdieu hävdade att en individs beteenden påverkas av externa sociala praktiker, Habitus. Individens beteende modifieras av nya erfarenheter och situationer inom ett visst sammanhang. Följaktligen är det möjligt för en maskin att ha en habitus, vilket kallas för Machine Habitus (Airoidi, 2021: 74).

*“Jag tror det största problemet är just att träna dem på information som redan finns på internet eller ofta, eller liksom information som redan finns. Så är även det jättesvårt att skapa en AI som inte är biased.” (Intervjuperson 1)*

Intervjuperson 1 betonar betydelsen av kunskap och data som används för AI-träning, ett perspektiv som också stöds av Collins (1992). Dessutom hävdar intervjuperson 1 att AI oundvikligen skulle anta specifika attityder och fördomar på grund av dess inmatning av redan existerande data, som speglar människors onlineaktiviteter och stämmer överens med deras verkliga handlingar. Collins (1992) betonar detta koncept ytterligare genom att citera ordspråket

"Garbage In, Garbage Out" (Collins, 1992: 165). Datorer uppvisar inflexibilitet, likt individer med liten intelligens, som ständigt upprepar samma fel.

*“Jag tänkte på en annan fara, som då handlar om att den är biased eller så, men inte just på grund av informationen den matas med utan mer av vem som bestämmer vad som är okej och inte okej när det kommer till svaren från AI och vad den kan providea oss med för information.” (Intervjuperson 2)*

Intervjuperson 2 understryker vidare betydelsen av att garantera att utvecklaren av AI-systemet förblir opartisk. Ändå hävdar Roselli et al (2019) att pliktbördan inte enbart bör läggas på skaparen av modellen, utan snarare bör bedömas av icke-tekniskt ledarskap och en granskande allmänhet eller tillsynsenhet. Intervjuperson 2 betonar att opartiskheten hos en AI-modell påverkas inte bara av data som används för träning, utan också av utvecklarens bakgrund, motiv och intention. Informanten ifrågasätter ansvarsskyldigheten för utvecklaren och kontrollanten av systemet.

*“För mig blir det väldigt negativt, när en dator eller algoritm ska bestämma hur en människa ska må eller inte tänka. /.../ ” (Intervjuperson 3)*

*“/.../ Så går man i stället till en “för dig” sida på Tiktok där man bara scollar med innehåll som Tiktok bara matar en.” (Intervjuperson 3)*

Intervjuperson 3 har ett liknande perspektiv som informanterna 1 och 2. Å andra sidan omdirigerar respondenten 3 fokus från själva data till de genererade resultaten, såväl som deras praktiska användningsområden och mål. Som nämnt av informant 3 kan AI:s inverkan på samhället jämföras med en kedjereaktion, särskilt i relation till algoritmen som är inkorporerad i sociala medieplattformar. Denna algoritm får näring och utbildas med hjälp av data som erhållits från individers internetbeteenden.

Elliott (2023) validerar informantens perspektiv. Tiktok är inte den första sociala medieplattformen som tillhandahåller rekommenderat innehåll genom algoritmer och spårning av individuell användning. Tidigare medier har dock förlitat sig på användarens konsumtion av innehåll och plattform. Därför går det att argumentera för att individens habitus (Bourdieu) eller det maskinella habitus (Airoldi, 2021) även baseras på individens onlinebeteende.

Denna algoritmiska övervakningsmetod omfattar gradvis och delvis interpersonella interaktioner och känslomässiga band, genom att automatiskt övervaka olika aspekter av användarnas beteenden och preferenser. Den snabba utvecklingen av algoritmisk övervakning får användare att aktivt söka efter högspecialiserade och precisa mikrokulturer som tillgodoser deras individuella sociala intressen. Ändå har detta fenomen sin egen uppsättning utmaningar, eftersom förmågan att utforska särskilda domäner kan stödja konstnärliga böjelser och underlätta bredare funktionalitet. Men det kan också oavsiktligt styra användare mot skadliga ideologiska perspektiv och online-gemenskaper (Elliot, 2023).

Floridi (2016) belyser en jämförbar synpunkt, specifikt relaterad till datavetenskap, kallad "grounding problem". Detta dilemma hänför sig till semantikens område. Enligt Floridi (2016) kommer datakonsumtion alltid att kräva ett semantiskt sammanhang, vilket påverkar betydelsen vi härleder från data och utdata från AI. Denna begränsning är inte exklusiv för människor utan sträcker sig även till maskiner. Detta beror på deras samexistens och engagemang i en osynlig dans.

## 5.2 Interaktioner

*“In short, we need to make this computer just like an ordinary person making their way in the social world!” (Collins, 1992:167)*

Citatet är utdraget från Collins (1992), som syftar till att presentera de mindre uppenbara aspekterna av sociologi, inklusive dess koppling till AI. Icke desto mindre hävdar Collins att artificiell intelligens är oförmögen att replikera mänskligt intellekt. Peter Kåres (2008) motbevisar detta argument genom att hävda att vetenskaplig kunskapsförmedling undantagslöst kräver en extern symbol för att underlätta interaktion, vilket motbevisar föreställningen att alla sociala interaktioner endast sker mellan individer. Collins artikel publicerades flera år innan Chat GPT blev tillgänglig för allmänheten, och den utmärker sig i att imitera intelligens. Men hans skepsis är fortfarande relevant.

Principerna från den sociala världen bör även tillämpas på den digitala världen. AI har förändrat människors interaktioner, både medvetet och omedvetet. Maskiner och artificiell intelligens kan delvis ha ersatt människors mening. Aioldi (2021: 70) menar att maskininlärningssystemen blir sociala agenter eftersom de lär sig från socialt strukturerade dataset och kontexter. Som ett resultat blir de en del av vår sociala värld. Således kan det hävdas att AI-system inte bara fungerar som en ersättare för sociala interaktioner mellan människor; de blir också en ny spelare i spelet.

*“Jag är lite rädd att det kan bli så att människor blir mindre interaktiva med varandra och blir mer beroende av det maskinella.” (Intervjuperson 3)*

Genom att bryta ner den roll som datorer, inklusive AI, har för huruvida individer skapar mening i sociala interaktioner kan symbolisk interaktionism inkorporeras (Blumer, 1969). I kontrast till Kåres (2008) påstående om att symbolisk interaktionism och AI sätts i dikotomi. Vad som beskrivs av informanten ovan påtalar på vad som händer när en maskin ersätter en individ. Kåres (2008) menar att det faktiskt finns en människa bakom maskinen, nämligen utvecklaren, vilket



samtliga informanter lyfter fram. Därmed kan den mänskliga faktorn, enligt Peter Kåres, i symbolisk interaktionism fortfarande vara relevant.

*“Människor dras isär från varandra i det interaktiva i verkligheten mellan människor och mer blir att man integrerar i stället med maskiner.” (Intervjuperson 1)*

Elliott (2023) konstaterar också att nya studier tyder på att datormedierade relationer och interaktioner kan eller kommer att förlora mening i sitt värde eftersom parasociala relationer och interaktioner förändras och reformeras. Frågan är om användningen av dessa medier, med inbäddade algoritmer eller rena generativa AI:er, syftar till att ersätta eller kompensera för otillräckliga sociala relationer. Därför kan fokus komma att flyttas till individens interaktion med maskinalgoritmerna, baserat på den mänskliga tendensen att skapa illusoriska dimensioner av mänsklig interaktion.

*”Om man kollar på typ Tiktok så kan det ju påverka en människa väldigt extremt för att där är det ju bara att man blir matad med massa information, och det är inte du som bestämmer vad det är för information.” (Intervjuperson 5)*

*“Om man blir matad med massa information om någonting väldigt länge så kan det vara så att det är det enda man tänker på, så kommer man vilja snacka om sådana saker och så vidare. /.../  
Det påverkar i sin tur hur människor interagerar med varandra.” (Intervjuperson 5)*

I takt med ökad digitalisering och växande AI-närvaro blir vi alltmer frånkopplade från varandra och våra egna övertygelser. Vår enda referenspunkt är att maskiner har dikterat våra tankar och känslor, och det är här vi konvergerar. Vi saknar individuella perspektiv och de slutsatser vi delar med oss av är inte längre enbart våra egna, utan snarare instruktioner från maskiner till dem som håller med oss. Vi saknar ett personligt ställningstagande, och den konsensus vi kommer fram till härrör i stället från de maskiner som förser oss med information, i motsats till våra egna kognitiva processer.

*“/.../ Det formar ju människor socialt. Ser de mycket av en viss grej, som en maskin bestämmer, så kommer människorna också vilja prata om det med sina vänner. Därför blir det en påverkan på det sociala hos människor, utifrån maskinell styrning.” (Intervjuperson 3)*

Enligt intervjuperson 3 har det ökade utnyttjandet av maskininlärningsalgoritmer i våra dagliga aktiviteter en betydande inverkan på våra sociala kontakter. Att minska människors beroende av dessa algoritmer för att förstå deras önskningar, mål och ambitioner kommer att leda till en avskildhet från mänskliga kopplingar.

I sitt arbete hävdar Elliot (2023) att individer successivt frigör sig från interpersonella kopplingar och förlitar sig på maskininlärningsalgoritmer för att analysera och förstå sina krav, ambitioner och begär. Intervjuperson 3 instämmer i detta perspektiv, men det är avgörande att erkänna att sociala kopplingar och spridning av idéer baserade på maskininläring inte är helt skadligt.

*“De flesta tror jag har en ganska positiv inställning till AI, man ser ju hur många som faktiskt använder Chat GPT nuförtiden. /.../” (Intervjuperson 4)*

Intervjuperson 4 förklarar att det finns anledning att vara positivt inställd till AI, med tanke på dess utbredda användning. Elliot (2023) förklarar vidare att AI inte är begränsat till generativ AI som Chat GPT, eftersom det finns modeller för specifika behov, till exempel emotionell AI, som sätter ett nytt mönster för AI. Dessa typer av AI är bekräftande för individer som söker emotionellt stöd, eftersom de inte är kritiska eller ifrågasätter en individs upplevelser. Detta nya universum av kamratskap kan stärka individer och hjälpa till med utmaningar eller frågor som vissa individer kanske inte känner sig bekväma att dela med sina fysiska vänner (Elliot, 2023).

*“Problemet med AI har ju blivit att det verkligen blir som en personlig assistent som dessutom kan allting.” (Intervjuperson 2)*

Intervjuperson 2 ser AI som en personlig assistent för individer som en negativ aspekt, eftersom det kan leda till att individer förlitar sig på maskiner för social interaktion i stället för människor. Även om AI kan ge skräddarsydda svar och olika typer av bekräftelse, kvarstår denna fara (Elliot, 2023).

Den symboliska interaktionismens karaktär är att hur mening skapas i sociala interaktioner genom symboler. Dessa symboler är ett slags socialt objekt. Därför kan AI tolkas som ett socialt objekt. Olika typer av AI-algoritmen finns överallt i vår omgivning, och genom olika typer av modeller som nyttjar vår data - medvetet och omedvetet, matas vi med information som datorn tycker att vi bör se. Detta sätter i gång en kedja av hur vi tolkar informationen vi får tilldelad, hur vi sprider den med våra jämlike, som i sin tur skapar trender. Dessa trender kan till synes vara harmlösa, hur man ska klä sig, fixa håret och liknande. Men exakt samma system används genom att påverka individers värderingar, politiska åsikter och liknande. Konsekvenserna och stegen från trend till samhällsnorm blir mindre, sedan fortsätter interaktionerna i alla sammanhang. Men vad som är viktigt att poängtera att vad som tidigare varit vad som satte i gång den egna reflektionen av en interaktion, för att sedan skapa symboler, då förutsatte att interaktionen var mellan människor (Blumer, 1969).

### **5.3 Utbildningens påverkan**

Detta avsnitt behandlas informanternas **generella inställning till AI** och hur den formats med utgångspunkt i deras utbildningsval, samt eventuell förändring sedan utbildningens start. Samtliga informanter har använt AI flera gånger i veckan, framför allt generativ AI, som Chat GPT. Syftet med användningen har varit som ett hjälpmedel både i privatlivet och studierna, det har varit hjälp med kodning, skriva mejl eller få en sammanfattning eller förståelse för olika uppgifter.

*“Ja, alltså det har förändrats under min studietid. Men det har inte förändrats hur jag använder AI på något vis.” (Intervjuperson 1)*

Informanterna hävdar att deras erfarenheter av mjukvaruutveckling under utbildningen också har påverkat deras allmänna inställning till AI, etiska aspekter och konsekvenser av felaktig hantering. Det framkommer hos envar studenter att en kombination av ökad egen AI-användning samt utbildningens naturligt progressiva utfall resulterar i en förändring både i generella attityder samt ger en ökad försiktighet.

*“/.../ Jag har alltid varit lite kritisk. /.../ Scoopet har ändrats, innan var det mer svart och vitt, antingen var det bra eller dåligt, men nu med djupare förståelse för programmering och så vidare är det lite klarare vad som är problemet. Då blir det också lättare att påverka, förstå och dra slutsatser kring. Det har absolut förändrats.” (Intervjuperson 3)*

De tredjeårsinformanter som inte har påbörjat sin specialisering kan ändå se en förändring i hur de uppfattar AI jämfört med de två informanter som är i sitt femte år och också har byggt sina egna enkla AI-modeller, medan de återstående tre informanterna inte har byggt några AI-modeller men har arbetat med mjukvaruutveckling. Således varierar förståelsen hos alla informanter, delvis på grund av deras studieår men också deras intresse för teknik och nyfikenhet utanför sina studier.

*“Jag har alltid tyckt att AI är ganska coolt, människor kan bygga sådana modeller. Men ju mer man lär sig om det, desto mer inser man både begränsningarna och vad det kan leda till.”*  
*(Intervjuperson 4)*

*”/.../ Definitivt att min uppfattning ändrats. /.../ Att av alla framgångar och upptäckter och allt nytt som kommit och det media har skrivit blir man rädd för i samhället, men allting jag har pluggat, där man faktiskt gått in på hur allt det här funkar så blir jag faktiskt inte lika rädd, för tekniken är egentligen, det är ju bara teknik och det är ju bara matematiska ekvationer som sitter där och räknar, det är ju bara det. /.../” (Intervjuperson 5)*

Intervjuperson 5 skiftar fokus mot externa påtryckningar från media om farorna med AI. Informanten menar att en djupare förståelse för tekniken också ger en känsla av säkerhet och förståelse för att tekniken inte är så farlig som allmänheten kanske tror eller fruktar.

*“De är kanske lite mer rädda för om det ens är nödvändigt med AI. /.../ Sen finns det ju också de som förstår sig på AI och också är rädda för AI, för de förstår riskerna som finns.”*

*(Intervjuperson 4)*

Det som kan tolkas från intervjuperson 4 är att det skapar en paradox av medvetenhet, där det finns en rädsla för det okända för individen, men de som faktiskt har kunskap om AI är också rädda eftersom de förstår begränsningarna. Denna medvetenhet kan utläsas och förstås hos alla intervjupersoner.

## **5.4 Etiska aspekter**

### **5.4.1 Förutsättningar och insikter från studierna**

Samtliga informanter uppgav att de inte upplever att deras utbildning har gett dem tillräckligt grund för etik inom mjukvaruutveckling, och framför allt AI-utveckling. Det går delvis att förklara genom att tre av informanterna inte påbörjat sin specialisering och läser grundblocket i sin utbildning. Däremot så har de två informanter som är i slutet av sin utbildning, och specialiserar sig inom AI, samma upplevelse, att den är knapphändig.

*“Nu går jag en etikkurs i AI, bara för att den verkade väldigt lätt att klara.” (Intervjuperson 5)*

Däremot har samtliga informanter uppgett att de har fått en ökad etisk medvetenhet sedan de påbörjade sin utbildning. Detta kan förklaras av väldigt många olika anledningar, bland annat att deras fördjupade förståelse av mjukvara från utbildningen påverkat deras ståndpunkt, men även utomstående faktorer som inte är begränsade till utbildningen.

Den generaliserade andre (Blumer, 1969) är en del av den symboliska interaktionismen som förklarar att en individ uppfattar ett problem, och i sin tolkning av detta tar personen gruppens gemensamma uppfattning i beaktning när ens egen bild formas. En av den symboliska interaktionismens grundpelare är att allt är föränderligt, vilket visas tydligt av informanterna genom att samtliga uppger att deras erfarenheter, både utanför och innanför studierna, har förändrats.

*”Jag tror inte det kommer vara och ge mig tillräcklig djupgående bild av etiken liksom bakom mjukvaruutveckling.” (Intervjuperson 2)*

*“Vi har haft två kurser som delvis varit om etik kring mjukvaruutveckling. Nu är det ju så aktuellt så AI är ju en del av det. Nu läser vi en kurs om affärsdriven mjukvaruutveckling, där vi har en hel uppsats om etik. /.../” (Intervjuperson 3)*

Trots att tredjeårsinformanterna upplyser om att de har gått, och går kurser i etik vid mjukvaruutveckling, anser ingen av informanterna att de har lärt sig är tillräckligt.

*”Vi har knappt gått igenom någon typ av etik, ehm kanske att någon lärare har sagt någon enstaka mening på någon föreläsning.” (Intervjuperson 5)*

*“Jag tycker att man borde lära sig mer. Om man ska vara en utvecklare av sådana här modeller tycker jag det är viktigt att man känner till konsekvenserna och riskerna som kan uppstå.”  
(Intervjuperson 4)*

*“Med tanke på att man kommer ju vara den som skapar de här modellerna i framtiden, och har väldigt mycket makt i sina händer.” (Intervjuperson 5)*

Förutom den generella förutsättningen att dessa studenter behöver ha en god grund för att förstå riskerna och konsekvenserna av denna typ av AI-utveckling, har informanterna ändå en

förståelse för riskerna och deras ansvar. Faran med AI är i flera steg, delvis hur data tas och sedan vad den kan användas till.

#### 5.4.2 Yttre påverkan, förutsättningar och medvetenhet

*“Allt handlar om att du som utvecklat den, hur du har utvecklat den, och vilken data du har använt.” (Intervjuperson 4)*

Svaren återspeglar det externa inflytandet och de förväntningar på de **etiska attityder** som studenterna har fått och det kan hävdas att det är oberoende av deras studier. Men eftersom dessa individer kommer att utveckla programvaran kan det hävdas att det finns en omedveten intern och extern förväntan på att inhämta kunskap på det berörda ämnesområdet. Valet av utbildning återspeglar informanternas intressen.

Externa influenser, liksom den symboliska interaktionismens karaktär av hur människor skapar mening genom symboler och hur de agerar utifrån detta, formar dessa attityder (Blumer, 1969). Dessa attityder formas genom interaktion med olika samhällsstrukturer, till exempel inom informanternas utbildning, samt externa förväntningar.

*“Hmm, asså jag tror att jag är etiskt medveten på det viset att jag skulle nog aldrig någonsin ha dåliga intentioner vid skapandet av mjukvara.” (Intervjuperson 2)*

*”/.../ Jag skulle aldrig utveckla någonting som använder data på fel sätt för att utnyttja människors privacy. /.../” (Intervjuperson 3)*

*“Mycket av min medvetenhet har kommit utanför studierna, men även en del inom studierna.” (Intervjuperson 4)*

Till skillnad från övriga informanter är intervjuperson 5 den som ser minst risker med AI-utveckling, gällande de etiska aspekterna.

*“För att jag är den som skapar teknologin, då måste jag såklart vara medveten och varsam om just dessa frågor men jag tycker inte att det är jag som ska bestämma sådant här.”*

*(Intervjuperson 5)*

Dessa slutsatser som intervjuperson 2 beskriver är i enlighet med HELGs slutsatser gällande att det är flera nivåer som AI behöver regleras på. Bestämmelser ska vara samförståndiga och ansvar ska inte läggas på den enskilde utvecklaren. Insyn i data är viktig och är det enda sättet att garantera är korrekt och inte använder skyddade eller inkorrekta data, den individ vars data är skyddad bör kunna se uppgifterna om den handlar om denne (Floridi et al., 2019).

*“Den viktigaste rollen är att vi som kommer att arbeta med det här kommer att skapa ett samtal av trust och transparens med folket som faktiskt bestämmer de här reglerna och bestämmer vad som är etiskt och vad som inte är etiskt och moraliskt. /.../” (Intervjuperson 2)*

Holmström et al. (2020) delar denna uppfattning och menar att det är viktigt för utvecklarna att koppla ifrån sig yttre förväntningar och påverkan för att i stället endast fokusera sig på det teknologiska.

*“/.../ Jag tycker att fokuset bör skiftas mer åt vilken typ av data som man använder eller skapar sina program kring. Fokuset borde skifta lite mer åt bara data hållet och inte endast mjukvaruprogram och dess användningsområden. /.../” (Intervjuperson 3)*

*“Man måste ju ha koll på vad det kan användas till, och så vidare, innan man faktiskt gör sakerna. Men det är en jättesvår fråga för då kanske det är att man vägrar göra det, men då är det en annan person som trycker på knappen i alla fall så görs det ändå. Det är främst företaget och statens ansvar. “(Intervjuperson 5)*



## 5.5 Framtid

Följande avsnitt kommer att fokusera på informanternas attityder, förväntningar och upplevelser inför framtiden. Även här går det att utläsa generella teman, med negativa och positiva inställningar.

*“Det blir mindre slutsatser som människor drar själva, vilket leder till att människor generellt kommer bli lite mer lika varandra, jag tror inte det gynnar oss i längden. Just det här att det blir svårt att dra egna slutsatser när någon annan, en dator, drar slutsatserna åt en.”*

*(Intervjuperson 3)*

Insikterna från Intervjuperson 3 kan också kopplas till hur människor kommer att interagera med varandra, och hur det kommer att förändras drastiskt i framtiden med den exponentiella utvecklingen av AI. Detta förklaras också i avsnittet ovan gällande interaktioner, eftersom dess nuvarande inverkan också är relevant för framtiden och det alltmer omfattande samhällliga inflytande som AI kommer att ha på våra dagliga sociala liv och den allmänna samhällsordningen.

*“Positivt! Det kommer att effektivisera mycket, men det är farligt att ta bort människan i det, men det tar också bort de mänskliga misstagen. /.../ Men vi måste vara tydliga med att vi inte får bli för beroende av AI heller, med tanke på att det kan finnas fel. /.../”* (Intervjuperson 4)

Det finns inte endast negativa aspekter av AI och dess inverkan i framtiden. Samtliga intervjupersoner lyfter fram att maskinintelligens kommer att underlätta och effektivisera ändamål där människor idag tvingas utöva monotona arbetsuppgifter. Algoritmerna kan möjliggöra nya idéer och perspektiv i stora samhällsutmaningar. Samtliga intervjupersoner är dock tydliga och försiktiga med att direkt adressera de negativa aspekterna efter att de positiva aspekterna har lyfts fram.

## 6. Slutsats och vidare forskning

Syftet med denna studie var att undersöka det intrikata förhållandet mellan artificiell intelligens (AI) och samhället, med särskilt fokus på studenter inom datateknik. Det blir tydligt att AI inte bara är ett enkelt tekniskt verktyg, utan snarare ett viktigt samhällsfenomen. Den efterföljande slutsatsen syntetiserar de viktigaste upptäckterna och funderar på hur de inte bara är i linje med, utan också ifrågasätter vår förståelse av AI:s inverkan på samhället.

Genom att i analysen dela upp de olika aspekterna för studiens syfte, går det att utläsa en röd tråd och hur alla dessa faktorer samspelar. Det är ett komplext samspel, och en osynlig dans mellan sociologisk teori, etik och studentens roll i AI-utveckling. Genom att sammanfatta detta som en kedja, från Bourdieu och Blumer, kan ett nytt koncept födas. Denna modell innebär alltså:

*Etisk datainsamling → Bias och Habitus → AI:s roll i samhället → Symbolisk interaktionism*

### 6.1 Etik

Vilket majoriteten av informanterna har belyst, krävs det en varsamhet och förståelse för etisk datainsamling, vilket syftar till en förståelse för det juridiska och moraliska sammanhanget vid datainsamling, hedra integritet och tillförsäkra delgivet samtycke. Artificiell intelligens speglar de data som den lärs ut på. Även om den data som används kommer från källor som är generiska och etiskt sunda, finns det ingen garanti för att få opartiska resultat.

Informanter betonade svårigheten att utveckla opartisk AI, särskilt när den tränas på redan existerande internetdata som speglar mänskliga onlinehandlingar och verkliga beteenden.

Informanternas utbildningserfarenheter har påverkat deras övergripande perspektiv på AI och dess etiska dimensioner. Studenter observerade en förändring i deras allmänna attityder om AI och utvecklade en större känsla av försiktighet, vilket påverkades av deras förstahandserfarenheter av mjukvaruutveckling under hela utbildningen.

## **6.2 Bias och Habitus**

Det är ofrånkomligt med partiskhet vars ursprung speglar utvecklaren och dess habitus - deras sociokulturella bakgrund. Dessa internaliserade fördomar och värderingar kan ouppsåtligt spridas till AI-system. Detta reflekterar samhälleliga bias. Betydelsen att säkerställa AI-opartiskhet bör inte bara vila på utvecklaren, utan bör också omfatta icke-tekniska ledarskaps- och övervakningsgrupper. Neutraliteten hos en AI-modell bestäms inte enbart av träningsdata, utan också av utvecklarens bakgrund, mål och ambitioner.

## **6.3 AI:s roll i samhället**

När AI är i bruk, tar den rollen som en aktiv agent, som fungerar som en social enhet inom den sociala sfären. Individens beslut och interaktioner kan utmana, modifiera eller stärka samhälleliga normer och fördomar. Att inse dess betydelse är avgörande på grund av den intrikata naturen hos dess förbindelser med människor. Detta kommer att underlätta förståelsen av dess funktion i samhället, inte bara i mellanmänskliga relationer utan också inom en bredare social ram. Denna insikt understryker det faktum att AI har en dubbel roll i att påverka och forma samhällets normer, seder och värderingar. Dessutom belyser den de etiska implikationerna som är involverade i AI-beslutsförandet och den inneboende predisposition som dessa system kan ha. Genom att undersöka AI från en mer sofistikerad och effektfull synvinkel kan vi få insikt i och effektivt hantera dess invecklade interaktion med mänskliga sociala system. Följaktligen resulterar detta i ett mer ansvarsfullt och rättvist införlivande av AI i vår dagliga tillvaro.

Effekten av AI på samhället kan liknas vid en oförutsedd kedja, särskilt med avseende på algoritmer som används i sociala medieplattformar, som matas och utbildas genom data som härrör från individuella onlinebeteenden. Detta införlivar gradvis interpersonella interaktioner och känslomässiga kopplingar genom att automatiskt övervaka olika aspekter av användarens beteende och preferenser.

#### **6.4 Symbolisk interaktionism**

Den symboliska interaktionismen är mycket viktig för att förstå AI:s roll och inverkan på samhället. AI:n är idag en social agent, som majoriteten av människor stöter på dagligen omedvetet. Detta innebär även att AI:n har en inverkan på sociala strukturer, utan att andra sociala agenter - människor, är medvetna om detta. Det växande utnyttjandet av maskininlärningsalgoritmer i vardagliga sysslor har ett betydande inflytande på sociala interaktioner, vilket potentiellt kan resultera i en frånkoppling från mänskliga relationer.

#### **6.5 Tvärvetenskapliga lösningar**

Som en av informanterna belöser, är det viktigt att förstå att utvecklarnas roll är att skapa system, och när dessa skapas utgår de från existerande regelverk. Dessa regelverk skapas inte av utvecklarna, utan genom tvärvetenskapliga lösningar.

#### **6.6 Vidare forskning och slutreflektioner**

Den snabba tillväxten av AI exemplifierar den dynamiska skärningspunkten mellan teknisk innovation och långvariga samhällsliga koncept. Denna studie understryker vikten av att upprätthålla en konversation mellan klassisk sociologi och nya AI-fenomen. Utvecklingen för hur individer agerar med digitala enheter är i ständig förändring, vilket accentuerar relevansen av implementation av sociologiska teorier i den digitala eran. Med tanke på AI-ekosystemets ständigt föränderliga natur är det klokt att vara försiktig när dessa koncept hastigt anammas.

Vidare forskning inom ämnet borde koncentrera på långsiktiga konsekvenser av AI på samhällsliga ramar, med fokus på hur AI-modeller påverkar klassiska sociologiska teorier. Giltighet av nya begrepp inom tvärvetenskapen bör undersökas. Vidare är forskning som fokuserar på etiska konsekvenser av AI från ett sociologiskt perspektiv av stor vikt. Det bör undersöka hur den exponentiella framkomsten av AI hotar våra nuvarande etiska ramverk. Syftet bör vara att åstadkomma en balans mellan sociologisk förståelse och vetenskaplig progression. Detta verifierar att utvecklingen av AI inriktar sig mot socialt och etiskt medvetna grundsatser.

## Referens- och litteraturlista

AI HELG (2019). *AI-frågor inrättad av Europeiska kommissionen i juni 2018: Etiska riktlinjer för tillförlitlig AI*. [online] *Europeiska Kommissionen*. Tillgänglig på:  
[file:///Users/astrid/Downloads/ethics\\_guidelines\\_for\\_trustworthy\\_ai-sv\\_8801B2F1-9D0A-11C8-9F29081DB6F5EBD9\\_60441.pdf](file:///Users/astrid/Downloads/ethics_guidelines_for_trustworthy_ai-sv_8801B2F1-9D0A-11C8-9F29081DB6F5EBD9_60441.pdf) [Hämtad 4 dec. 2023].

Airoldi, M. (2021). *Machine Habitus*. John Wiley & Sons.

Berger, P. och Luckmann, T. (1999). *Kunskapssociologi: hur individen uppfattar och formar sin sociala verklighet*. 2nd ed. Översatt av S. Olsson. Wahlström & Widstrand.

Blumer, H. (1969). *Symbolic interactionism: perspective and methods*. Englewood Cliffs, N.J.: Prentice-Hall.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press.

Collins, R. (1992). *Sociological insight: an introduction to non-obvious sociology*. [online] New York: Oxford University Press. Tillgänglig på:  
[https://search.alexanderstreet.com/view/work/bibliographic\\_entity%7Cbibliographic\\_details%7C4704889#page/163/mode/1/chapter/bibliographic\\_entity%7Cdocument%7C4704900](https://search.alexanderstreet.com/view/work/bibliographic_entity%7Cbibliographic_details%7C4704889#page/163/mode/1/chapter/bibliographic_entity%7Cdocument%7C4704900) [Hämtad 27 nov. 2023].

Ejvegård, R. (2009). *Vetenskaplig metod*. Lund: Studentlitteratur.

Elliott, A. (2022). *Algorithmic Intimacy*. John Wiley & Sons.

European Commission (2021). *The Act*. [online] The Artificial Intelligence Act. Tillgänglig:  
<https://artificialintelligenceact.eu/the-act/>.

Fishbein, M. and Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, Mass.: Addison-Wesley Pub. Co.

Floridi, L. (2014). *The fourth revolution: how the infosphere is reshaping human reality*. Oxford: Oxford University Press.

Floridi, L. (2023). *The Ethics of Artificial Intelligence*. Oxford University Press.

Jansson, P. (2014). 'Att ligga steget före'. Fokusgruppsintervjuer med specialpedagoger om deras uppfattningar om förebyggande och hälsofrämjande arbete. *gupea.ub.gu.se*. [online] Tillgänglig på: <https://gupea.ub.gu.se/handle/2077/35680> [Hämtad 12 dec. 2023].

Lindgren, S. and Holmström, J. (2020). Social Science Perspective on Artificial Intelligence. *Journal of Digital Social Research*, 2(3). doi:<https://doi.org/10.33621/jdsr.v2i3.65>.

Lindstedt, I. (2019). *Forskningens hantverk*. 2nd ed. Studentlitteratur AB.

LTH (2023). *Civilingenjör i datateknik*. [online] LTH, Lunds Tekniska Högskola. Tillgänglig på <https://www.lth.se/utbildning/datateknik300/#:~:text=Som%20f%C3%A4rdig%20civilingenj%C3%B6r%20i%20datateknik> [Hämtad 12 dec. 2023].

McCarthy, J. (2007). What is Artificial Intelligence? [online] pp.1–15. Tillgänglig på: <https://www-formal.stanford.edu/jmc/whatisai.pdf> [Hämtad 6 dec. 2023].

Natalie (2023). *ChatGPT — Release Notes | OpenAI Help Center*. [online] [help.openai.com](https://help.openai.com). Tillgänglig på: [https://help.openai.com/en/articles/6825453-chatgpt-release-notes#h\\_4799933861](https://help.openai.com/en/articles/6825453-chatgpt-release-notes#h_4799933861).

Skärvad, PH., Lundahl, U. (2016). *Utredningsmetodik*. Lund: Studentlitteratur.

Roselli, D., Matthews, J. och Talagala, N. (2019). Managing Bias in AI. *Companion Proceedings of the 2019 World Wide Web Conference*, pp.539–544. doi:<https://doi.org/10.1145/3308560.3317590>.

[www.safe.ai](https://www.safe.ai). (n.d.). *Statement on AI Risk | CAIS*. [online] Tillgänglig på: <https://www.safe.ai/statement-on-ai-risk>

## **Bilagor**

### **Bilaga 1: Informationsbrev**

Hej! Jag, Astrid Boström Andersson, är student vid Lunds universitet, ägnar mig nu åt att skriva min kandidatuppsats i sociologi. Syftet med min forskning är att undersöka civilingenjörstudenterperspektiv vid LTH på etiska överväganden i utvecklingen av artificiell intelligens (AI) och dess påverkan på samhället. Studien syftar till att undersöka sambandet mellan teknik, etik och samhälle.

Eftersom du nu är studerar datateknik på LTH skulle jag vilja höra av dig om ditt intresse att delta i studien. Intervjun pågår vanligtvis i ungefär en timme och kan genomföras i vilken miljö som helst, oavsett om det är digitalt eller fysiskt. Genom ditt deltagande i intervjun har du gett ditt uttryckliga samtycke.

Ditt engagemang är valfritt och kan avslutas när som helst. Dina uppgifter kommer att förbli ospårbara till dig. Efter ditt medgivande kommer intervjun att spelas in och sedan transkriberas för att fungera som grunden för uppsatsen. Inspelningen och transkriptionen kommer att lagras säkert och kan endast nås av behöriga personer. Dessutom får du chansen att granska din utskrift för att verifiera att det inte har förekommit några fall av missförstånd eller feltolkningar av dina perspektiv.

Transkriberingen och inspelningen kommer att raderas vid godkännande av avhandlingen.

Om du vill delta i forskningen är det av största vikt. Om du har några frågor, vänligen kontakta mig via e-post på [as8238bo-s@student.lu.se](mailto:as8238bo-s@student.lu.se) eller per telefon på 070 842 36 43.

Tack så mycket!

## Bilaga 2: Beteckning av informanter

Vem	Relevans	Datum	Längd
Intervjuperson 1	Student i åk 3	23/11	43 min
Intervjuperson 2	Student i åk 3	23/11	65 min
Intervjuperson 3	Student i åk 3	23/11	73 min
Intervjuperson 4	Student i åk 5	23/11	43 min
Intervjuperson 5	Student i åk 5	27/11	54 min

## Bilaga 3: Transkribering

Transkribering lämnas på begäran.



#### **Bilaga 4: Intervjuguide**

Har du använt AI?

Hur ofta använder du AI?

Vad är din inställning till AI?

Vad är det största problemet med AI? Hur kan detta lösas?

Hur tror du att AI kommer att påverka framtiden?

Kan du exemplifiera fenomen, händelse eller dylikt som påverkat din etiska ståndpunkt vid användning och skapande av AI?

Har du lärt dig, eller fått insikter, under din studietid (genom kurser, externa föreläsningar) som gett dig en tillräcklig grund för etik inom mjukvaruutveckling?

Tror du att övergången från mänsklig till maskinell styrning av data kan relateras till samhället som stort?

Känner du att din attityd mot AI har förändrats under din studietid?

Upplever du dig själv som etisk medveten vid skapande av ny mjukvara? Hur tänker du kring de etiska aspekterna vid användning och skapande av AI-mjukvara?

Vilka etiska aspekter är mest relevanta för civilingenjörstudenter inom datavetenskap?

Tror du att det finns ett samband mellan demografiska faktorer och ökad användning av AI? Hur en individ tänker kring etiska frågor rörande detta område?

Tror du att det finns ett samband mellan mänskliga och samhällsliga faktorer som påverkar hur en individ ser på etiska aspekter vid AI-utveckling?

Flera stora techgudar har gått ut och varnat för AIs snabbt växande framkomst, vad tycker du om detta?