

Gabriele Braslauskaite

Forventninger til kunstig intelligens i norsk media på 1980-tallet

Bacheloroppgave i Historie

Veileder: Thomas Brandt

Mai 2024

Gabriele Braslauskaite

Forventninger til kunstig intelligens i norsk media på 1980-tallet

Bacheloroppgave i Historie
Veileder: Thomas Brandt
Mai 2024

Norges teknisk-naturvitenskapelige universitet
Det humanistiske fakultet
Institutt for historiske og klassiske studier



Kunnskap for en bedre verden

Forord

Denne besvarelsen markerer siste del av bachelorstudiet mitt i historie ved Norges teknisk-naturvitenskapelige universitet (NTNU), campus Dragvoll. Denne bacheloroppgaven har vært svært givende når det gjelder mestring, lærdom og personlig utvikling. Jeg har i løpet av dette semesteret fått mye kunnskap om interessante teorier og anvendelser innenfor teknologi og kunstig intelligens, noe som har vekket min nysgjerrighet til å selvstendig utforske temaene ytterligere.

Jeg vil først og fremst takke veileder Thomas Brandt for eksepsjonell støtte, veiledning og inspirasjon gjennom denne arbeidsprosessen.

Takk til familien min som har holdt motet mitt oppe under skrivesperrer, og takk til

Nasjonalbiblioteket for god tilgjengelighet til alle tidsskriftene jeg hadde bruk for.

Til sist, vil jeg takke Anders Aakervik Lyngsnes for moralsk støtte, og Tormod Skorpe

Skjolden for å ha bidratt til å dyrke min interesse for teknologi og informatikk.

Innholdsfortegnelse

Forord	1
1.0 Innledning	3
1.1 <i>Problemstilling</i>	3
1.2 <i>Kilder, metode og avgrensning</i>	4
1.3 <i>Oppgavens oppbygning</i>	4
1.4 <i>Historiografi</i>	5
1.5 <i>Forhistorie</i>	7
2.0 Tidlig KI-forskning – en berg-og-dalbane	9
2.1 <i>Tidlige prestasjoner og AI-vintre</i>	10
3.0 Analyse	12
3.1 <i>Ekspertsystemer</i>	14
3.3 <i>Ekspertsystemer i helsevesenet</i>	17
3.4 <i>Etikk og diskusjoner og ekspertsystemer i forhold til mennesket</i>	22
4.0 Konklusjon og diskusjon	27
5.0 Referanseliste	31
6.0. Bildekilder	32
7.0 Litteraturliste	32

1.0 Innledning

Betegnelsen «kunstig intelligens» (KI) dukker opp hyppig i dagens medier, og har blitt en integrert del av både vår hverdag og arbeidsliv. Til tross for det, fortsetter diskusjonene om nye mulige bruksområder og fremtidsløsninger med KI. Allerede nå har vi bl.a. ansiktsgjenkjenning på telefonen, og utviklingen av selvkjørende biler er allerede godt i gang. Likevel, fortsetter vi å være fascinert over fremtidens maskiner og har stadig forventninger til at de skal kunne være like intelligente som oss, men også kunne bli en høyere form for intelligens. Slike fascinasjoner og forventninger til KI er ikke noe nytt, men har vedvart i flere århundrer med varierende intensitet. Med dagens store fokus på kunstig intelligens, kan det derfor være interessant å ta et tilbakeblikk på hvilke tanker og ideer som florerte i den tidlige fasen av KI-utviklingen.

Begrepet «kunstig intelligens» ble etablert på 1950-tallet, men selve utviklingen av denne teknologien var langsom og var preget av flere opp- og ned turer. Selv om tiårene både før og etter 1950 har hatt sin betydning, vil denne besvarelsen se nærmere på 1980-tallets Norge og de forventningene som ble uttrykt i norske medier. Dette tiåret var preget av stort engasjement for KI-utvikling i store deler av verden, noe som resulterte i en mengde interessante diskusjoner. Sentralt i disse diskusjonene stod utviklingen og bruken av en spesifikk type kunstig intelligens kjent som «ekspertsystemer». Til tross for at land som USA var bedre rustet for en slik utvikling enn Norge på den tiden, var Norge likevel aktivt involvert i diskusjonene rundt det.

Denne besvarelsen vil utforske denne periodens diskusjoner om kunstig intelligens nærmere, med særlig fokus på ekspertsystemer og hvordan de ble omtalt i lys av ulike aspekter som etikk, arbeidsliv og anvendelser.

1.1 Problemstilling

Som oppgaven vil tydeliggjøre mer etter hvert, var dette tiåret sterkt preget av fremskritt innenfor teknologi og datavitenskap. Forskningsspørsmålet som denne besvarelsen skal ta utgangspunkt i, vil dermed være: Hvordan ble kunstig intelligens omtalt i mediene i Norge på 1980-tallet? I lys av forskningsspørsmålet, har jeg formulert tre problemstillinger: Hvordan ble endringer i norsk arbeidsliv diskutert i lys av forventningene til kunstig intelligens? Hvilke etiske dilemmaer/spørsmål dukket opp i diskursen om kunstig intelligens på 1980-tallet?

Hvilke konkrete anvendelser av kunstig intelligens så man for seg på 1980-tallet?

En slik innsnevring med nøye utvalgte perspektiver vil gi meg mulighet til å gå mer i dybden på temaet.

1.2 Kilder, metode og avgrensning

Undersøkelsen er gjort innenfor tidsrammen 1980-1989, av hensyn til struktur og de generelle oppgavekravene, med unntak av det forhistoriske kapitlet. Denne tidsrammen forholdte jeg meg til under søket etter kilder, men endte opp med å innsnevre tidsperioden til andre halvdel av 1980-tallet.

Tematikken er i utgangspunktet ganske omfattende, men første runden med søk avdekket at «ekspertsystemer» var et av de mest sentrale diskusjonstemaene knyttet til KI i denne perioden. Av den grunn, ønsker jeg i denne besvarelsen å gå nærmere inn på akkurat det. Hva som menes med ekspertsystemer vil bli nærmere gjort rede for i kapittel to og tre.

For å kunne gå noe mer i dybden på dette temaet, har jeg valgt ut fem avis- og tidsskriftsartikler; *Glade og sinte roboter...* av Peter Beck, rapporten *Kunstig intelligens og ekspertsystemer. Hva er det, og hva benyttes det til?* av Jan-Erik Kosberg, *Ekspertsystemer – Hva er det?* av Jan Christiansen, *Hvorfor en datamaskin aldri vil kunne tenke som et menneske* av Hubert og Stuart Dreyfus, og *Medisinske ekspertsystemer* av Truls Østbye og Knut Stavem. De tre første tidsskriftene er fra 1985, mens de to siste er fra 1987 og 1989. Alle kildene ble funnet på Nasjonalbiblioteket.no, der søkeordene «kunstig intelligens» og «ekspertsystem» var tilstrekkelig for å få opp ønskede resultater. Metoden for å finne frem til tidsskriftsartiklene har vært å gå inne på tidsskrift-kategorien på nettbiblioteket og benytte filterfunksjonen for å kun få opp tidsskrifter fra årene mellom 1980 og 1989. Det kan være viktig å bemerke at enkelte avisartikler, som *Glade og sinte roboter...* av Peter Beck, er kun tilgjengelig i «aviser»-kategorien.

Under arbeidet med kildene, var jeg også nødt til å ta hensyn til avisenes og tidsskriftets formål og eventuell politiske assosiering. Nasjonalbiblioteket har tilgang til en rekke norske tidsskrifter og aviser som omtaler kunstig intelligens på 1980-tallet, der enkelte aviser som for eksempel *Materialisten* og *Røde fane*, har en tydelig ideologisk profil. Politisk bakgrunn er viktig å være bevisst over i historiske analyser av kilder, da dette kan ha noe å si for hvilke typer meninger og tilnærminger som fremheves. Til tross for tilgjengeligheten til kilder som dette, har jeg valgt å fokusere på mindre ideologisk eksplisitte presseorganer for å studere hvordan KI ble diskutert i både fagtidsskrifter og mer allmenn dagspresse.

1.3 Oppgavens oppbygning

Kapittel én er den innledende delen av besvarelsen, der det gis en kort introduksjon til temaet kunstig intelligens. Videre i kapitlet presenteres historiografien med begrunnelser for litteraturvalgene, samt litt forhistorie for å legge grunnlaget for kapittel to.

Kapittel to består også av en historisk skildring, men med større vekt på historie som er mer direkte relevant for den valgte tidsperioden. Dette kapittelet danner grunnlaget for den endelige analysen i kapittel tre, der det utføres en analyse av de fem utvalgte tidsskriftsartiklene, supplert med sekundærlitteratur. I kapittel tre presenteres og diskuteres daværende tanker rundt temaer som robotrettigheter, automatisering, ekspertsystemer, samspill mellom mennesket og maskinen, samt etiske implikasjoner. Til slutt kommer kapittel fire, som utgjør den konkluderende del som også rommer noen diskusjoner og sammenligninger. De aller siste kapitlene består av referanse- og litteraturliste, samt bildekilder.

1.4 Historiografi

I mitt arbeid med denne problemstillingen har jeg, i tillegg til norske tidsskriftsartikler fra 1980 årene, benyttet noen bøker og artikler for å få en dypere forståelse av temaet kunstig intelligens. For å få en historisk oversikt, har jeg benyttet meg av boka *Maskiner som tenker* av Inga Strümke, supplert med artiklene *History of AI Winters* av Milton Lim og *Artificial Intelligence* av Stephanie Dick og *Artificial Intelligence and Keeping Humans «in the Loop»* av Robert Mazzolin, for å få bedre innsikt i ulike teorier og perspektiver innen KI som har vært relevant for besvarelsen.

Strümkes *Maskiner som tenker* går i dybden på en rekke spørsmål angående kunstig intelligens. Boka gir blant annet grundig forklaring på hva kunstig intelligens er, viktige aspekter og bidragsytere i utviklingen, samt hvordan det kommer til å påvirke den moderne verden. Kapittel én og to («Litt nøye utvalgt historie» og «hvordan vi prøver å gjøre maskiner intelligente»), har vært til særlig nytte, da det gis gode forklaringer og bakgrunn for ting som algoritmer, ekspertsystemer og ELIZA-effekten. Bokas historiske skildringer har blitt mye brukt som grunnlag for videre forskning fra min side. Selv om Strümke gir et oversiktlig innblikk i ulike aspekter ved kunstig intelligens, er det viktig å bemerke at Strümke ikke er en utdannet historiker. Hennes historiske fremstilling kan derfor inneholde forenklinger eller skjevheter, da hennes primære fokus er på KI mer generelt i dag, og ikke på en nøyaktig historisk redegjørelse. Av den grunn, har jeg ikke utelukkende basert meg på Strümkes historieskildring. En annen bok som ble brukt i denne besvarelsen er *Robot Rights* skrevet av den amerikanske akademikeren David J. Gunkel i 2018. Boka bidro til en dypere forståelse av de forskjellige sidene i den pågående debatten om robotrettigheter, samt at den skapte mer grunnlag for refleksjoner rundt avisartikkelen *Glade og sinte roboter...* av Peter Beck.

Artikkelen *Artificial Intelligence* av historiker Stephanie Dick handler om historien til kunstig intelligens og diskuterer særlig nødvendigheten av å se historien i lys av en lengre utvikling enn at det «startet» i Dartmouth i 1955. Dick fremhever viktige faktorer innen KI-utvikling som etter hennes mening undervurderes i mange av historieskildringene. Disse faktorene inkluderer bl.a. det overordnede målet med automatisering, samt endringer i måten vi definerer og oppfatter intelligens. Dicks synspunkter er interessante å se på som et tillegg til andre historiske fremstillinger, da det tilbyr et kritisk perspektiv og dypere forståelse av feltet og dets historie.

Artikkelen *History of AI Winters* av teknologiskribenten Milton Lim dekker tidlig KI historie og de mange opp- og nedturene innenfor KI-forskning. Lim gir en kortfattet oversikt over de viktigste milepælene og utfordringene innenfor KI-forskning, der han trekker frem begivenheter som Turingtesten, læringsalgoritmen perceptron, AI-vintre og Moravecs paradokset. Selv om Lim ikke går like mye i dybden på AI-vintrene som Strümke, bidrar hans artikkel til et mer variert spekter av litteratur.

I artikkelen *Current research on Gödel's incompleteness theorems* presenterer professor Yong Cheng en oversikt over aktuell forskning på Gödels ufullstendighetsteoremer, med fokus på klassifisering av ulike bevis for teoremene, grenser og anvendelighet. Chengs artikkel har vært relevant, da den gir konkrete definisjoner og eksempler på teoremene. I all hovedsak ble artikkelen valgt ut på grunn av personlig interesse for å utdype min forståelse av Gödels arbeid.

I artikkelen *The carousel of ethical machinery* av professor i informatikk, Luís Moniz Pereira, fremmes et litt annerledes perspektiv på KI-historie. Artikkelen drøfter de etiske implikasjonene ved teknologisk fremskritt, der det særlig legges vekt på økonomiske og sosiale konsekvenser som følge av bl.a. automatisering. Artikkelen gir en kort fremstilling av greske mytologiske forløpere når det gjaldt tidlig bevissthet rundt potensialet til intelligente maskiner. Artikkelen gir også et mer dyptgående perspektiv på KI-historie sett fra en mer uvanlig vinkel. En annen artikkel som kobler opp gresk mytologi og kunstig intelligens, er *Fra Eliza til ChatGPT*, skrevet av redaktørsjef for *Tidsskrift for Den norske legeforening*, Are Brean. I denne artikkelen gjør Brean en historisk sammenligning av Eliza og ChatGPT, der han bruker det greske sagnet om Pygmalion for å illustrere utfallet av utviklingen til denne typen teknologi.

Artikkelen *Imagining the thinking machine: Technological myths and the rise of artificial intelligence* av Simone Natale og Andrea Ballatore, diskuterer rollen til teknologiske myter i utviklingen av kunstig intelligens. I selve besvarelsen har denne artikkelen kun blitt sitert én

gang, men på samme måte som Cheng og Lims artikler, har denne bidratt til å skape et mer nyansert bilde av temaet.

Den siste artikkelen som skal presenteres i dette kapitlet er *Artificial Intelligence and Keeping Humans «in the Loop»* av teknologistrategen Robert Mazzolin. Denne artikkelen utforsker spørsmål om KI-autonomi, menneskelig tilsyn og etiske implikasjoner ved å overlate beslutninger til KI-systemer. Med andre ord, diskuteres de forskjellige sidene ved «Human-in-the-loop» paradokset, som ofte dukker opp i min analyse. Selv om det legges mest vekt på KI-styrte våpen, kan eksemplene som oppgis også anvendes på det som blir diskutert i denne besvarelsen.

1.5 Forhistorie

«Informasjonssamfunnet» eller «kunnskapssamfunnet» er kjente betegnelser på den moderne epoken vi befinner oss i nå. Til tross for at begreper som bl.a. kunstig intelligens og algoritmer kan virke moderne for de aller fleste, strekker historien seg lengre bak enn det man skulle trodd.

Tanker om en intelligent/tenkende maskin fantes trolig før kunstig intelligens engang ble et fagfelt. Det har vært en rekke med elementer, og ikke minst individer, som har både direkte og indirekte bidratt til å legge grunnlaget for utviklingen. Det mest logiske å starte med kan være oppfinnelsen av algoritmen, da det er mer direkte knyttet til utviklingen av KI. Selve tanken om en slik intelligens var allikevel en annen sak. I artikkelen «The carousel of ethical machinery», skrevet av Luís Moniz Pereira, settes dagens forståelse av kunstig intelligens i sammenligning med antikkens myter, historier og tanker rundt teknologisk oppfinnelse. Pereira nevner bl.a. at i den hellenistiske perioden ble det produsert maskiner drevet av vann eller luft, som medførte en fascinasjon, samt økt bevissthet blant grekerne om både effektiviteten og fare ved maskineri. Historien om Hefaistos, den greske guden av ilden, smedkunsten og håndverket¹, nevnes særlig her som en slags tolkning av tidligere forståelser om teknologisk innovasjon;

«Hephaestus, the God of technological invention, foretelling already our future, built a fleet of tripod stools “without driver” that responded to commands to deliver food and wine to the Gods, who were not to be encumbered with such trivia. More remarkable still was the bundle of life-sized gilded female robots he had created to carry out his

¹ Bouvrie, 2018

orders. According to Homer, these servants of the divine were— in every way—“as real young women, with sensations and reason, strength, and even voice.”²

Poenget med dette eksempelet er ikke for å påstå at kunstig intelligens har sine røtter i antikken, men heller for å illustrere at tanken om menneskeskapt intelligens har eksistert ganske lenge. Dette kapitlet skal allikevel legge mest vekt på de mer direkte bidragene til utviklingen av kunstig intelligens.

Så tidlig som i 1931, presenterte Kurt Gödel en teori som viste at «ingen matematiske systemer kan være både komplette og konsistente»³, for å sitere Strümke. Denne teorien (eller teoriene), er kjent som ufullstendighetsteoremer, og har vist seg å spille en viktig rolle innenfor ikke bare matematikk, men også bl.a. logikk, filosofi og datavitenskap. Teoremet kan være vanskelig å forstå seg på ved første øyekast, men enkelt forklart, viser den bl.a. at det er begrensninger på hva som kan oppnås med et formelt system. Et system, om den er kompleks nok, kan altså ikke bevise sin egen konsistens, samt oppdage motsigelser som for eks. $0 = 1$, ved bruk av sine egne regler.⁴ Alan Turing lot seg fascinere av Gödels teoremer og forsøkte å anvende det på datavitenskap. I 1937 beviste Turing to innflytelsesrike ting – umuligheten av et dataprogram som kan finne ut om en algoritme kan fullføres (the halting problem)⁵, og at det kan finnes maskiner som kan løse ethvert matematisk problem så lenge den er formulert som en algoritme (Turingmaskinen).⁶ Så langt var man allikevel ikke i mål helt enda, da man kun hadde et konsept som enda ikke hadde blitt realisert. Matematikeren John von Neumann var den som sto bak arkitekturen vi bygger moderne datamaskiner ut fra. Den såkalte Neumann-arkitekturen besto av to komponenter – én for lagring og én for beregning. I tillegg besto arkitekturen av enheter som gjorde det mulig for et menneske å faktisk kommunisere med maskinen. Kommandoer fra mennesket tas inn og overføres til beregningsenheten, mens en annen enhet hentet informasjon fra beregningsenheten slik at mennesket fikk nytte av den.⁷ Den første Neumann-maskinen sto klar i 1952 og var revolusjonerende for datavitenskapens historie.⁸ Det kan på mange måter trekkes noen paralleller mellom den «originale» Neumann-arkitekturen og de moderne datamaskinene vi har i dag, men med det forbehold at teknologien har blitt langt mer avansert opp gjennom årene.

² Pereira, 2020, s. 185

³ Strümke, 2023 s. 7

⁴ Cheng, 2020, s. 113

⁵ Strümke, 2023 s. 7

⁶ Strümke, 2023 s. 8

⁷ Strümke, 2023 s. 8

⁸ Strümke, 2023 s. 9

2.0 Tidlig KI-forskning – en berg-og-dalbane

Kunstig intelligens som et eget fagfelt sies å ha sine røtter i en liten workshop ved Dartmouth College i USA på 1950-tallet, ledet av forskeren John McCarthy. Med likesinnede medstudenter forsket McCarthy på disse, enda ganske ukjente, «intelligente datamaskinene». Det var også i denne sammenhengen at begrepet *artificial intelligence (AI)*, eller kunstig intelligens (KI), først ble brukt.⁹ Denne skildringen av KI-historiens «begynnelse» er den mest anerkjente skildringen som ganske mange tar utgangspunkt i. Det er allikevel viktig å anerkjenne at utviklingen av kunstig intelligens som fagfelt er et resultat av flere tiår med både direkte og indirekte innsats og bidrag fra ulike fagfelt.

Vitenskapshistoriker Stephanie Dick skrev i 2019 en artikkel der hun diskuterte historien til kunstig intelligens og hevdet at den er mer kompleks enn det historiske narrativ vi er kjent med.

«[...] artificial intelligence wasn't born at Dartmouth in 1955, as the standard account would have us believe, but rather participates in much longer histories of what counts as intelligence and what counts as artificial.»¹⁰

Poenget til Dick her, er at utviklingen av KI må sees i lys av en lengre utvikling enn kun Dartmouth-narrativet. Dick sier eksplisitt, at historiske beretninger om kunstig intelligens ofte overdriver betydningen av Dartmouth-workshopen, og at til og med deltakerne selv mente at workshopen ble mer til en skuffelse;

«McCarthy recollected that «anybody who was there was pretty stubborn about pursuing the ideas that he had before before he came, nor was there, as far as I could see, any real exchange of ideas.»¹¹

Dick fremhever også viktigheten av vår forståelse av hva intelligens er og hvordan det har endret seg opp gjennom tiden. Dicks budskap bør allikevel ikke tolkes som en fullstendig bagatellisering av workshoppens historiske rolle, men heller som noe som ikke var et særlig betydelig gjennombrudd. Det fremheves altså at det eksisterte andre tilnærminger til det fra før av, som for eks. innen matematikk- og filosofiforeningene.¹²

Samtidig som det er viktig å anerkjenne mindre og større bidrag som kom før, bør man ikke undervurdere det som ble oppnådd ved Dartmouth-workshopen. Selv om det bygget på tidligere kunnskap og kanskje ble mer anerkjent enn andre mindre bidrag, bidro det samtidig

⁹ Strümke, 2023 s. 17

¹⁰ Dick, 2019 s. 3

¹¹ Dick, 2019 s. 3

¹² Dick, 2019 s. 2

til å legge grunnlaget for at KI forskning ble satt på dagsorden, la grunnlag for eksperimentering og skapte større økonomisk interesse.

2.1 Tidlige prestasjoner og AI-vintre

Et kjent dataprogram som ble utviklet på 1960-tallet av Joseph Weizenbaum var Eliza. Ved å identifisere spesifikke ord i setningene som programmet fikk servert, kunne Eliza konstruere et svar ut fra det.¹³ Figur 1 på neste side illustrerer hvordan et slikt program fungerte. Denne nye teknologien ble raskt tatt i bruk for terapeutiske formål, men falt gradvis i popularitet og ble et offer for de samme fallgruvene som andre KI-systemer på den tiden. Som de fleste KI-systemene på det tidspunktet, var Eliza meget begrenset når det gjaldt teknologisk nivå, språk og nytteverdi. Betydelig store summer ble investert i ulike typer kunstig intelligens i løpet av 60-tallet, men finansieringen og interessen avtok da man virkelig innså kompleksiteten i å skape velfungerende KI-systemer. Både på 60- og 70-tallet var det håp om at tilnærmingene skulle føre til maskiner med generell intelligens.¹⁴ Slik ble det altså ikke, og det ble fort tydelig at det *er* grenser for hva en maskin kan gjøre.

Eliza var et tidlig eksempel på et program som simulerte menneskelig samtale. Til tross for store begrensninger innen datateknologi, demonstrerte Eliza potensialet til kunstig intelligens og representerte et betydelig steg i kommunikativ teknologi.

¹³ Strümke, 2023 s. 38

¹⁴ Strümke, 2023 s. 21

```
Welcome to

      EEEEE LL      IIII ZZZZZZ  AAAAA
      EE     LL      II     ZZ   AA  AA
      EEEEE LL      II     ZZ   AAAAAA
      EE     LL      II     ZZ   AA  AA
      EEEEE LLLLLL IIII ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   The research on AI is consuming me
ELIZA: Tell me more about that.
YOU:   I want to research even more!
ELIZA: What would it mean to you if you got to research even more ?
YOU:   That would be cool
ELIZA: Does talking about this bother you ?
YOU:   No?
ELIZA: Are you saying no just to be negative?
YOU:   no?
ELIZA: Why 'no' ?
YOU:   █
```

15

Figur 1: En kort samtale mellom meg og Eliza. Eliza var en simpel samtalepartner som var avhengig av spesifikke nøkkelord for å kunne stille oppfølgings spørsmål. Dersom man ikke svarte spesifikt nok, slik som jeg gjorde her, svarer Eliza gjentakende på en nesten passivt aggressiv måte.

Forskning innenfor dette feltet ble for det meste lagt til siden før det «boomet» igjen på 80-tallet.¹⁶ Slike nedganger i interesse og finansiering, er kjent som såkalte «AI winters». Ifølge artikkelen «History of AI winters» av Milton Lim, har det i KI historie vært to betydelige AI vintré – den første i 1974 – 1980, den andre i 1987 – 1993. Disse nedgangsperiodene har vært viktig for KI historien, da det også kunne ha hatt en læringseffekt for forskerne. Et eksempel på dette er da ekspertsystemer som XCON, LISP-maskiner og Symbolics ble svært populære på 1980-tallet. Ekspertsystemene var i utgangspunktet nyttig for å løse spesifikke problemer som for eksempel å diagnostisere sykdommer, men det viste seg etter hvert at ekspertsystemene var rett og slett for dyre å vedlikeholde, samt at de slet med å lære og var vanskelig å oppdatere.

I løpet av bare ett år ble de dyre Lisp-maskinene byttet ut med rimeligere og mer brukervennlige stasjonære datamaskiner fra produsenter som Apple og IBM. Sistnevnte eksempel fungerer også som et eksempel på det fundamentale Moravecs paradokset. Paradokset representerer en begrensning i KI-forskning, som går ut på at det er lett å få en maskin til å utføre oppgaver som er vanskelig for mennesker, men at det er vanskelig å få en maskin til utføre oppgaver/handlinger som er lett for mennesker.

¹⁵ Landsteiner, 2005

¹⁶ Strümke, 2023 s. 41

«It is comparatively easy to make computers exhibit adult level performance on intelligence tests, playing checkers or calculating pi to a billion digits, but difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility [...]»¹⁷

Lims artikkel sier ikke noe eksplisitt om hvilken retning industrien for spesialiserte maskiner gikk etter kollapset av markedet, men det kan nok så trygt antas at forskere ble nødt til å etter hvert vurdere andre tilnærminger. Dette peker også mot en mulig årsak til utfordringene innen KI-forskning. Det kan tolkes slik at det kanskje var for stort fokus på feil plass, og da spesielt hvis man tar utgangspunkt i Stephanie Dicks skildring om at KI-forskning først og fremst hadde som mål om å skape systemene så menneskelig som mulig.¹⁸ Samtidig, var det ganske mye som indikerte at det å etterligne menneskelig kognisjon i for eks. utviklingen av ekspertsystemer, var nærmest umulig å oppnå. Dette skal undersøkes nærmere senere i besvarelsen.

3.0 Analyse

Kunstig intelligens var et omdiskutert tema på tvers av alle samfunnsgrupper på 80-tallet, og gjorde seg synlig/hørt i de fleste mediene. Diskusjonene var slett ikke svart og hvitt, da det ofte syntes en blanding av frykt, kritikk, optimisme eller nysgjerrighet. Dette kapitlet skal se nærmere på de ulike omtalelsene.

Tidlig KI-forskning var ikke kun en berg-og-dalbane utviklingsmessig, men også i måten den generelle befolkningen reagerte på det. Som nevnt tidligere, var dette et svært omdiskutert tema, og forventningene var også på samme måte varierende. Dette fenomenet reflekteres godt i datidens medier, som for eks. aviser og andre tidsskrifter, som det skal sees eksempler på. I avisartikkelen *Glade og sinte roboter...* publisert i *Aftenposten* 1985, dekkes en rekke temaer rundt «robotisering» av yrker, computerhjerner på menneskelig nivå, robotmennesker, og, interessant nok, robotrettigheter.¹⁹ Diskusjonen skjedde med utgangspunkt i en annonse i et amerikansk tidsskrift, som denne avisartikkelen siterer;

«Bestill angreps-roboten «Doberman» snarest. Den vil straks gå til angrep på alle inntrengere i Deres leilighet, firma eller der De måtte ønske beskyttelse».²⁰

¹⁷ Lim, 2018

¹⁸ Dick, 2019 s. 2

¹⁹ Beck, 1985

²⁰ Beck, 1985

Frykten for automatisering, eller i dette tilfellet, robotisering av yrker var ikke noe særlig overraskende i dette tiåret. Det var enda en tid med lite forskning i dette området, og hvordan fremtiden skulle påvirkes av maskinene var vanskelig å predikere nøyaktig. Bekymringene ble kanskje særlig større da forskere uttalte seg om det, som følgende sitat illustrerer;

«Enkelte yrker står allerede idag i fare for å bli «robotisert», om vi bare slipper «Doberman»s mer vennligsinnede slektninger til for alvor. Sekretær- og kontoryrket – og spesielt jusen – nevnes av forskerne som særlig godt egnet for robotene.»²¹

Hvorvidt det var snakk om ekspertssystemer eller faktiske robotmennesker som de såkalte «robot-sekretærene» nevnt i starten av artikkelen, er ikke helt klart. Det kan antas at det kanskje menes begge, men med tanke på hvor mye det snakkes om de fremtidige robotmenneskene, så var nok antakelsene mer dystopiske;

«Om noen få tiår kan vi lage roboter som ligner mennesker [...] Robotene vil ikke kunne erstatte menneskets hjerne, men det forventes at de kan lages meget «smarte». Enkelte yrker kan til og med bli erstattet av roboter, hevdes det. Noen forskere mener at vi i fremtidens samfunn kan gå i supermarkedet (under) for å kjøpe vår egen robot-sekretær, robot-jurist, robot-lærer, robot-soldat ...»²²

Selv om avisartikkelen antydte muligheten for robotmennesker som skulle «traske» blant oss, trekkes det et klart skille mellom mennesket og maskinen, som for eks. i sitatet ovenfor. Det anerkjennes allikevel samtidig at dens intelligens hadde potensialet til å overgå menneskets; «om 10-15 år kommer en kunstig hjerne til å arbeide like fort som menneskets – og ikke lenge derefter enda fortere, mener Koch.»²³ Det dukker allikevel opp en rekke med spørsmål når det er snakk om disse robotmenneskene. I artikkelen nevnes nemlig også spørsmålet om hvorvidt roboter skulle få samme rettigheter som mennesker; «Jeg tror faktisk vi må ta stilling til slike spørsmål allerede om 10-20 år, sier Christof Koch.»²⁴

Det blir et dilemma når man skal sette et så skarpt skille mellom mennesket og robotmennesket når det kommer til intelligens, samt kognitive/motoriske ferdigheter, og samtidig forholde seg til noe menneskeskapt med hjernekapasiteten til å gjøre den akkurat menneskelig nok. Det dukker også opp et spørsmål om hvilke kriterier man skal ta utgangspunkt i, altså om man skal tenke på det som en utvidelse av rettigheter forbundet med mennesker, eller rettigheter knyttet til ikke-menneskelige ting som har nytteverdi for oss som

²¹ Beck, 1985

²² Beck, 1985

²³ Beck, 1985

²⁴ Beck, 1985

for eks. miljøet. I boka *Robot Rights* av akademiker David J. Gunkel, påpekes det at med en økende innflytelse av maskinenes posisjon i forskjellige områder i samfunnet, så vil det som resultat dukke opp en del viktige spørsmål;

«At what point might a robot, algorithm, or other autonomous system be held accountable for the decisions it makes or the actions it initiates? When, if ever, would it make sense to say, “It’s the robot’s fault”? [...] When, in other words, would it no longer be considered nonsense to inquire about the rights of robots?”²⁵

Dette kunne godt være tilfellet også på 1980-tallet. Det var tross alt på 1980-tallet at intelligente systemer virkelig begynte å få en sterkere rolle i samfunnet, med ledende fokus på å gjenskape menneskelig intelligens. I lys av dette er det ikke utenkelig at det kunne skape sterke reaksjoner hos enkelte. Forskernes bidrag med å forsterke bekymringene og forventningene, slik det fremstår i avisartikkelen, spilte nok også en rolle i dette.

Avisartikkelen inneholder flere perspektiver på robotteknologien, som til tross for at de kunne stride mot hverandre litt, viste høy tro på hva kunstig intelligens kunne få til på både godt og vondt. Et interessant aspekt ved denne avisartikkelen var nok diskusjonen om robotens plass i samfunnet, da det i enkelte deler av teksten, virket som at roboten ble ansett mer som en erstatning av mennesket, enn en utvidelse.

3.1 Ekspertsystemer

Ekspertsystemer er et uunngåelig tema i kunstig intelligens historie. Ekspertsystemenes eksistens var på flere måter revolusjonerende for effektivisering av arbeidslivet på 1980-tallet, og da spesielt innenfor bl.a. helsesektoren og ingeniørvirksomheten. Ekspertsystemene spilte ikke kun en rolle i forsøkene på å effektivisere og videreutvikle, men bidro også til å dyrke forventninger til hva kunstig intelligens kunne bli til. Førsteamanuensis i medieteorier og historie, Simone Natale, og geografisk dataforsker, Andrea Ballatore, skrev i 2020 artikkelen *Imagining the thinking machine: Technological myths and the rise of artificial intelligence*, som diskuterer rollen av teknologiske myter i utviklingen av kunstig intelligens mellom 1950 og tidlig på 1970-tallet. Artikkelen beskriver ekspertsystemenes rolle på 1980-tallet slik;

«The wave of ‘expert systems’ in the 1980s generated viable and profitable applications, but fostered at the same time novel expectations about AI.»²⁶

Ekspertsystemer var kanskje ikke like utbredt i Norge på dette tidspunktet, men impulser fra utlandet fostret raskt forventninger som kun økte med årene.

²⁵ Gunkel, 2018, s. 5

²⁶ Natale & Ballatore, 2020, s. 13

Før det går nærmere inn på dette temaet, kan det først og fremst være lurt å avklare definisjonen på ekspertsystemer. Dette blir også relevant for diskusjonen rundt måten kunstig intelligens ble omtalt innenfor helsevesenet og andre sektorer senere i besvarelsen. En god kilde for å forklare ekspertsystemenes betydning slik de fremsto i 1985, er Elab rapporten av Jan-Erik Kosberg, med tittelen *Kunstig intelligens og ekspertsystemer. Hva er det, og hva benyttes det til?*

I rapporten, defineres ekspertsystemer som følger;

«Et ekspertsystem er et intelligent datamaskinprogram som bruker eksisterende kunnskaper og konkluderende prosedyrer for å løse problemer som er så vanskelige at de ellers ville kreve menneskelig ekspertise.»²⁷

Forklart litt enklere, tas kunnskapen fra menneskelige eksperter og pakkes inn i et boks som står tilgjengelig for en enda større brukergruppe. Noen av eksemplene på slike ekspertsystemer gitt i rapporten var bl.a. selskapets General Electric (GE) bruk av programmet CATS-1/DELTA for vedlikehold av diesellokomotiv, og ekspertsystemer i helsesektoren med evnen til å diagnostisere med en tillitsrate på 85%.²⁸ Sett fra et moderne perspektiv, er ikke en slik tillitsrate særlig høy, og det hadde nok vært mer skepsis rettet mot dette om noe lignende ble presentert i dag.

Rapporten la stor vekt på diskusjoner rundt utvikling av KI, markedspregninger, samt såkalte «femte generasjons prosjekter» knyttet til avanserte datamaskiner i USA, Japan og Europa. Til tross for fokuset på amerikanske studier, gis det også en oversikt over Norges planer rundt satsing på kunnskapsbasert teknologi og ekspertsystemer, noe som dette kapitlet særlig skal behandle.

Når det gjaldt norsk satsing på automatisering og kunstig intelligens, la NTNMF (Norges Teknisk-Naturvitenskapelig Forskningsråd)²⁹ frem tre satsingsområder; mikroelektronikk, datateknikk, kommunikasjons- og informasjonssystemer, og automatisering.

Ekspertsystemer beskrives som et eget «spesielt» område, men som opplevde økt interesse og forventninger;

«Ekspertsystemer er et felt hvor det satses svært mye på FoU internasjonalt. I Norge befinner vi oss i en oppstartingsfase idet en del miljøer nå interesserer seg for feltet uten at en hittil har utviklet konkrete systemer [...]»³⁰

²⁷ Kosberg, 1985 s. 4

²⁸ Kosberg, 1985 s. 4

²⁹ Li, 2022

³⁰ Kosberg, 1985 s. 18

Dette tyder på at behovet og interessen for slik teknologi økte fortere enn selve utviklingen. I rapporten trakk NTNf internasjonale eksempler, der slik teknologi var «i ferd med å gi resultater» i for eks. medisinsk diagnostikk. I tillegg, påpekes det at teknologien kunne effektivisere felt som økonomi, geologi og mekanisk konstruksjon, samt løse problemer på en mye enklere måte enn det man klarte før.³¹ Det kan tolkes slik at satsingen på teknologi fikk det meste av sin motivasjon fra utenlandske resultater. Å ta utgangspunkt fra utenlandsk forskning o.l. er ikke nødvendigvis noe negativt, men for sterk avhengighet kan potensielt føre til at landet står sårbart for en rekke utfordringer. Det skal ikke gås inn i dybden på det her, men hvis man tenker på det så er det på den ene siden lite effektivt på lang sikt. Det kunne bl.a. gå utover nasjonale behov for innovasjon, samt kontroll over egen teknologisk utvikling. På den annen side er det en urealistisk forventning at land som Norge på spesielt den tiden skulle kunne utføre uavhengig forskning og utvikling innenfor ekspertssystemer fra «scratch». Dette var åpenbart noe som aktørene her var klar over allerede, da de påpekte selv at Norge hadde for svakt utgangspunkt til å sette noe slikt i gang helt selv. Man kan altså se at det var et stort fokus på å få bygd opp kompetanse for selvstendig utvikling, samtidig som at det ikke skulle neglisjeres at utenlandske erfaringer var like viktig.

Eksempler som ekspertsystemet MYCIN for medisinsk diagnostikk, som er nevnt i rapporten, trigget muligens tanker om potensielt andre muligheter systemene kunne gi, eller inspirerte til å utforske nye bruksområder for det man allerede hadde tilgang til. Gitt landets «svake utgangspunkt» oppga NTNf tre hovedområder Norge bør satse på innenfor ekspertssystemer; generell opplæring i nøkkelmiljøer, etablere fast kontakt med spissmiljøer i utlandet, og støtte et utvalg av konkrete, bruksrettede prosjekter som kan gi både erfaring og resultater på kort sikt.³² Med andre ord ble det oppfordret til å satse på økt kompetanse innenfor feltet, dra nytte av utenlandsk ekspertise og støtte konkrete prosjekter for å generere så mye erfaring som mulig. Mot slutten av rapporten ble det i tillegg oppfordret til bl.a. utvikling av problemorienterte FoU- programmer (POP), som innebar konkret problemløsning ved hjelp av ekspertssystemer. Disse programmene skulle identifisere konkrete problemer innen ulike fagområder som for eks. medisin og ingeniørfag.³³

«Tiden skulle nå være moden for at FoU-instituttene og brukerinteressentene går sammen om å formulere slike problemorienterte programmer. En felles satsing på de identifiserte områder skulle kunne gi norske FoU-institutter, industribedrifter og

³¹ Kosberg, 1985 s. 17

³² Kosberg, 1985 s. 18

³³ Kosberg, 1984 s. 19

næringsliv forøvrig en bedre plattform å stå på i møte med de kommende års utfordringer.»³⁴

Som dette sitatet illustrerer, var det et stort engasjement for norsk satsing innenfor ekspertssystemer. Hvorvidt tiden egentlig var moden for en slik utvikling kan sees på fra flere sider. På den ene siden, kan man argumentere for at tiden ikke var fullt ut moden for en slik satsing. Basert på tilgjengelig informasjon, er det lite som tyder på at ekspertssystemer ble tatt i bruk i noe særlig omfang på 1980-tallet. Det er dermed bemerkelsesverdig, at til tross for store ønsker om utvikling og samarbeid, at det var dette som var tilfellet. På den annen side, kan man argumentere for at selv om utviklingen ikke kom så langt på 1980-tallet, så var tiden inne for å i det minste legge et grunnlag for fremtidig innovasjon. På samme måte som Dartmouth-workshopen satte kunstig intelligens på dagsorden, så kunne en langtidsplan som dette ha bidratt til å sette utvikling av ekspertssystemer på kartet for fremtidig forskning.

3.3 Ekspertsystemer i helsevesenet

Med en relativt rask utvikling innen teknologi, ble implementeringen av KI i norsk arbeidsliv mer og mer sannsynlig. Et ressurskrevende gren som helsevesenet, kan være et godt eksempel på dette tilfellet. I avisartikkelen *Ekspertsystemer – hva er det?* fra 1985, diskuterte forsker og metodeleder ved Statens Institutt for Folkehelse, Jan Christiansen, bruksområdene for ekspertsystemer, samt etiske/moralske problemstillinger som kom med det. Med tanke på at artikkelens lengde og rikt innhold, skal det fokuseres på nøye utvalgte aspekter som kan gi svar på problemstillingene. Christiansens diskusjoner skal i tillegg suppleres med avisartikkelen «Medisinske ekspertsystemer» fra 1989, skrevet av professorene Truls Østbye og Knut Stavem.³⁵ I denne avisartikkelen diskuterte Østbye og Stavem hvordan spesifikt Norge bør satse på medisinske ekspertsystemer, noe som også Christiansen tok opp kort mot slutten.

Innholdet i Christiansens avisartikkel inneholder både bevissthet rundt ekspertsystemenes begrensninger, samtidig som det fremhever potensielle bidrag de kan komme med.

Sammenlignet med avisartikkelen nevnt tidligere, Glade og sinte roboter... (1985), kan man trekke små forskjeller i måten ekspertsystemer ble omtalt. Her ble ikke ekspertsystemene ansett som en trussel i like stor grad når det gjaldt å «erstatte» mennesker, selv om ordet «erstatte» benyttes eksplisitt som en enten/eller i Christiansens definisjon av

³⁴ Kosberg, 1985 s. 20

³⁵ Duke Global Health Institute, (u.å.); Universitetet i Oslo, 2012

ekspertsystemer.³⁶ Til tross for det, var samarbeidet mellom ekspertsystemet og mennesket i denne avisartikkelen i større fokus enn det vi har sett tidligere.

Helt i starten påpekte Christiansen forskjellen mellom mennesket og ekspertsystemet når det gjaldt utføring av oppgaver. Han sammenlignet teknologien med øks, kniv, blyant, papir» osv., som er alle midler til å hjelpe oss utføre oppgaver raskere og mer effektivt. Ifølge Christiansen, lå altså forskjellen mellom datamaskinen og mennesket i det kvantitative, ikke det kvalitative. Datamaskinen jobber tross alt ut fra den informasjonen den blir gitt, og har ikke evnen til å tenke selvstendig eller gå i dybden i sitt arbeid slik som mennesker;

«Det er ikke noe spesielt med datamaskinen hvis vi sammenligner med de tradisjonelle verktøyene for tankearbeid – blyant og papir, fred og ro, bøker. Ingen kvalitativ forskjell. Den store utfordringen i datateknologien er en kvantitativ forskjell.

Datamaskinen setter oss i stand til å utføre oppgaver som ikke bare ville tatt umåtelig lang tid med de tradisjonelle metodene, men som ville være for kompliserte og involvere for mange steg til at et menneske eller en gruppe mennesker ville klare det.»³⁷

Det som er vel så interessant, er hvordan ekspertsystemer ble omtalt i forhold til arbeidsliv og den menneskelige ekspertisen som var del av det. Christiansen oppsummerte en rekke med systemer som var i vanlig bruk innenfor bl.a. helsevesenet: databasesystemer, simulering og modellbygging, beslutningsstøtte, systemer for bevisføring og mønstergjenkjenning. Det endelige ekspertsystemet ble da som et slags produkt av disse systemene kombinert;

«Ekspertsystemene er rådgivningshjelpemidler på samme måte som beslutningsstøttesystemene, men med adskillig større «egenvilje» når det gjelder å presentere og innhente relevant informasjon. De benytter seg av teknikker fra modellbygging, databaseteknologi, automatisk bevisføring og mønstergjenkjenning.»³⁸

Med andre ord, menes det at ekspertsystemene var avanserte nok til å faktisk søke etter relevant informasjon og presentere det forståelig. Det virker kanskje ikke så imponerende lenger fra et moderne perspektiv, men på den tiden kunne det regnes som et stort skritt innen intelligente systemer, samtidig som en trussel for enkelte yrker.

Christiansen presenterte videre noen bruksområder for ekspertsystemene, som bl.a. tolkning av situasjonsbeskrivelser, diagnose, opplæring, planlegging m.m. Det ble også oppgitt noen eksempler på hvordan et slikt system kunne se ut;

³⁶ Christiansen, 1985, s. 4

³⁷ Christiansen, 1985, s. 3

³⁸ Christiansen, 1985, s. 4

«Et typisk eksempel vil være et system som skal gi pleiepersonale råd om øyeblikkelig hjelp i de tilfellene der det ikke finnes eksperter – leger – til stede. Et annet eksempel er rådgivning til allmennpraktikere når det gjelder behandling av sykdommer som vanligvis krever konsultasjon av en spesialist.»³⁹

Selv om dette på mange måter var effektivt og lovende, var det fortsatt snakk om et system som jobbet ut fra ferdigprogrammerte regler og kunnskap. Christiansen påpekte selv at «ekspertsystemer vil garantert ikke være feilfrie», noe som ville pålegge et større ansvar for eventuelle brukere. Med ansvar menes altså bevissthet rundt driftsstabilitet og kvaliteten på rådene systemet ga. Samtidig som dette forebygget et potensielt problem, skapte det et nytt et. Hensikten med ekspertsystemer var for å effektivisere og avlaste, men om man konstant var nødt til å overprøve systemet, ble mennesket påført mer arbeid og hele hensikten ble da borte;

«En restriktiv holdning på dette området fører jo til at brukeren av systemet alltid må være i stand til å overprøve systemets beslutninger. Brukeren må altså være en bedre ekspert enn systemet selv! Dette er selvsagt absurd - all nytten av systemet vil jo være borte i en slik situasjon!»⁴⁰

Spørsmålet om hvor ansvaret skulle ligge var altså veldig komplekst og ikke minst situasjonsavhengig.⁴¹ Argumentet som Christiansen legger frem her, tar for seg dilemmaer rundt både ansvar og autonomi. Mens man hadde et ønske om at systemet skulle ta beslutninger på ekspertnivå, slik den var programmert til, var allikevel ikke tilliten stor nok til at det var trygt å la systemet handle autonomt uten tilsyn. Dette store paradokset innenfor ekspertsystemer er kjent som «human-in-the-loop», som stiller spørsmål om i hvilken grad full autonomi i slike systemer er ønskelig og mulig. Teknologistrateg Robert Mazzolin skrev en artikkel i 2020 med tittelen *Artificial Intelligence and Keeping Humans «in the Loop*, der han bl.a. utforsker spørsmål om KI-autonomi, menneskelig tilsyn og etiske implikasjoner ved å overlate beslutninger til KI-systemer.⁴² Mazzolins artikkel får frem at saken er tosidig med gode, men også motstridende poeng. På den ene siden, kan for eks. KI-styrte våpen som handler uten menneskelig innblanding, behandle konfliktsituasjoner uten emosjonelle eller psykologiske impulser som vanligvis kan påvirke beslutningsevnen. På den annen side, påpeker Mazzolin også en svakhet ved denne tilnærmingen. Selv om systemet kan handle

³⁹ Christiansen, 1985, s. 4

⁴⁰ Christiansen, 1985, s. 5

⁴¹ Christiansen, 1985, s. 5

⁴² Mazzolin, 2020

«logisk», vil den samtidig mangle evnen til risikoaversjon, noe som potensielt kan skape unødvendige konflikter, eller forlenge dem uendelig.⁴³

Et siste viktig poeng av Mazzolin som bør inkluderes er følgende;

«As AI systems are charged with making decisions with life-and-death consequences, be it in combat settings, medical facilities or simply on public roads, we are faced with the unpalatable prospect of dividing human lives into more and less valuable groups.»⁴⁴

Et potensielt autonomt system kan altså ha kapasiteten til å «filtrere» ut enkelte saker, for eks. innen medisin. En menneskelig medisinsk ekspert, kan naturligvis også prioritere enkelte pasienter fremfor andre, men det er viktig å huske på at et system ikke har den samme emosjonelle eller etiske vurderingsevnen som en menneskelig ekspert.

Ansvar var likevel ikke det eneste problemet som forelå, ifølge Christian. Et annet problem angikk brukervennlighet. For at ekspertsystemer skulle være nyttige og mulige å bruke, burde brukergrensnittet ha vært utformet på en slik måte som kunne forstås uten informatikkunnskaper. Brukergrensesnitt er en betegnelse på kontaktflaten mellom brukeren og et system, og innebærer ting som ikoner og knapper, tastevalg osv., avhengig av typen brukergrensesnitt.⁴⁵ ifølge Christiansen var brukersnittet, sammen med systemets forklaringssegenskaper den viktigste faktoren når det gjaldt utbredelse av ekspertsystem. «Forklaringsegenskaper» innebærer i dette tilfellet maskinens evne til å formidle/forklare sine beslutninger til brukeren.

Christiansens vektlegging av brukergrensesnittet er ikke urimelig, med tanke på at det dreier seg om systemer som forventes å kunne være ansvarlig for medisinske tjenester. Den konkrete bekymringen dreide seg om hvorvidt input og output skulle være brukervennlige på tvers av fagfelt. Ifølge avisartikkelen, var mange av systemene programmert i programmeringsspråket LISP. Programmeringsspråket var ikke i seg selv et problem, men om det skulle forventes LISP-kunnskaper for å bruke systemet, kunne det, ifølge Christiansen, by på en rekke utfordringer. Hvis brukerne var forventet å kunne «skrive programbiter i LISP eller lære seg notasjonene i 2. ordens predikatalkyle»⁴⁶, kunne det potensielt hindre bruk av ekspertsystemet i praksis. Isåfall, hadde det da ingenting å si om rådene du fikk av systemet var gode eller ikke.⁴⁷ Her kan det være viktig å poengtere at det Christiansen prøvde å få frem,

⁴³ Mazzolin, 2020

⁴⁴ Mazzolin, 2020

⁴⁵ Nätt & Rossen, 2023

⁴⁶ Christiansen, 1985, s. 5

⁴⁷ Christiansen, 1985, s. 5

var *hva* som var viktig for at et ekspertsystem skulle være brukbart i ulike fagområder. Christiansen snakket altså om hvordan brukersnittet *bør* være, i det tilfellet ekspertsystemene fikk stor utbredelse.

Denne artikkelen inneholder en viss mengde kritikk, og forventningene er hovedsakelig realistiske, samtidig som de på enkelte tidspunkt, fremtreder som optimistiske. Sammenlignet med det vi har sett tidligere i denne besvarelsen, blir ikke kunstig intelligens satt på pidestall i like stor grad. Det forventes at ekspertsystemene skal kunne på et tidspunkt utføre mer avanserte oppgaver, mens det samtidig påpekes hva som kan forventes innen da. Diskusjonen går altså veldig mye i sirkel, spesielt angående ekspertsystemets evne til å effektivisere medisinsk praksis. Christiansen ga typiske eksempler på bl.a. et system som kunne gi rådgivning angående behandling, men anerkjente samtidig at slike systemer kunne ta feil som oss mennesker. Som resultat, mente Christiansen at dette ville pålegge større ansvar for brukerne, men hvis ansvarsfordelingen ble for uklar, ville det medbringe en rekke etiske problemstillinger.

Det samme finner man igjen også i avisartikkelen *Medisinske ekspertsystemer* av Truls Østbye og Knut Stavem. Som nevnt tidligere, diskuterte Østbye og Stavem hvordan spesifikt Norge bør satse på medisinske ekspertsystemer. Det ble også påpekt hva som måtte til økonomisk og ressursmessig for at dette en dag skulle lykkes både innen- og utenlands. På samme måte som i Elab-rapporten fra tidligere, anerkjennes Norges svake teknologiske utgangspunkt. Følgende sitat gir en god oppsummering av budskapet til Østbye og Stavems;

«I Norge har vi sannsynligvis ikke de ressurser som kreves for å utvikle større kommersielle medisinske ekspertsystemer. Om vi ønsker å utvikle medisinske ekspertsystemer, bør vi derfor konsentrere oss om begrensede områder. Vi må også legge forholdene til rette for fremtidig bruk av slike systemer i Norge. Grunnutdanning i EDB for helsepersonell er et viktig første skritt.»⁴⁸

Holdningen mot potensiell fremtidig tilgang på ekspertsystemer i Norge var ganske optimistisk, til tross for det svake utgangspunktet. Det bør samtidig has i bakhodet at historisk sett, skyldes det mest sannsynlig ønsket om å ikke falle bak utviklingsmessig, samt ønsket om å opprettholde konkurransedyktighet. Det legges tross alt stor vekt på at det er «en forutsetning at vi tidlig satser på tilrettelegging for eksport, ellers vil aldri de høye utviklingskostnadene kunne dekkes.»⁴⁹

⁴⁸ Østbye & Stavem, 1989 s. 3000

⁴⁹ Østbye & Stavem, 1989 s. 3002

For at dette skulle være mulig, ble det stilt høye forventninger til kvaliteten på de fremtidige ekspertssystemene, særlig angående kunnskapsrepresentasjon, kunnskapsinnsamling, medisinsk beslutningsstøtte, og menneske-maskin-interaksjon. Det var, med andre ord, fokus på at medisinske ekspertsystemer skulle være i stand til å presentere informasjon forståelig, kontinuerlig oppdateres med ny informasjon uten brukerens inngrep, ha utformede regler og modeller for beslutningsstøtte, samt ha problemfri menneske-maskin-interaksjon.⁵⁰

Det som legges vekt på her, er altså mye av det samme som Christiansen legger vekt på i sine diskusjoner. Det virker som at det lå en forventning om å utvikle systemene slik at mennesket, eller kanskje heller brukerne, kunne i større grad være utenfor «løkken».

3.4 Etikk og diskusjoner og ekspertsystemer i forhold til mennesket

Så langt har dette kapittelet omhandlet utvikling av ekspertsystemer, datidens potensielle anvendelser til kunstig intelligens i helsevesenet, samt etiske problemstillinger som oppsto med det. I dette delkapittelet er målet å gå mer i dybden på hva slags etiske spørsmål rundt bruken av KI ble omtalt, og hvordan andre bruksområder, som for eks. undervisning og forsvar, ble diskutert. For dette skal det tas utgangspunkt i en artikkel i avisen *Samtiden*, skrevet av professorene Hubert og Stuart Dreyfus, med tittel «Hvorfor en datamaskin aldri vil kunne tenke som et menneske». Innholdet består av en kritisk drøfting av «den utbredte oppfatning at mennesker og maskiner tenker på samme måte – når det gjelder forretningsvirksomhet, undervisning, forsvar, medisin, jus og samfunnsspørsmål i sin alminnelighet.»⁵¹ Med tanke på at det er mye innhold, er det tilstrekkelig å konsentrere seg om eksemplene rundt undervisning og forsvar.

I likhet med tidligere artikler som har vært nevnt underveis, som for eks. «Ekspertsystemer – hva er det?» og «Glade og sinte roboter ...», forsøkes en viss grad av definering på hva som skiller mennesket fra den «tenkende» maskinen. Det skilles i dette tilfelle mellom fem ferdighetsstadier som et menneske passerer i løpet av livet: nybegynner, viderekommen, kompetent, dyktig og ekspert.⁵² For å forklare disse ulike stadiene brukes bl.a. sjakkspiller-eksemplet.

En nybegynner lærer seg fakta i forhold til bestemte ferdigheter, og handlingsmønstre som er basert på disse fakta. En sjakkspiller lærer for eksempel regelen om å «...alltid ofre brikkene dine hvis den samlede verdi av de brikkene du slår, overskrider verdien av de brikkene du taper.» Som nybegynner vet man som regel ikke at denne regelen kan brytes i visse

⁵⁰ Østbye & Stavem, 1989 s. 3001

⁵¹ Dreyfus & Dreyfus, 1987 s. 13

⁵² Dreyfus & Dreyfus, 1987 s. 13

situasjoner. Man lærer seg, med andre ord, regler og å handle ut fra dem, men ikke å bryte dem når det trengs.⁵³ De reglene man lærer som nybegynner setter en i stand til å samle erfaring. En viderekommen sjakkspiller vil, basert på tidligere erfaring, lettere unngå å bli utsatt for samme triks og angrep, men vil enda spille etter reglene. Når man nærmer seg det kompetente stadiet, vil man oppleve at regler må iblant settes til side. En mer kompetent sjakkspiller, vil kanskje ignorere de reglene de lærte som nybegynnere, og ha større tendens til å spille dristig. Kompetente spillere vil også føle større grad av ansvar for sine feil, i motsetning til en nybegynner som kanskje har større tendens til å skyld på reglene. I det dyktige stadiet, vil man hente frem hele situasjoner fra tidligere og bruke dem i aktuelle sammenhenger uten å dele det opp i regler og komponenter. En dyktig utøver handler altså ut fra intuisjon, men tenker også analytisk. Man har med andre ord, evnen til å definere et problem uten å bevisst konsentrere seg om det.

Det siste og høyeste ferdighetsnivået er ekspert. En ekspert vet som oftest hva de skal gjøre, da de har kommet til et punkt der de har samlet masse erfaring og innsikt i det de holder på med. De ferdighetene de har tilegnet seg er nå en del av dem. En sjakkspiller på ekspertnivå vil ofte ikke være bevisst på at de manipulerer brikker engang, og viker unna farer praktisk talt automatisk.

Dette er bare en kort oppsummering av de forskjellige stadiene slik de blir forklart av Dreyfus og Dreyfus. Intelligens ble altså definert som noe mer enn bare kalkulasjon og fornuftstenkning;

«Når voksne tilegner seg ferdigheter, ser man tydelig at de går *fra* en analytisk framgangsmåte der de bevisst følger bestemte regler, *til* atferd basert på ubevisst gjenkjennelse av tidligere situasjoner.»⁵⁴

Ekspertene går altså noen ganger tilbake til nybegynner eller kompetent nivået når de tenker seg frem til løsninger. På den andre siden, er allikevel også fornuftstenkning en viktig del av intelligens, da ikke alle situasjoner bør behandles på autopilot.

Videre, refererer Dreyfus og Dreyfus til to forskere, Allan Newell og Herbert Simon, som forsøkte å gå nærmere inn på datamaskinens evne til å manipulere generelle symboler. De fant ut at det var mulig å «la symboler representere elementære fakta om verden og bruke regler for å få fram forhold mellom slike fakta».⁵⁵ Newell og Simon mente at en datamaskin som er matet med fakta og regler, kunne i prinsippet løse problemer, forstå mønstre og trekke logiske

⁵³ Dreyfus & Dreyfus, 1987 s. 14

⁵⁴ Dreyfus & Dreyfus, 1987 s. 16

⁵⁵ Dreyfus & Dreyfus, 1987 s. 16

slutninger på samme måte som et menneske. I følge Dreyfus og Dreyfus, handlet dette grunnleggende sett om datamaskinens manglende evne til å skille mellom hva som var relevant og ikke, og kunne dermed *kun* brukes til spesifikke oppgaver. Newell og Simons tilnærming ble satt til side, og det ble forsøkt å få datamaskinen til å etterlikne ekte innsikt og intuisjon. Det ble allikevel slik at ingen kom så langt, og man endte opp med å møte på samme problemet som Newell og Simon gjorde.

Ifølge Dreyfus og Dreyfus, for at man skulle kunne sette likhetstegn mellom mennesket og maskinen, så måtte den forstå hva det ville si å være et menneske. Det vil si, at datamaskinen måtte kunne handle både innenfor og utenfor reglene, avhengig av situasjonen. Det var allikevel nettopp dette som var problemet – datamaskinen/programmet kunne ikke bruke sunn fornuft eller tenke fritt. Det kunne bare handle ut fra de reglene og faktaene som ble programmert i det. Sett ut fra det, er det ikke så utenkelig hvorfor de mange forsøkene på å gjenskape menneskelig intelligens ble mislykket.

«Når alt kommer til alt, har all intelligent atferd sitt utspring i vår følelse av hva vi *er*. Vi kan aldri uttrykke dette i nøyaktig form av klare regler og fakta. Det er derfor umulig å programmere en datamaskin med den type innsikt. Det er heller ikke mulig å programmere den til å mestre endringer i dagligdagse situasjoner.»⁵⁶

Dette bringer oss tilbake til det såkalte Moravecs paradokset som ble omtalt tidligere i denne besvarelsen. Dette utsagnet er inne på kanskje en av de største begrensningene i KI-forskning som rett og slett går ut på at de vanskelige problemene er enkle, mens de enkle problemene er vanskelige, slik Lim formulerte det.⁵⁷ Den typen intuisjon som rommer i den menneskelige eksperten var altså, i følge Dreyfus og Dreyfus, ikke noe som kunne programmeres. På bakgrunn av det, blir da det store spørsmålet om hvorvidt et ekspertsystem var en virkelig ekspert.

Ekspertsystemene ble lagd for å kunne utøve ekspertise innenfor et bestemt felt, og gitt ordet «ekspert» i betegnelsen, skulle man antatt at de da fungerte på samme, eller lignende, nivå som levende eksperter. I følge Dreyfus og Dreyfus var det allikevel litt mer komplisert enn dette. Det var ingen tvil om at datamaskinen presterte bedre enn et menneske i nybegynnerstadiet, men det var ikke stand til å måle seg med et menneske på ekspertnivå. Begrunnelsen som blir gitt for dette er ganske enkel. En person på ekspertnivå følger ikke noen regler, i motsetning til en maskin som er bundet til faste regler og prosesser den skal

⁵⁶ Dreyfus & Dreyfus, 1987 s. 19

⁵⁷ Lim, 2018

følge. Dette betyr ikke at menneskelige eksperter bokstavelig talt aldri følger noen regler, men heller at de for det meste handler intuitivt.

«Nybegynnere trekker slutninger ved å bruke regler og fakta akkurat som en datamaskin, men eksperten vet intuitivt hva han skal gjøre uten å følge noen regler. Ekspertene må gå tilbake til begynnerstadiet for å kunne sette opp de reglene de fremdeles husker, men ikke lenger bruker.»⁵⁸

Tar man utgangspunkt i denne logikken, kunne man konkludert med at en maskin aldri burde ha like stort ansvar som den levende eksperten, noe som blir sterkt fremmet i avisartikkelen. I følge Dreyfus og Dreyfus, kunne ekspertssystemene bli en trussel for bl.a. undervisning på skolen, men hevdet samtidig at de hadde sin plass i undervisningen. Det blir erkjent at det kunne bidra i noen undervisningsopplegg som omfatter, for eksempel, bruk av malerkost eller skrivemaskin, men at det samtidig skulle være en grense for hvor mye det skulle brukes i læring.

«Datasimulering i slike sammenhengen kommer helt sikkert til å øke i utbredelse i framtida og vil hjelpe elever i alle aldre og innen alle fagområder til å utvikle sin intuisjon. Men siden ferdigheter bare kan læres gjennom praksis, er det sunn fornuft å operere med virkelige gjenstander. Begynnerundervisning i elektrofag, for eksempel, bør foregå med batterier og lyspærer.»⁵⁹

Det som menes med sitatet ovenfor, er at ekspertssystemer eller «datasimulering» ikke skulle lære elevene, men heller fungere som en ressurs i tillegg til det vanlige undervisningsopplegget. Som argument mot de såkalte «AI-entusiastene», som mente at datamaskiner skulle spille en større rolle i undervisningen, slår Dreyfus og Dreyfus fast at for at det i det hele tatt skal være mulig, så må forskere finne ut hvordan man skal få maskinene til å snakke, lytte, vite og veilede.⁶⁰

Ekspertsystemene ble ansett som å ha operert fra nybegynner til kompetent-nivå, noe som betydde at hvis ekspertsystemene skulle erstatte lærere, hadde de produsert kun «dyktige nybegynnere» og potensielt hindre enkelte elever i å bli noe mer enn kompetente.⁶¹

Et annet eksempel på et bruksområde for ekspertssystemer i denne avisartikkelen var innenfor militæret. Helt spesifikt, gikk dette ut på utvikling av «totalt selvstyrte land-, sjø- og luftfartøy i stand til å utføre kompliserte, langtrekkende rekognoserings- og angrepstokt.»⁶² Her snakkes

⁵⁸ Dreyfus & Dreyfus, 1987 s. 23

⁵⁹ Dreyfus & Dreyfus, 1987 s. 24

⁶⁰ Dreyfus & Dreyfus, 1987 s. 25

⁶¹ Dreyfus & Dreyfus, 1987 s. 26

⁶² Dreyfus & Dreyfus, 1987 s. 27

det ikke om det norske militæret, men det er allikevel interessant å se på hvordan norske medier forholdte seg til slike ideer.

Dreyfus og Dreyfus stilte seg kritiske til bruken av slik teknologi i militæret, da den først og fremst enda ikke var i stand til å utføre de oppgavene som man ønsket;

«Dagens datamaskiner er dessuten like lite i stand til å behandle «usikre data» som de var for få år siden da vårt datastyrte raketangrepssystem tolket radarreflekser fra den oppadstigende månen som et fiendtlig angrep.»⁶³

Blant kritikken, ble det også uttrykt bekymringer for hvilke konsekvenser et fullstendig uavhengig, datastyrt forsvarssystem skulle ha i en reell kampsituasjon.

«I en kampsituasjon bruker dagens militære ledere sunn fornuft og erfaring i tillegg til relevant informasjon om det som foregår. Det skremmende ved tanken på et fullstendig uavhengig, datastyrt forsvarssystem er at ekspertenes intuisjon vil måtte vike for beslutninger tatt på et mye lavere ferdighetsnivå.»⁶⁴

Ideelt sett, kunne dette fungert om man så på det som «mennesket som samarbeider med maskinen» istedenfor «maskinen som erstatter mennesket.» Ut fra måten det formuleres i avisartikkelen, ser det allikevel ut til å være et tilfelle der man så på det mer som en truende erstatning. Her også dukker human-in-the-loop paradokset opp. Dreyfus og Dreyfus sine bekymringer ser ut til å stamme fra det faktum at systemene enda var ganske underutviklet, noe som økte risikoen for at noe kunne gå galt om man puttet for mye tillit inn i systemenes autonomi. Dette er et poeng som også Mazzolin trekker frem;

«[...] whether humans should be kept in the loop or not will depend upon how adept AI becomes at the crucial tasks of discriminating between different data sets to properly “self learn”, and noticing attempts at manipulation.»⁶⁵

I motsetning til måten ekspertsystemer i undervisning ble omtalt, stilles det mer skepsis til å la ekspertsystemer spille en betydelig rolle i militæret. I denne sammenhengen blir plutselig forventningene til KI mer pessimistiske, i den forstand at det ikke skulle medføre positive endringer.

«Med mindre de nåværende forestillingene om AI blir tilbakevist, risikerer vi en framtid der datamaskiner vil ta avgjørende militære og sivile beslutninger som helst bør overlates til mennesket.»⁶⁶

⁶³ Dreyfus & Dreyfus, 1987 s. 28

⁶⁴ Dreyfus & Dreyfus, 1987 s. 28

⁶⁵ Mazzolin, 2020

⁶⁶ Dreyfus & Dreyfus, 1987 s. 28

Det sies aldri konkret hva disse «forestillingene» innebærer, men det kan trygt tolkes at det er snakk om oppfatninger eller en tro på at systemer faktisk kunne bli så avanserte at de kunne ta avgjørelser på vegne av menneskelige eksperter. Det oppfordres på et vis til å ikke sette kunstig intelligens på en pidestall, spesielt når de bl.a. enda var så upresise og lite i stand til å behandle usikre data. Det som kanskje kan være tilfellet, er at skepsisen rettes mer mot de sistnevnte forestillingene om et selvstendig og selvstyrt system i mer avgjørende situasjoner, heller enn et system som kan til en viss grad bistå i avgjørende beslutninger.

Denne artikkelen er full av mange eksempler på hvilke forventninger som ble stilt til KI i fremtiden, og det ser ut til at pessimismen og positivismen varierer med bruksområdene. Budskapet er allikevel klart, og det er at datamaskiner kun egnet seg som et verktøy og eventuelt en liten utvidelse av menneskelig kapasitet, enn noe som kunne erstatte menneskelig beslutningstaking. Dette er, som nevnt, avhengig av bruksområdet. I undervisningssammenheng beskrives ekspertsystemet som en mulig utvidelse av menneskelig kapasitet, mens det rettes mer skepsis mot bruken av slike systemer i militære avgjørelser dersom systemet får «det siste ordet».

Som i flere av de tidligere nevnte avisartiklene i denne besvarelsen, synliggjøres et behov for å sikre at teknologien skulle tjene menneskelige interesser. Det var altså stort fokus på å ikke la utviklingen gå for mye ut av kontroll, der det i stor grad ble oppfordret til å sikre etiske retningslinjer, så vel som ansvars- og bruksprotokoler.

4.0 Konklusjon og diskusjon

Formålet med denne oppgaven har vært å gå mer i dybden på de forskjellige måtene kunstig intelligens ble omtalt på av norske medier på 1980-tallet. For å kunne finne svar på noe så omfattende som dette, ble spørsmålet delt inn i disse tre følgende problemstillingene:

«Hvordan ble endringer i norsk arbeidsliv diskutert i lys av forventningene til kunstig intelligens?»

«Hvilke etiske dilemmaer/spørsmål dukket opp i diskursen om kunstig intelligens på 1980-tallet?»

«Hvilke konkrete anvendelser av kunstig intelligens så man for seg på 1980-tallet?»

For å besvare disse problemstillingene ble det tatt utgangspunkt i en rekke tidsskrifter og sekundærlitteratur. Avisene og tidsskriftene ga et direkte innblikk i de forskjellige måtene KI ble omtalt på 1980-tallet, mens sekundærlitteraturen ble benyttet til å få innsikt i KI-historie

og ulike synspunkter rundt temaet. Fra et svært bredt utvalg av eksempler, der alle potensielt kunne svare på mine problemstillinger, ble det gjort et smalere utvalg. Det er altså utallige mange små og store eksempler som kan i ulike grad fortelle oss hvilke tanker som florerte rundt KI i denne perioden, som bl.a. debatter, tegneseriestriper og jobbbannonser, som vi ser et eksempel på i figur 2;



Figur 2: En Pusur stripe i Åndalsnes Avis i 1989 som gir en humoristisk kommentar til den økende interessen for kunstig intelligens.

Denne besvarelsen har forsøkt å vise et visst mangfoldig utvalg av kildene, men uten noe videre forskning kan ikke utvalget sies å være representativt. Under letingen etter passende kilder til problemstillingene, har jeg altså lest og sett gjennom langt flere tidsskrifter enn det som er blitt nevnt i denne besvarelsen. I en litt større forskningssammenheng, hadde det dermed vært interessant å se nærmere på disse i tillegg.

Som vi har sett, var diskusjonene om KI i 1980-tallets Norge svært mangesidig, preget av alt fra skepsis, nysgjerrighet, optimisme og frykt. I denne perioden var ekspertsystemene en dominerende form for kunstig intelligens, til tross for lite fremskritt i området. I norsk kontekst var mangelen på ressurser og kompetanse en betydelig barriere, som det ble tydeliggjort av både Kosberg, og Østbye og Stavem. Dette var imidlertid en utfordring som man var klar over og var innstilt på å overkomme, selv om det skulle vise seg at tiden ikke var helt moden for det. Det var derfor et reelt engasjement for å utvikle en teknologisk infrastruktur, som innebar bl.a. forskningssatsingsområder og KI-systemer. Dette store engasjementet var trolig drevet av en rekke årsaker, deriblant et indirekte utenlandsk press om å holde seg ajour med den raske teknologiske utviklingen og et stadig mer konkurransedyktig

⁶⁷ Åndalsnes Avis, 1989

marked. Et økende omfang av innovative systemer i bl.a. næringslivet, medførte sannsynligvis en del press på norske bedrifter og institusjoner.

Til tross for det, var det likevel mange som så på ekspertsystemene med optimisme, og betraktet dem som potensielt effektiviserende verktøy i en rekke områder. I denne besvarelsen har vi sett på eksempler på både eksisterende bruksområder, og potensielle anvendelser som ble forventet i fremtiden. Noen fagområder man så for seg skulle være effektiviserende anvendelser av ekspertsystemer var bl.a. innenfor medisinsk diagnostikk, økonomi, geologi og mekanisk konstruksjon. Andre potensielle anvendelser innebar ekspertsystemer i form av roboter/robotmennesker, som kunne erstatte sekretærer, lærere, jurister eller til og med soldater. Tanker rundt fremtidig teknologi av denne typen brakte naturligvis med seg en del bekymringer. Samtidig som mange var optimistiske til den teknologiske utviklingen, var det mange som var skeptiske til ideen av bl.a. automatiserte yrker, samt at mange var bekymret for hvilke etiske implikasjoner det ville bringe med seg. Det var lite som tydet på at norsk arbeidsliv undergikk konkrete endringer som følge av utbredelsen av kunstig intelligens. Det som heller var tilfellet, var en oppfordring til å være forberedt på at slike endringer skulle forekomme.

Under analysen og sammenligningen av kildene ble det observert noen paralleller mellom 1980-tallets og nåtidens teknologiske landskap. For dette vil jeg trekke inn én siste artikkel med tittelen *Fra Eliza til ChatGPT*⁶⁸. I denne artikkelen gjør redaktørsjef for *Tidsskrift for Den norske legeforening*, Are Brean, en historisk sammenligning av Eliza og ChatGPT, der han påpeker både potensialet, men også begrensningene til teknologien;

«Det er ingen tvil om at ChatGPT og lignende chatboter vil gi store muligheter innen klinisk medisin og forskning i årene som kommer. Men ennå er det et langt stykke igjen før Pygmalions skulptur er blitt helt levendegjort.»⁶⁹

For kontekst, refererer Pygmalion til det gamle greske sagnet om en billedhugger som forelsker seg i en kvinnestatue han selv har laget. Gudinnen Afrodite besvarte bønnene hans og statuen blir levendegjort.⁷⁰

Breans sammenligning av moderne chatboter med sagnet om Pygmalion, indikerer at teknologien har en lang vei å gå før den kan måle seg med menneskelig intelligens. I følge Brean har ikke engang dagens KI kommet så langt som å oppnå det som ble forsøkt med Eliza. Brean erkjenner potensielt gode bruksområder for chatboter som for eks. innen klinisk

⁶⁸ Brean, 2023

⁶⁹ Brean, 2023

⁷⁰ Torjussen, 2019

medisin, men påpeker samtidig at bruk av chatboter i forskningssammenheng kan innebære risiko for faktafeil og plagiat. I dag er det ikke uvanlig å finne spor av KI i forskningsartikler, men det kan innebære en risiko for svekket pålitelighet i moderne ekspertsystemer som er avhengige av ekspertinformasjon. Dersom et medisinsk ekspertsystem blir matet med feilinformasjon generert av KI, kan det føre til alvorlige konsekvenser for behandlingen og potensielt være truende for pasientens helse. En slik situasjon kan også føre til en uklar ansvarsfordeling, som Jan Christiansen diskuterte tidligere i denne besvarelsen. En slik kompleksitet kan potensielt svekke tilliten til både utviklere og brukere av systemet. Med dagens risiko for feilinformasjon grunnet KI, blir tilsyn og oppmerksomhet rundt intelligente systemer enda viktigere. Av den grunn, *kan* det tenkes å være nødvendig å forbli «in the loop».

5.0 Referanseliste

- Beck, P. (1985, 14. mai). Glade og sinte roboter. Aftenposten. Mer og mer lik mennesker, s. 11.
<https://www.nb.no/items/63bb27167ea541c974e1dd8d8f906dfc?page=9&searchText=%22glade%20og%20sinte%20roboter%22>
- Bouvier, d. S. (2018, 4. oktober). Hefaistos. Store Norske Leksikon. Hentet 8. mars, 2024, fra <https://snl.no/Hefaistos#:~:text=Hefaistos%2C%20gresk%20gud%20for%20ilden,Hefastos%27%20kult%20stammer%20fra%20Lilleasia>.
- Brean, A. (2023). Fra Eliza til ChatGPT. Tidsskrift for Den norske legeförening. <https://tidsskriftet.no/2023/04/fra-redaktoren/fra-eliza-til-chatgpt>
- Cheng, Y. (2020). Current research on Gödel's incompleteness theorems. *The Bulletin of Symbolic Logic*, 27(2), 113–167. <https://doi.org/10.1017/bsl.2020.44>
- Christiansen, J. (1985). Ekspertsystemer – hva er det? Helsetjenesten: fagtidsskrift for helsevesenet, 4(4), 3-6.
<https://www.nb.no/items/d4bcec1e7ec1e5c6d083bc5cf78e50bc?page=3&searchText=LISP>
- Dick, S. (2019). Artificial Intelligence. Harvard Data Science Review, 1(1).
<https://hdsr.mitpress.mit.edu/pub/0aytgrau/release/3>
- Dreyfus, H. & Dreyfus, S. (1987). Hvorfor en datamaskin aldri vil kunne tenke som et menneske. *Samtiden*. s.13-28.
<https://www.nb.no/items/0e4aa0e59178e119474c728294dfc59a?page=29&searchText=dreyfus>
- Duke Global Health Institute. (u.å.). Truls Ostbye (Østbye).
<https://globalhealth.duke.edu/people/ostbye-ostbye-truls>
- Gunkel, D. J. (2018). *Robot Rights*. The MIT Press.
- Kosberg, J.E. (1985). *Kunstig intelligens og ekspertsystemer: hva er det, og hva benyttes det til?* (ELAB Rapport STF44 A85024).
- Lim, M. (2018). History of AI Winters. Actuaries Institute.
<https://www.actuaries.digital/2018/09/05/history-of-ai-winters/>
- Li, T. (2022, 13. mai). *Norges forskningsråd*. Store Norske Leksikon. Hentet 15. Februar, 2024, fra https://snl.no/Norges_forskningsrad

Mazzolin, R. (2020, 23. november). Artificial Intelligence and Keeping Humans “in the Loop”. *Centre for International Governance Innovation*.

<https://www.cigionline.org/articles/artificial-intelligence-and-keeping-humans-loop/>

Natale, S. & Ballatore, A. (2020). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence*, 26(1), 3–18.

<https://journals.sagepub.com/doi/epub/10.1177/1354856517715164>

Nätt, H. T. & Rossen, E. (2023, 15. februar). brukergrensesnitt. *Store norske leksikon*.

<https://snl.no/brukergrensesnitt>

Pereira, M. L. (2020). The carousel of ethical machinery. *AI & Society*, 36(1), 185-196.

Strümke, I. (2023). *Maskiner som tenker*. Kagge Forlag AS

Torjussen, S. (2019, 4. november). Pygmalion. *Store norske leksikon*.

<https://snl.no/Pygmalion> (hentet 3. mars, 2024)

Universitetet i Oslo. (2012, 30. august). Knut Stavem.

<https://www.med.uio.no/klinmed/personer/vit/kstavem/>

Østbye, T. & Stavem, K. (1989, 20. oktober). Medisinske ekspertsystemer. *Tidsskrift for Den norske legeforening, Klinikk og forskning*, 109(29), 3000-2.

<https://www.nb.no/items/ea2b07df2048a43a2ff07eef559dc442?page=59&searchText=%22ekspertsystemer%22>

6.0. Bildekilder

Landsteiner, N. (2005). ELIZA Terminal. Mass:werk – media environments.

<https://www.masswerk.at/elizabot/eliza.html>

Åndalsnes Avis. (1989, 14. september). 64(107), 1-16.

<https://www.nb.no/items/876b3e93a00de0294a8f92298aee34e5?searchText=%22kunstig%20intelligens%22&page=0>

7.0 Litteraturliste

Universitetsbiblioteket i Oslo Bibliografisk avdeling. (1994). *Norske tidsskrifter 1971-1983 = Norwegian periodicals 1971-1983*. Biblioteket.

Weinreb, D. & Moon, D. (1981). *Lisp Machine Manual* (4 utg.).

