



Självständigt arbete (examensarbete), 15 hp, för Kandidatexamen i
Informatik
VT 2019

What's up AI?!

En undersökning kring människoliknande beteende hos chatbotar och dess påverkan på användarupplevelsen

Nina Svenningsson

Author

Nina Svenningsson

Title

What's up AI?! – A study of human behavior in chatbots and its impact on the user experience

Supervisor

Montathar Faraon

Examiner

Kerstin Ådahl

Abstract

Artificial intelligence (AI) is spreading through our society and is being used in more and more areas. Chatbots are one popular form of social AI that uses natural language to communicate with users. There are different opinions about whether a chatbot should talk and act like a human or not. One side argues that chatbots should evolve in order to simulate human behavior and intelligence. On the other hand, there are those who argue that chatbots should be transparent about being machines, and that it even can be positive for the user experience if chatbots are not too human-like. A behavior that is too human-like could also make users uneasy. This bachelor thesis aims to explore the discrepancy that has been observed in results from earlier research in order to give indications of which factors contribute to a positive user experience in chatbots, in relation to how human-like they are perceived to be. The results suggest that there's a wide range in people's preferences of how chatbots should behave. This makes it hard to name specific human-like factors that can be appealing to a large group of users. Today, if you want to design a chatbot that appeal to as many users as possible, you should give it an official/formal tone where it answers short, to the point and uses sophisticated words and well-constructed sentences. To move forward, the recommendation is to explore how chatbots could adapt their behavior to each user, in order to create a positive user experience.

Keywords

Chatbot, AI, user experience, uncanny valley, humanness, UX-design, conversation design

Författare

Nina Svenningsson

Titel

What's up AI?! – En undersökning kring människoliknande beteende i chatbotar och dess påverkan på användarupplevelsen

Handledare

Montathar Faraon

Examinator

Kerstin Ådahl

Sammanfattning

Artificiell intelligens (AI) sprider sig genom samhället och används mer och mer inom flera olika områden. Chatbotar är en populär form av social AI som använder sig av naturligt språk för att kommunicera med användare. Det finns olika åsikter kring hurvida en chatbot ska prata och bete sig människolikt eller inte. Den ena sidan argumenterar att chatbotar ska fortsätta utvecklas för att kunna simulera mänskligt beteende och intelligens, å andra sidan argumenteras att chatbotar ska vara tydliga med att de är maskiner och att det även kan vara positivt för användarupplevelsen om chatbotar inte är alltför människolika. Dessutom finns det teorier om att ett beteende som är för människolikt kan skapa obehag hos användaren. Detta examensarbete syftar till att undersöka den diskrepans som har observerats i resultaten från vetenskapliga studier för att ge indikationer på vilka faktorer som bidrar till en positiv användarupplevelse i interaktionen med chatbotar i förhållande till hur människolika de uppfattas. Resultatet visar på att det finns en stor bredd i användares preferenser för hur en chatbot bör bete sig vilket gör det svårt att nämna specifika människoliknande faktorer som är tilltalande för en större grupp användare. Om man idag vill designa en chatbot som tilltalar så många användare som möjligt bör man ge den en officiell/formell ton samt låta den svara kort, koncist och med sofistikerade ordval och välkonstruerade meningar. För att gå vidare rekommenderas bland annat att undersöka hur chatbotar kan anpassa sitt beteende efter olika användare för att skapa en positiv användarupplevelse.

Ämnesord

Chatbot, AI, användarupplevelse, uncanny valley, kusliga dalen, människolikhet, UX-design, konversationsdesign

Innehåll

1.	Inledning.....	6
1.1.	Syfte och frågeställning.....	7
1.2.	Avgränsningar	8
1.3.	Begreppsdefinitioner	8
1.4.	Disposition.....	10
2.	Litteraturgenomgång	10
2.1.	Artificiell intelligens.....	10
2.2.	Användarupplevelse	14
2.3.	Chatbotar och virtuella assistenter	15
2.3.1.	Förväntningar kring prestanda och beteende.....	17
2.3.2.	En mänsklig maskin	19
2.4.	Sammanfattning.....	21
2.4.1.	Faktorer som används i undersökningen.....	22
3.	Metod.....	23
3.1.	Litteratursökning	23
3.2.	Design av undersökning	24
3.3.	Deltagare	24
3.4.	Material	24
3.5.	Procedur.....	25
3.6.	Etiska överväganden.....	26
4.	Resultat och analys	27
4.1.	Småpratande	29
4.2.	Uppträdande/sätt.....	30
4.3.	Sofistikerade ordval och välkonstruerade meningar	32
4.4.	Dynamisk dialog.....	33
4.5.	Passande känslor för sammanhanget.....	35
4.6.	Jämförelse mellan olika grupper	36
4.7.	Designriktlinje för design av chatbotar	37
4.8.	Sammanfattning.....	38
5.	Diskussion	39

5.1.	Faktorer kopplade till människolikhet.....	39
5.1.1.	Småpratande	39
5.1.2.	Upptredande/sätt.....	41
5.1.3.	Sofistikerade ordval och välkonstruerade meningar	41
5.1.4.	Dynamisk dialog.....	42
5.1.5.	Passande känslor för sammanhanget	43
5.2.	Personlighet och ”den kusliga dalen”.....	44
5.3.	Begränsningar i studien	45
5.4.	Implikationer för UX-design	46
5.5.	Konsekvenser för individer och samhället	47
6.	Slutsatser och framtida forskning.....	48
7.	Källförteckning.....	52
8.	Bilaga 1 – Enkäten på originalspråk (engelska).....	56
9.	Bilaga 2 – Enkäten översatt till svenska.....	61

1. Inledning

Artificiell intelligens (AI) är intelligens som uppvisas av maskiner och datorer. AI har en tendens att leda tankarna till dystopiska science-fictionböcker och filmer där robotar tar över världen, och i samhället finns det en oro för hur det ska gå för mänskligheten om datorer eller robotar blir för smarta (Bostrom, 2017; Hecht, 2018; Polson & Scott, 2018, s. 16; Tegmark, 2017). AI finns dock redan runt omkring oss och används av många människor varje dag, även om alla inte är medvetna om till vilken utsträckning det sker (Hecht, 2018; McNeal & Newyear, 2013).

Sedan födelsen på 1950-talet har AI utvecklats i takt med teknologiska framsteg, och på 1990-talet startade en ny digital revolution som gjorde att tekniken gav möjlighet att utveckla mer komplexa system som exempelvis kan förstå språk och se eller uppfatta sin omgivning (Makridakis, 2017). Under de senaste åren har mycket hänt kring utvecklingen av så kallade chatbots (Akma, Hafiz, Zainal, Fairuz, & Adnan, 2018; Følstad & Brandtzaeg, 2017; Luger & Sellen, 2016; McNeal & Newyear, 2013; Zamora, 2017). En chatbot är en typ av social AI som är gjord för att prata med människor (Westerman, Cross, & Lindmark, 2018), som man interagerar med med hjälp av naturligt språk och för en dialog med i syfte att nå ett visst resultat (Dale, 2016). Dessa digitala tjänster används mer och mer och det finns förutsägelser som säger att mer än 50% av alla företag årsvis kommer lägga mer pengar på utvecklingen av chatbotar än på att utveckla mer traditionella appar redan år 2021 (Panetta, 2017).

När användare har fått artikulera vad de önskar i en chatbot har de bland annat angett att de ska vara högpresterande, smarta, sömlösa och lätta att tycka om (Zamora, 2017). Inom designfältet och forskningen finns en diskrepans mellan vilka förväntningar man har på interaktionen med och utvecklingen av chatbotar och vilken riktning som är den bästa att gå för att tillmötesgå användarnas behov.

För att utveckla chatbotar och skapa ett mer utbrett användande och mervärde för användare förespråkas på vissa håll angreppssättet att skapa chatbotar som kan simulera mänskligt beteende på ett så trovärdigt sätt som möjligt (Akma et al., 2018; Paikari & van der Hoek, 2018). Det finns dock ingen konsensus kring om det faktiskt är önskvärt att chatbotar försöker imitera mänskligt beteende (Demeure, Niewiadomski, & Pelachaud, 2011). Det är inte ovanligt att de chatbotar som existerar idag faktiskt är tydliga med att de inte är mänskliga, antingen genom att de deklarerar vad de är och hur de kan hjälpa till när de presenterar sig vid dialogens start, men också genom att påminna om att de är chatbotar när en användare ställer frågor eller chatboten inte förstår något (Thompson, 2018). Oavsett nivå av mänsklighet så anses personlighet vara en viktig byggsten för att skapa en framgångsrik chatbot som tilltalar användarna (Jain, Kumar, Kota, & Patel, 2018; Liao et al., 2018; Mone, 2016).

När personlighet används för att attrahera användare finns det däremot också risk att förväntningarna blir större än vad chatboten kan infria. Om en chatbot till exempel skämtar förväntar sig användare att den är smart för att den kan hantera något så komplext som humor (Luger & Sellen, 2016). När chatboten inte lever upp till förväntningarna blir användare snabbt besvikna och frustrerade (Luger & Sellen, 2016; McDuff &

Czerwinski, 2018). Det kan också finnas risker att användare tycker att det är obehagligt om chatboten försöker vara alltför människolik men sedan inte klarar av att bemöta de förväntningar som detta skapar (Luger & Sellen, 2016; McDuff & Czerwinski, 2018; Mone, 2016; Strait, Vujovic, Floerke, Scheutz, & Urry, 2015; Waddell, 2017). Chatboten utgår från en typ av interaktion som är väldigt lik konversationer med andra människor, och när användare kan interagera med en chatbot på samma villkor som en vanlig konversation men responsen inte liknar den som de hade förväntat sig av en dialog med en annan människa, så skapas en dissonans. Chatbotens sanna, robotiska natur blir plötsligt väldigt tydlig vilket i sammanhanget blir något av en chock (Waddell, 2017).

Användare upplever inte idag att deras förväntningar på chatbotars kompetens blir infriade (Jain et al., 2018; Luger & Sellen, 2016). När de inte känner sig säkra på vad chatboten faktiskt kan göra blir förväntningarna otydliga och även detta påverkar användarupplevelsen (Zamora, 2017). Luger och Sellen (2016) förespråkar att det ska göras tydligare i design och utformning vilken sorts intelligens eller kompetens en chatbot faktiskt har i syfte att ge användaren mer rimliga förväntningar och i förlängningen en bättre användarupplevelse. En typisk användare har idag svårt att få en bra bild över hur chatboten fungerar och de får inte heller tillräcklig med feedback för att kunna avgöra detta (Jain et al., 2018; Luger & Sellen, 2016).

I forskningen idag finns ingen tydlig gemensam uppfattning om huruvida faktorer som får en chatbot att verka människolik är positivt eller negativt. Det existerar också studier som säger emot varandra. Till exempel argumenterar Westerman, Cross & Lindmark (2018) för att en chatbot eventuellt kan bli mer trovärdig om den är mer robotisk än människolik och även Mone (2016) har sett att en maskin inte nödvändigtvis behöver vara människolik för att vara engagerande. Det finns å andra sidan studier som visar att användare attraheras av människoliknande chatbotar (Akma et al., 2018) och dialoger med dessa där de kan använda sig av ett språk som möjliggör människolik konversation (Jain et al., 2018). Piccolo, Mensio & Alani (2018) har däremot sett en koppling mellan frustration hos användaren och ett för stort mått av människolikhet. Det finns inte heller nödvändigtvis en relation mellan hur människolik och emotionellt intelligent ett system är och hur omtyckt det är (McDuff & Czerwinski, 2018). Neururer, Schlögl, Brinkschulte & Groth (2018) argumenterar att det kan vara avhängigt på kontext om en chatbot ska härma mänskligt beteende eller utveckla en egen personlighet.

Mot bakgrunden av föregående diskrepans och resonemang syftar detta examensarbete till att identifiera relevanta faktorer i upplevelsen av en chatbots beteende och personlighet och undersöka dessa utifrån ett användarperspektiv för att utläsa mönster som kan bistå i skapandet av designriktlinjer för chatbotar. Ett speciellt fokus kommer ligga på de faktorer som gör att chatbotens beteende upplevs människolikt och om dessa faktorer är önskvärda för att uppnå en positiv användarupplevelse.

1.1. Syfte och frågeställning

Syftet med detta examensarbete är att undersöka och analysera diskussionen inom forskningen och den diskrepans som observerats i resultaten från den vetenskapliga litteraturen, med fokus på om det är önskvärt med mänskligt beteende i chatbotar. Detta

för att ge indikationer på vilka faktorer som bidrar till en positiv användarupplevelse i chatbotar i förhållande till hur människolika de uppfattas, och formulera en designriktlinje som kan stötta framtida arbete med design av chatbotar. Utifrån litteraturen kommer faktorer att identifieras och sammanställas för att agera som underlag i en enkät som undersöker användares uppfattning om och inställning till dessa faktorer när de demonstreras i en chatbot. Svaren ska analyseras för att identifiera mönster som i sin tur kan utgöra grunden för att föreslå designriktlinjer som kan guida designers vid framtida design av chatbotar med fokus på en positiv användarupplevelse.

Frageställningen lyder som följer; *Vilka faktorer som gör att chatbotar upplevs som människolika kan användas för att skapa en positiv användarupplevelse mellan användare och chatbot?*

Användarupplevelse definieras här som den upplevelse en användare har av att interagera med en chatbot, hur interaktionen påverkar deras sinnesstämning och på vilket sätt de kan nå ett önskat resultat genom interaktionen. En *chatbot* är en virtuell personlig assistent som interagerar med en användare genom naturlig språkinlärning (eng. Natural Language Processing, eller NLP), det kan gälla både textbaserade och röststyrda agenter (Dale, 2016).

1.2. Avgränsningar

I studien har en avgränsning gjorts till att endast undersöka situationer där användaren är fullt medveten om att hen interagerar med en chatbot. Det har till exempel inte gjorts en jämförelse mellan samtal där användaren pratar med en människa och sedan en liknande konversation där de interagerar med en chatbot. Undersökningen och frågorna har designats för att göra det tydligt att deltagarna ska reagera utifrån hur det skulle kännas att få de olika svaren från just en chatbot. Fokus har också legat på studier som undersöker chatbotar som använder sig av Natural Language Processing (NLP) och alltså interagerar med användaren via språk (text eller röst) och inte via digitala gränssnitt där användaren klickar sig fram genom att trycka på färdigartikulerade svarsalternativ eller länkar (ofta i form av knappar) som skapas av chatboten.

När det gäller de faktorer som undersöks i enkäten har en avgränsning gjorts till de faktorer som i litteraturgenomgången har setts bidra till hur människolik en chatbot är. Utifrån dessa faktorer har sedan faktorer valts som på ett rättvist sätt kan undersökas i en enkät.

1.3. Begreppsdefinitioner

- **Användarupplevelse** används för att beskriva upplevelsen en användare har av att interagera med en chatbot, hur interaktionen påverkar deras sinnesstämning och på vilket sätt de kan nå ett önskat resultat genom interaktionen.
- **Artificiell intelligens (AI)** är intelligens när den uppvisas av maskiner och datorer (Gams, Gu, Härmä, Muñoz, & Tam, 2019), det kan också vara mjukvaruprogram som resonerar och simulerar tänkande (Govers, 2018, s. 8).
- **Chatbot** används för att beskriva en virtuell personlig assistent som interagerar med en användare genom naturlig språkinlärning (eng. Natural Language

Processing, eller NLP), det kan gälla både textbaserade och röststyrda agenter (Dale, 2016). Det är inte ovanligt att en chatbot använder sig av ett samtals- eller dialogsystem som liknar det som existerar i samtals- och meddelandeappar (Dale, 2016; Følstad & Brandtzaeg, 2017).

- **Dynamisk dialog** är ett begrepp som i den här uppsatsen innebär en dialog där båda parter har kompetensen och möjligheten att ställa relevanta följdfrågor och förstå och hålla kontext genom konversationen.
- **Faktorer** används i uppsatsen som ett begrepp för att benämna de olika egenskaper eller tekniska aspekter som kan sammankopplas med upplevelsen av en chatbot. Till exempel kan personlighetsdrag som “familjär”, “människoliknande” eller “humoristisk” inkluderas som faktorer men också mer tekniska/praktiska förväntningar som “effektivitet”, “snabbhet” och “relevanta svar”.
- **Interaktion** definieras i detta examensarbete som den aktivitet där en användare kommunicerar med ett digitalt medium, till exempel en chatbot, för att uppnå ett visst resultat. Interaktionen består då till stor del av dialog via språk men kan också innebära att användaren använder sig av ett grafiskt gränssnitt, till exempel att trycka på knappar, för att få en uppgift utförd.
- **Konversationsdesign** är ett fält inom design som syftar till att undersöka och designa interaktion som sker via språkliga konversationer, detta kan även inkludera sociala interaktioner som sker mellan användare och deras omgivning (Nishida, 2015).
- **Kusliga dalen** (eng. the uncanny valley), är en teori av Mori (1970, 2012) som beskriver ett obehag som skapas när människor interagerar med människoliknande robotar. Till en viss gräns upplevs människolikhet i robotar engagerande och attraktivt, men om det blir för stora likheter så kan detta istället frambringa obehag eller olustkänslor hos användaren.
- **Natural Language Processing** (NLP), eller *naturlig språkinläring*, är en samling tekniker som möjliggör för analysering av text eller språk på en eller flera nivåer (Schuetzler, Grimes, Giboney, & Buckman, 2014), och som hjälper människor att interagera med datorer. Komplexiteten kan gå från sofistikerade metoder för att förstå vad människan skriver eller pratar om, till att söka efter ämnesord och matcha med passande svar (Schuetzler et al., 2014).
- **Personlighet** i samband med chatbotar refererar i detta examensarbete till, om inget annat nämns, en *människolik* personlighet hos chatboten.
- **UX-design** (eng. user experience design) är ett fält inom design där fokus ligger på att skapa en behaglig och positiv användarupplevelse, som håller en hög kvalitet, för den som använder en viss tjänst eller produkt (Norman, 2013, s. 5).
- **UX-designer** (eng. user experience designer) är en designer som jobbar med användarcentrerad design med fokus på användarens upplevelse.
- **Virtuell assistent** är ett digitalt hjälpmedel som användare interagerar med i syfte att få hjälp att lösa olika uppgifter. Assistenterna kan vara röstbaserade, textbaserade eller multimodalt förkroppsligade (Jain et al., 2018).

1.4. Disposition

I nästa sektion presenteras litteraturgenomgången. Där kommer ämnesområdet diskuteras utifrån ett bredare perspektiv med start inom fältet för artificiell intelligens för att sedan smalna av mot chatbotar och användares förväntningar på chatbotar.

Därefter följer metodavsnittet som beskriver hur insamlingen av empiri har gått till, med utformningen av en enkät och de etiska överväganden som har tagits under övervägande i datainsamlingen.

I resultat- och analysavsnittet sorteras och analyseras de svar som kommit in via enkäten i syfte att hitta mönster som kan visa på vilka faktorer och funktioner som användare upplever som positiva när de använder chatbotar. Insikterna sätts i relation till de faktorer som i litteraturgenomgången har identifierats som faktorer som påverkar hur människolik en chatbots beteende upplevs. Utifrån detta artikuleras en designriktlinje som kan guida designers inför framtida design av chatbotar.

I diskussionsavsnittet avhandlas hur resultaten kan relateras till en eventuell existens av "den kusliga dalen" hos chatbotar och vilka implikationer detta kan skapa för designers vid design av chatbotar. Resultaten från undersökning relateras också till studierna från litteraturgenomgången.

Avsnittet för slutsatser och framtida forskning sammanfattar de slutsatser som har dragits utifrån resultat och analys, och artikuleras områden som behöver fortsätta utforskas för att skapa en bättre förståelse för hur chatbotar kan designas och hur de bör bete sig.

2. Litteraturgenomgång

2.1. Artificiell intelligens

Artificiell intelligens (AI) är intelligens som uppvisas av maskiner och datorer (Gams et al., 2019), det kan också vara mjukvaruprogram som resonerar och som simulerar tänkande (Govers, 2018, s. 8). I många fall är AI uppbyggt efter försök att efterlikna mänsklig intelligens (Gams et al., 2019) och om man ska förstå vad AI innebär behöver man också skapa sig en förståelse för vad denna intelligens innebär. Det finns många olika försök till att definiera intelligens, men ingen konsensus (Tegmark, 2017, s. 65). För att försöka skapa en bild av hur mänsklig intelligens ser ut kan man dela upp intelligens i sex olika områden: *abstrakt intelligens*, *praktisk intelligens*, *emotionell intelligens*, *estetisk intelligens*, *kinestetisk intelligens* och *social intelligens* (Neururer et al., 2018). Tillsammans kan dessa områden bygga en mer komplett bild av hur mänsklig intelligens kan vara uppbyggd och en kompetent AI behöver kunna hantera dessa olika typer av intelligens (Neururer et al., 2018).

AI är inte ett nytt område, redan på 1950-talet myntades begreppet av John McCarthy på Dartmouthkonferensen och samma år tionde publicerade Alan Turing (1950) en artikel som av många forskare anses vara födelsen av AI (Gams et al., 2019). Turing (1950) undersökte möjligheterna att utveckla datorer som kunde tänka och visa intelligent beteende. För att analysera och svara på frågan om datorer kan tänka skapade han ett test,

Turingtestet, där en dator och en människa blir utfrågade av en jury som försöker utröna vem av dem som är en dator. Om juryn inte lyckas avgöra vem som är vem så skulle det innebära att datorn lyckas med testet och eventuellt kan anses kunna tänka. Många forskare och utvecklare har försökt skapa datorer som klarar av Turingtestet, men det var inte förrän 2014 som det första lyckade försöket skedde (Warwick & Shah, 2016). Det är svårt för artificiell intelligens att imitera mänsklig konversation och mänskligt beteende då det finns så många olika aspekter att ta hänsyn till (Vinciarelli et al., 2015; Warwick & Shah, 2016). En utmaning i dialoger mellan människa och maskin är att mänskliga konversationer bygger mycket på icke-verbal kommunikation, utöver själva språket, och detta är svårt för en robot att plocka upp (Schuetzler et al., 2014). Mänskligt beteende bygger på kognitiva, sociala och psykologiska faktorer som dessutom är beroende av kontext och skiljer sig mellan olika situationer (Vinciarelli et al., 2015). De kognitiva kompetenser det krävs för att imitera en människa är avancerade och inbegriper både sociala färdigheter, intelligens och språk (Lortie & Guitton, 2011).

För att en dator ska klara av Turingtestet behöver den övertyga en tredjedel av domarna att den är en människa. Chatboten som var den första maskinen att lyckas lura domarna kallades Eugene Goostman och utgav sig för att vara en ung pojke från Ukraina, vilket skulle kunna förklara eventuella stavfel eller varför han eventuellt inte skulle kunna svara på alla frågor (Shah, Warwick, Vallverdú, & Wu, 2016; Warwick & Shah, 2016). Det finns tvivel på om testet faktiskt kan mäta hur intelligent en maskin är, eller om det snarare handlar om att en programmerare försöker övertala en annan människa om att de pratar med en människa och inte en maskin (Braga & Logan, 2017). Trots att det nu finns maskiner och chatbotar som har lyckats klara Turingtestet betyder inte det att de är högtintelligenta. Turing själv artikulerade att testet handlar om att imitera en människa, dessutom krävs endast att de klarar av att hålla upp fasaden i fem minuter vilket är en relativt kort tid, vilket gör att det är större chans att de lyckas klara av att hålla upp imitationen under testets gång (Warwick & Shah, 2016).

Tidiga försök att utveckla AI försvårades av bland annat bristfällig teknik, dåliga metoder och begränsningar i den existerande hårdvaran (Bostrom, 2017, ss. 20-28; Govers, 2018, s. 10). I början av 1990-talet startade dock en ny era genom, bland annat, introduktionen av mikroprocessorer som kunde inkorporera grafik och musik vilket öppnade upp helt nya användningsområden och möjligheter för att skapa applikationer som klarar av mer komplexa uppgifter som exempelvis att förstå språk och kunna se eller uppfatta sin omgivning (Makridakis, 2017).

Under utvecklingen av AI har fyra typer av artificiella intelligenser formulerats, nämligen; *mekanisk intelligens*, *analytisk intelligens*, *intuitiv intelligens* och *empatisk intelligens* (Huang & Rust, 2017). *Mekanisk intelligens* bygger på en förmåga att automatiskt kunna genomföra uppgifter, ofta repetitiva, och anpassar sig inte eller lära sig nya saker. Man kan se exempel på detta i till exempel serviceyrken, hos kundtjänst och taxiförare, och AI har idag bland annat tagit över uppgifter byggda på mekanisk intelligens i självbeställning via skärmar i snabbmatsrestauranger. I *analytisk intelligens* finns förmågan att analysera och processa information och på detta sätt lära och anpassa efter nya situationer, vilket passar bra i komplexa, konsekventa och förutsägbara

sammanhang. Maskininlärning är en typ av *analytisk intelligens* och Toyota har till exempel utvecklat en smart funktion i sina bilar som kan diagnostisera fel och underlätta för tekniker. AI med *intuitiv intelligens* är skapad för att vara mer flexibel. Det krävs en annan typ av förståelse än enbart analysering av data, som kreativitet och en holistisk förståelse för att ta beslut intuitivt och att lösa problem. Yrkesgrupper som läkare och advokater har arbetsuppgifter som kräver intuitiv intelligens. *Empatisk intelligens* innefattar att kunna avläsa och reagera på andra människors känslor, och existerar i nuläget bara i de mest avancerade AI. Ett exempel på detta är chatbotar som kan interagera med människor och lära sig från erfarenheterna. Det pågår dock en debatt kring om maskiner kan vara emotionella och känna empati, och å andra sidan om detta faktiskt spelar någon roll så länge de klarar av att agera som om de kan det (Huang & Rust, 2017).

I samhället finns en oro över vad som ska hända om eller när AI slår igenom, men den finns redan bland oss och många använder AI dagligen utan att reflektera över det (Hecht, 2018; McNeal & Newyear, 2013). Det kan handla om att fråga Apples röstassistent Siri om vägen eller självkörande bilar, men också mer dolda funktioner som när spam-mail rensas ut till skräpposten, vid internetbetalningar när kortinformationen analyseras för att upptäcka potentiella bedrägerier eller när användare får rekommendationer kring vilken film de kan tänkas vilja se närmast på Netflix (Gams et al., 2019; Hecht, 2018).

Ett av målen med AI är att ersätta, komplettera och förstärka i princip alla uppgifter som idag genomförs av människor, vilket om det faktiskt genomförs kommer att innebära en dramatisk förändring för människor och samhället (Makridakis, 2017). Frågor kring hur människans roll kommer att se ut i framtiden om eller när datorer kan prestera lika bra eller bättre än människor, och dessutom till ett billigare pris, skapar oro i vissa läger men det finns andra som är positiva och välkomnar utvecklingen (Makridakis, 2017). Makridakis (2017) delar in dessa åsikter i fyra olika grupper: *optimisten*, *pessimisten*, *pragmatikern* och *tvivlaren*.

Optimisten tror att AI kan revolutionera samhället och skapa en utopi där människan kan utrota sjukdomar och lägga mindre tid på arbete och mer tid på vad de vill göra. *Pessimisten* ser en framtid där människorna låter datorer ta över fler och fler beslut, och allteftersom de blir mer komplexa tar datorerna också över mer kontroll från människorna. I *pessimistens* framtid riskerar människan att bli sekundär till datorerna. *Pragmatikern* tror att AI kan hjälpa människan att utvecklas och att vi kan ta tillvara på kraften hos datorer för att förstärka våra egna kunskaper. På detta sätt ligger människan steget före och om det skulle bli riskabelt tror *pragmatikern* att vi kommer kunna stänga ner farliga datorer, alltså utgör AI ingen överhängande risk. *Tvivlaren* tror inte att AI är ett hot då hen inte anser att superintelligent AI är en möjlighet. Superintelligens innebär att en maskins intelligens överträffar mänsklig intelligens (Bostrom, 2017, s. 43; Gams et al., 2019). Dock tror inte *tvivlaren* att superintelligent AI är något att oroa sig för då det mänskliga medvetandet inte går att sammanfatta i regler och datorer kan inte vara kreativa eller innovativa då de inte kan bryta regler. Alltså tror inte *tvivlaren* att datorer kan uppnå mänsklig intelligens, och ännu mindre superintelligens.

Tegmark (2017, ss. 40-48) har gjort en liknande indelning men i tre kategorier: *digitala utopister*, *teknikskeptiker* och *medlemmar av rörelsen för välvillig AI*. De *digitala utopisterna* är inte oroliga för utvecklingen av AI utan anser att det är ett naturligt kommande steg i människans utveckling, och att utkomsten kommer bli positiv för människan. *Teknikskeptikerna* är inte heller oroliga, men av den anledningen att de inte tror att det är möjligt att skapa AI med en generell intelligens som överträffar människans inom en överskådlig framtid. Det ligger alltför långt borta för att det ska vara något att oroa sig över. *Rörelsen för välvillig AI* ser att det kan finnas säkerhetsrisker och komplexa frågor som behöver besvaras om vi ska kunna ha en positiv framtid med övermänskligt intelligent AI och argumenterar för att riskerna behöver diskuteras och forskning behöver föras kontinuerligt för att det ska finnas en chans att människan kan hantera dessa komplexa frågor när det väl är dags. Enligt *rörelsen för välvillig AI* behöver vi agera nu för att ha en chans till en framtid med välvillig AI.

För att närma oss högre intelligent AI så har också de sätt vi kommunicerar med datorer utvecklats. Natural Language Processing (NLP), eller naturlig språkinlärning, är en samling tekniker som möjliggör för analysering av text eller språk på en eller flera nivåer (Schuetzler et al., 2014). Genom att arbeta med NLP kan datorer interagera med människor (AbuShawar & Atwell, 2016; Polson & Scott, 2018). I vissa fall används sofistikerade metoder för att förstå vad människan skriver eller pratar om, i andra fall letar systemet istället efter nyckelord och matchar dessa med de svar som den har lärt sig för att kunna simulera att den förstår (Schuetzler et al., 2014).

Chatbotar är ett exempel på en AI som använder sig av NLP för att kunna kommunicera och interagera med människor. En chatbot är en typ av social AI som är gjord för att prata med människor (Westerman et al., 2018), som man interagerar med med hjälp av naturligt språk, och för en dialog med i syfte att nå ett visst resultat (Dale, 2016). Chatbotar utnyttjar samma typ av dialogmönster som de som återfinns i samtals- och meddelandeappar och de blir mer och mer populära (Dale, 2016; Følstad & Brandtzaeg, 2017).

Sammanfattningsvis så krävs det hög intelligens och en kapacitet att hantera flera olika sociala färdigheter på ett följsamt sätt genom olika situationer och i olika kontext för att en AI ska kunna simulera mänsklig intelligens (Lortie & Guitton, 2011; Vinciarelli et al., 2015; Warwick & Shah, 2016). Det är dock svårt att bygga maskiner och system som kan simulera ett mänskligt beteende (Lortie & Guitton, 2011; Schuetzler et al., 2014; Vinciarelli et al., 2015; Warwick & Shah, 2016) och det finns motsägande åsikter kring om AI kan bli lika intelligent som en människa eller inte (Makridakis, 2017; Tegmark, 2017). Via Turingtestet testas om maskiner faktiskt kan lura en jury på ett sådant sätt att de inte vet om det är en människa eller maskin som de talar med (Turing, 1950). Dessa resultat är inte nödvändigtvis en indikation på att det existerar intelligenta maskiner, då testet snarare mäter hur väl ett system eller en dator kan simulera intelligens under en kort tid (Braga & Logan, 2017). Samtidigt visar det att utvecklingen går framåt då det idag finns maskiner som har genomfört Turingtestet med ett resultat där tillräckligt många i juryn inte är säkra på om de är en människa eller maskin (Warwick & Shah, 2016). Med NLP har sättet användare kan interagera med maskiner utvecklats (Schuetzler et al., 2014)

och chatbotar blir mer och mer populära (Dale, 2016). Utvecklingen av AI är ett område som oroar många människor och det finns en rädsla för vad som ska hända om eller när AI slår igenom “på riktigt” (Hecht, 2018; Polson & Scott, 2018, s. 16). Detta är något som möjligen kan påverka användarupplevelsen och interaktionen med dessa system.

2.2. Användarupplevelse

Inom UX-design (eng. user experience design) ligger fokus på att skapa en behaglig och positiv användarupplevelse som håller en hög kvalitet för den som använder en viss tjänst eller produkt (Norman, 2013, s. 5). Don Norman, professor och designexpert, beskriver även hur interaktionsdesign behandlar hur användare interagerar med teknik, med fokus på att förbättra användarens förståelse över vad som kan göras, vad som händer och vad som precis hänt (Norman, 2013, s. 5).

Lynch och Horton (2016, s. 10) anger i sin bok om webb- och UX-design följande faktorer som viktiga för att skapa en lyckad UX-design (i exemplet rör det sig om designen av en hemsida):

1. *Lärandevänlig* (eng. learnability) - Hur lätt är det för en ny användare att lära sig hur hemsidan fungerar?
2. *Orienteringsvänlig* (eng. ease of orientation) - Hur lätt är det för användaren att orientera sig på hemsidan?
3. *Effektivitet* (eng. efficiency) - Hur snabbt kan användaren nå sitt mål?
4. *Minnesvänlig* (eng. memorability) - Kan en återvändande användare snabbt komma in i hur hemsidan fungerar?
5. *Tillgänglighet* (eng. accessibility) - Kan användare med fysiska eller sensoriska begränsningar tillgodogöra sig hemsidan på ett effektivt sätt?
6. *Misstagsförlåtande* (eng. error forgiveness) - Är hemsidan förlåtande vid vanliga användarmisstag, och gör användare många misstag på hemsidan?
7. *Glädje* (eng. delight) - är det trevligt för användaren att använda hemsidan?

Framförallt är de fyra faktorerna *lärandevänlig*, *orienteringsvänlig*, *effektivitet* och *tillgänglighet* viktiga att ha i åtanke under hela designprocessen (Lynch & Horton, 2016).

Norman (2013, s. 38) beskriver det som han kallar för “klyftan till utförande” (eng. “gulf of execution”), när användaren behöver skapa sig en förståelse för hur en produkt kan användas och “klyftan till utvärdering” (eng. “gulf of evaluation”) där användaren ska lista ut vilken status artefakten har i stunden och vad som hände efter att hen interagerat med den. Som designer finns det ett antal designprinciper man kan följa för att hjälpa användaren till bättre förståelse och användarupplevelse (Norman, 2013, s. 72). Även om dessa principer kanske är gjorda för objekt med fysiska attribut eller i alla fall med gränssnitt, finns det några faktorer som kan vara viktiga att ta med sig till designen av virtuella personliga assistenter och chatbotar. Den första är *discoverability* och innebär att det ska vara möjligt för användaren att utläsa vad hen kan göra med en produkt/ett objekt. Den andra är *feedback*, det ska vara tydligt vilka konsekvenser en aktion får och användaren ska kunna utläsa i vilket tillstånd produkten eller objektet hen interagerar med är. Den tredje, *conceptual model*, innebär att användaren ska kunna skapa sig en bild av

hur produkten eller objektet fungerar och förstå vad som kommer hända när hen genomför en specifik aktion. Dessa faktorer har visat sig vara problematiska att förmedla till användare som interagerar med chatbotar, de har svårt att skapa sig en korrekt bild över hur chatboten fungerar och de får inte tillräcklig feedback för att kunna avläsa vad chatboten är kapabel till (Jain et al., 2018; Luger & Sellen, 2016).

En annan aspekt av användarupplevelsen som debatteras i förhållande till AI är hur människoliknande de är och hur detta påverkar användarupplevelsen. Till exempel argumenterar Westerman et al. (2018) för att en chatbot eventuellt kan bli mer trovärdig om den är mer robotisk än människolik och även Mone (2016) har sett att en maskin inte nödvändigtvis behöver vara människolik för att vara engagerande. Det finns å andra sidan studier som visar att användare attraheras av människoliknande chatbotar (Akma et al., 2018) och dialoger med dessa där de kan använda sig av ett språk som möjliggör människolik konversation (Jain et al., 2018). Piccolo et al. (2018) har däremot sett en koppling mellan frustration hos användaren och en stor människolikhet. Det finns inte heller nödvändigtvis en relation mellan hur människolik och emotionellt intelligent ett system är och hur omtyckt det är (McDuff & Czerwinski, 2018). Neururer et al. (2018) argumenterar att det kan vara avhängigt på kontext om en chatbot ska härma mänskligt beteende eller utveckla en egen personlighet. En människolik chatbot kan påverka användaren till att interagera mer som i en konversation med andra människor, samma pauser mellan meningar och om man stannar upp för att tänka etc. Om chatboten istället är för statisk känner användarna inte samma behov av att hålla uppe en naturlig konversation, vilket till exempel gör det lättare att ta extra tid för att formulera och skapa trovärdiga svar i de fall där användaren försöker att ljuga av någon anledning, eller framställa sig själv på ett inte helt autentiskt sätt (Schuetzler et al., 2014).

Teorin om "den kusliga dalen" (eng. the uncanny valley) bygger på att det till en viss gräns är tilldragande och engagerande ju mer människoliknande AI blir, men om det går över en gräns och blir för människolikt så blir det istället obehagligt (Mori, 1970, 2012). Detta är ett fenomen som inte bara har observerats inom området för fysiska robotar som bygger på AI, utan tendenser till samma teori har även diskuterats och analyserats i förhållande till virtuella assistenter i form av chatbotar (Candello, Pinhanez, & Figueiredo, 2017; McDuff & Czerwinski, 2018; Mone, 2016; Skjuve, Haugstveit, Følstad, & Brandtzaeg, 2019). För att förstå på vilka sätt en chatbot kan upplevas som obehaglig krävs en större förståelse för hur chatbotar fungerar generellt, och vilka förväntningar det finns på dem från användarens håll. Detta behandlas i nästa avsnitt.

2.3. Chatbotar och virtuella assistenter

Det finns olika typer av virtuella assistenter och dessa kan delas in i (1) *röstbaserade* - som till exempel Apples Siri, Google Assistant, Amazons Alexa, (2) *textbaserade* - där man interagerar via skrivna meddelanden och (3) *multimodal förkroppsligad assistent* - som har ett visuellt grafiskt gränssnitt och kommunicerar även icke-verbalt genom att använda sig av till exempel gester (Jain et al., 2018). En chatbot är en typ av social AI som är gjord för att prata med människor (Westerman et al., 2018). En chatbot pratar man med med hjälp av naturligt språk, och för en dialog med i syfte att nå ett visst resultat

(Dale, 2016). Flexibiliteten detta skapar gör att chatbotar används inom många olika områden idag (Akma et al., 2018; Skjuve et al., 2019) såsom *marknadsföring* (Cancel, 2019), exempelvis SproutSocial (sproutsocial.com) och Dialogtrail (dialogtrail.com), *kundservice* (Chakrabarti & Luger, 2015), *bibliotek* (McNeal & Newyear, 2013) och *sjukvård* (Pereira & Díaz, 2019), exempelvis MedWhat (medwhat.com) och Endurance¹.

De senaste åren har utvecklingen av chatbotar nått stora framgångar och de används mer och mer (Akma et al., 2018; Følstad & Brandtzaeg, 2017; Luger & Sellen, 2016; Zamora, 2017), men det är ingen ny teknik i grunden. Redan 1966 programmerades "Eliza" till att simulera samtal som en terapeut och systemet var så övertygande att flera av deltagarna i studierna trodde att de pratade med en riktig människa (McNeal & Newyear, 2013; Weizenbaum, 1966). Framgångarna med Eliza inspirerade andra till att försöka skapa mer människoliknande chatbotar som kunde klara Turingtestet (McNeal & Newyear, 2013). Dock har det dröjt till senare år för utvecklingen av chatbotar att ta fart, men idag finns en ny tilltro till att chatbotar har kommit tillbaka för att stanna (Dale, 2016). Detta kan möjligtvis bero på framgången som samtals- och meddelandeappar har fått (Dale, 2016; Følstad & Brandtzaeg, 2017), som idag används av två miljarder människor (Statista, 2019). Från och med 2014 började fler företag och utvecklare utnyttja möjligheten att skapa chatbotar inuti dessa applikationer, vilket har ökat tillgängligheten och sänkt tröskeln för användare till att faktiskt komma åt och använda chatbotar (Klopfenstein, Delpriori, Malatini, & Bogliolo, 2017; Zamora, 2017).

Chatbotar kan ha olika nivåer av intelligens och interagera på olika sätt. Paikari och van der Hoek (2018) föreslår en klassificering av interaktionsstilar på chatbotar där de delar in dessa inom fyra olika områden: (1) *enformig* (eng. dull) - chatboten svarar genom att använda enstaka ord och upprepar ofta samma eller liknande fraser, (2) *alternativt ordförråd* (eng. alternate vocabulary) - chatboten har en större variation av svar, de säger samma sak men det blir större variation för användaren, (3) *relationsbyggare* (eng. relationship builder) - ordförråd och samtalsämnen går mellan att vara lättsamma och professionella, chatboten kan designa konversationsflöde och försöker skapa ett mer intressant flöde genom att spontant ge information eller underhålla med ett skämt och (4) *människolik* (eng. human-like) - där chatboten lär sig av tidigare erfarenheter och konversationer och kan på så sätt interagera med mer subtila konversationsmönster med högre grad av meningsfulla dialoger.

I samma studie artikuleras även tre användningskategorier för chatbotar med olika ändamål (Paikari & van der Hoek, 2018):

1. *Information* - som hjälper användaren att hitta relevant information för olika situationer;
2. *Samarbete* - med mål att hjälpa användare att kommunicera och samarbeta genom att stödja interaktion och kommunikation under arbetets gång och;

¹endurancerobots.com/azbnmaterial/a-robot-companion-for-senior-people-and-patients-with-alzheimer-s-disease/

3. *Automatisering* - som hjälper till att automatisera delar av arbetet med återkommande uppgifter eller ändringar som behöver göras mellan olika artefakter.

Trots att chatbotar fortsätter att utvecklas och används mer och mer så finns det inte en tydlig gemensam bild inom vetenskapen kring vilket faktiskt ändamål och värde de bidrar med för användare (Zamora, 2017).

2.3.1. Förväntningar kring prestanda och beteende

Det är svårt att förutsäga hur människor reagerar på interaktionen med chatbotar då detta beror mycket på personlighet (Lortie & Guitton, 2011; Mone, 2016). Dock finns det ett antal studier som har undersökt vilka faktorer som påverkar interaktionen och användarupplevelsen mellan människa och chatbot. I följande sektion sammanfattas faktorer som är representerade i litteraturen.

Chakrabarti och Luger (2015) utvärderade artificiella konversationer med hjälp av Grices samtalsmaximer, som kan delas in i de fyra kategorierna *kvantitet*, *kvalitet*, *relevans* och *uppförande*. I undersökningen identifierades ett antal faktorer som bidrog till att en analys av en artificiell konversation var lyckad, nämligen *tydliga och okomplicerade svar*, *relevanta svar* och *lagom långa svar* (inte för mångordiga men inte heller för sparsamt uttryckta). I dialogen med en chatbot uppskattar användaren också sofistikerade ordval och välkonstruerade meningar och om en chatbot kan hantera dialogen på detta sätt uppfattas den också mer människolik (Lortie & Guitton, 2011).

Zamora (2017) undersökte vilka förväntningar förstagångsanvändare hade på en chatbot innan användning och formulerade utifrån resultatet fyra kategorier av förväntningar; *högpresterande*, *smart*, *sömlös* och *personlig*. I sin tur finns det ett antal faktorer som utmärkte sig speciellt inom varje kategori. *Högpresterande* pekade på om en applikation som var *snabb*, *effektiv* och *pålitlig*. *Smart* inbegrep faktorer som rörde intelligens, *kunnig*, *exakt* och *förutseende*. I den *sömlösa* kategorin värderades *jämna/smidiga upplevelser* som är *enkla* och *flexibla*. I kategorin *personlig* värderades en chatbot som skulle kunna *förstå användaren* (på ett personligt plan) och som är *lätt att tycka om*.

Skjuve et al. (2019) identifierar tre faktorer som påverkade användarens upplevelse av konversationer med chatbotar: (1) *innehållet i konversationerna* - orelevanta eller kryptiska svar, (2) *uppträdandet i förhållande till personlighet* - personlig eller opersonlig i dialogen (detta kunde också påverka hur människolik chatboten upplevdes) och (3) *flöde i konversationen* - om det blir för långa pauser mellan svar. Dessa faktorer kan också påverka varandra. Om exempelvis chatboten ger orelevanta svar kan den uppfattas som sarkastisk eller hånande (Skjuve et al., 2019), vilket är en egenskap som kan skapa misstro mot chatboten (Demeure et al., 2011). När en chatbot minns något som användaren sagt skapar detta engagemang. Det kan vara så enkelt som att den minns användarens namn, när en chatbot visar empati kan detta även skapa empati av samma sort mot chatboten (Portela & Granell-Canut, 2017).

Jain et al. (2018) rekommenderar att titta på följande fem faktorer vid design av chatbotar:

1. Tydliggöra chatbotens kompetens och förmågor
2. Behålla kontext genom en konversation
3. Möjliggöra för småprat (främst vid initiala och avslutande hälsningar och för att implementera humor hos chatboten)
4. Hantera när dialogen misslyckas
5. Avsluta konversationen på ett smakfullt sätt

De såg också att användare föredrog en chatbot som hade ett människoliknande sätt att konversera på eller en engagerande upplevelse som tog tillvara på den bekanta interaktionen i meddelandekontexten med konversationsroller där användare och chatbot turas om att tala (Jain et al., 2018). Användarna förväntade sig att chatboten skulle ställa relevanta frågor för att avgränsa sökandet och styra konversationen åt rätt håll, de uppskattade också om chatboten upplevdes ha en personlighet (Jain et al., 2018). Personligheten förväntas matcha chatbotens syfte, till exempel att den är mer strikt och professionell om det gäller nyheter och humoristisk i underhållningssammanhang (Jain et al., 2018). Dessa faktorer sågs alla ha en koppling till hur människolik chatboten upplevdes.

En positiv användarupplevelse vid interaktion med en chatbot kännetecknas av att användaren får hjälp som är bättre och mer effektiv än det alternativ som använts för att lösa samma uppgift tidigare, om processen blir mer komplex eller krånglig kommer de flesta användare inte att fortsätta använda chatboten (Zamora, 2017). En annan anledning till att inte fortsätta använda en viss chatbot är om förväntningarna användaren har haft på chatboten inte infrias. Då är sannolikheten liten att användaren fortsätter använda chatboten (Zamora, 2017). Om en chatbot utger sig för att vara smart eller använder mänskliga attribut som till exempel humor, kommer förväntningarna kring dess intelligens drastiskt att öka och om användaren då inte upplever att dessa förväntningar matchar med chatbotens faktiska prestation så infrias inte förväntningarna, med besvikelse som följd (Luger & Sellen, 2016). Det finns ofta inte tillräcklig information eller feedback för att användare ska kunna avläsa och få en korrekt bild av vad chatboten klarar av (Jain et al., 2018; Luger & Sellen, 2016). Flera studier visar att användare har väldigt höga förväntningar på vad en chatbot och andra virtuella assistenter ska kunna klara av och i de flesta fall kan dessa förväntningar inte infrias med dagens teknik (Jain et al., 2018; Luger & Sellen, 2016). Till exempel uttrycker deltagare i en studie utförd av Bickmore et al. (2018) en stor frustration men även förvåning över att virtuella assistenter som Apples Siri, Amazons Alexa och Googles Assistant inte klarade av att svara på medicinska frågor. Flera av deltagarna uttryckte att assistenterna borde ha bättre svar, eller åtminstone något svar, när det gällde så pass allvarliga medicinska frågor som de blev tillfrågade.

2.3.2. En mänsklig maskin

En aspekt som påverkar upplevelsen av och engagemanget för interaktionen med chatbotar är om de har en tilltalande personlighet (Jain et al., 2018; Liao et al., 2018; Mone, 2016). Dock är det inte säkert att det är passande i alla sammanhang (Neururer et al., 2018). De behöver inte nödvändigtvis vara människoliknande (McDuff & Czerwinski, 2018; Piccolo et al., 2018), samtidigt finns det nackdelar om chatboten agerar för lite likt mänskligt beteende (Mone, 2016). Strait et al. (2015) undersökte användares upplevelse av fysiska robotar i förhållande till hur människoliknande de är och såg att ett visst mått av människoliknande var positivt för att fånga användarens uppmärksamhet, men om robotarna var för människoliknande så upplevde de obehag och undvek också att interagera med de robotarna. Detta är något som bekräftar Masahiro Moris teori om "den kusliga dalen" (Mori, 1970, 2012). Moris teori bygger på att det är engagerande och positivt ju mer människoliknande robotar är, men detta gäller bara upp till en viss gräns. När de blir för människolika så skapar det istället obehag hos människor. Chatbotar interagerar dock med människor på ett annat sätt än vad fysiska robotar gör, och därför behövs det forskas mer kring hur teorin om "den kusliga dalen" kan överföras till chatbotar och vilka skillnader som kan påverka användarens upplevelse i förhållandet till hur människolika de är (Skjuve et al., 2019). En liknande reaktion, som vid "den kusliga dalen", kan uppstå i förhållande till den emotionella respons som chatboten ger. Till en viss gräns är det positivt att chatboten simulerar känslor men om det blir en för stor emotionell respons så finns det risk att det skapar obehag istället, speciellt om det inte matchar med konventioner och oskrivna regler för mänsklig konversation, då människor är mycket bra på att känna om det är något som inte riktigt stämmer (McDuff & Czerwinski, 2018).

Om en chatbot visar tecken på emotionell intelligens så finns det en risk att användare tolkar det som att den är kapabel till mer än vad den är (Luger & Sellen, 2016; McDuff & Czerwinski, 2018). När chatbotar demonstrerar emotionell intelligens kan det till och med tolkas som att de är mer medvetna i produktionen av sina svar (Yang, Ma, & Fung, 2017). Emotionella reaktioner som är relevanta för sammanhanget skapar också ett mer trovärdigt intryck och användare uppfattar att chatboten är mer kompetent och varm i sitt bemötande (Demeure et al., 2011).

Flera studier i tidigare forskning diskuterar hur man kan gå tillväga för att skapa mer människoliknande chatbotar och virtuella assistenter (Bickmore & Cassell, 2005; Lortie & Guitton, 2011; Paikari & van der Hoek, 2018). Däremot saknas det konsensus kring om detta faktiskt är rätt väg att gå för att bygga förtroende (Neururer et al., 2018; Piccolo et al., 2018; Skjuve et al., 2019).

Att skapa förtroende och ett engagemang som gör att användare faktiskt vill interagera med en chatbot beror delvis på design av personlighet men också designen av konversationerna. Faktorer som har setts kunna bidra till ett ökat förtroende hos användaren är småprat (Bickmore & Cassell, 2005; Portela & Granell-Canut, 2017), att chatboten är öppen och transparent med att den är en maskin (Mone, 2016), att chatboten kan reagera och svara på tecken som indikerar på emotionella och sociala beteenden (McDuff & Czerwinski, 2018) och att chatboten är autentisk och matchar användaren på

ett personligt plan och i konversationerna (Neururer et al., 2018). Autentisk innebär i detta fall att chatboten är medveten om och klarar av att hantera kontext, att den är transparent med sitt syfte och att den kan agera som en person med värderingar, egen attityd och kultur (Neururer et al., 2018).

En aspekt som gör det komplicerat att analysera och designa användarupplevelsen av chatbotar är att alla människor fungerar olika och uppfattningen om, eller upplevelsen av, en chatbot ofta ligger lika mycket i ögonen på betraktaren som i beteendet i chatboten (Lortie & Guitton, 2011). Till exempel finns det forskning som visar att småprat bara gynnar förtroendet för chatboten för extroverta användare, hos introverta påverkade denna faktor inte (Bickmore & Cassell, 2005). Om en chatbot kan anpassa sin personlighet till olika typer av användare skulle det kunna öka acceptansen hos en bredare grupp av användare (Demeure et al., 2011).

Det finns indikationer på att ett för stort mått människolikhet kan påverka användarupplevelsen negativt (Piccolo et al., 2018). Det finns även studier som ifrågasätter om mer människoliknande automatiskt betyder mer omtyckt (McDuff & Czerwinski, 2018). McDuff och Czerwinski (2018) ser delvis svårigheter att motsvara förväntningar som skapas då upplevelsen av den emotionella intelligensen inte motsvarar chatbotens övergripande intelligens och kompetens, men påpekar också det kan existera en "kuslig dal" när det gäller emotionell intelligens då människor är väldigt känsliga för när något inte känns helt rätt.

Det finns forskning som har börjat undersöka hur, och om, "den kusliga dalen" kan överföras till interaktionen med chatbotar (Skjuve et al., 2019). Resultaten pekar på att utvecklingen inte har gått tillräckligt långt och chatbotar är idag inte tillräckligt människoliknande för att de ska kunna gå över gränsen till "den kusliga dalen" (Skjuve et al., 2019). Trots detta så finns det förväntningar från användare att chatbotar idag kan, och kommer, bete sig på ett människoliknande sätt (Paikari & van der Hoek, 2018). Skjuve et al. (2019) såg att det fortfarande är skillnad på "den kusliga dalen"-syndrom vid användande av chatbotar mot upplevelsen av fysiska/visuella robotar. En koppling till "den kusliga dalen" kunde dock ses i förhållande till responstid. Användare kunde känna sig illa till mods när de fick vänta på svar, möjligtvis för att användaren inte visste vad de skulle förvänta sig eller hur de skulle bete sig (Skjuve et al., 2019). I andra fall skapades ett obehag hos användaren om chatboten svarade alltför snabbt och gick emot de informella regler om turordning och betänketid som människor är vana vid från konversationer med andra människor (Mone, 2016).

Om en chatbot ska bete sig som en människa och härma mänskligt beteende eller om den ska utveckla en egen personlighet kan bero på kontext (Neururer et al., 2018). I vissa fall där till exempel dess enda syfte är att representera en organisation eller ett företag så kan chatboten ta över en passande personlighet, i andra fall där den står mer på egna ben kanske också en egen personlighet bör skapas (Neururer et al., 2018). Det finns dock frågor och diskussioner, både inom forskning och designfältet, kring om det är acceptabelt utifrån ett etiskt perspektiv för en chatbot att inte vara transparent med om den är en chatbot eller en människa (Neururer et al., 2018; O'Neill, 2019). Westerman et al. (2018) argumenterar att det kan finnas fördelar att låta en chatbot bete sig mer som en robot än

som en människa då detta kan skapa större trovärdighet. Samtidigt finns det studier som har visat att människor har en tendens att lita mer på andra människor än på tekniska system eller chatbotar, även i de fall där man kan se att systemet eller chatboten presterar bättre än en människa (Dietvorst, Simmons, & Massey, 2015).

2.4. Sammanfattning

Litteraturgenomgången har visat att forskningen kring chatbotar har en del motsättningar. Det finns inte en gemensam konsensus kring om chatbotar borde bli mer människoliknande eller om det är bättre att låta chatbotar vara tydliga med vad de är, vad de kan och låta dem utveckla en egen sorts personlighet (Akma et al., 2018; Jain et al., 2018; Luger & Sellen, 2016; McDuff & Czerwinski, 2018; Mone, 2016; Neururer et al., 2018; Piccolo et al., 2018; Skjuve et al., 2019; Westerman et al., 2018). Det är också svårt att bygga chatbotar som kan simulera mänskligt beteende och mänsklig konversation på ett realistiskt sätt då denna intelligens är komplex och uppbyggd av avancerade färdigheter och kompetenser (Lortie & Guitton, 2011; Schuetzler et al., 2014; Vinciarelli et al., 2015; Warwick & Shah, 2016).

Syftet med litteraturgenomgången var att identifiera faktorer som påverkar användarnas uppfattning om upplevelsen att samtala och interagera med chatbotar. De identifierade faktorerna syns nedan i tabell 1.

Tabell 1. Sammanfattning från litteraturgenomgången av de faktorer som påverkar användarupplevelsen vid interaktion med chatbotar.

Faktor	Studie
Tydliga & okomplicerade svar. Relevanta svar. Inte för mångordig men inte för fåordig.	Chakrabarti och Luger (2015)
Snabb. Effektiv. Pålitlig. Kunnig. Exakt. Förutseende. Jämna/smidiga upplevelser. Enkelhet. Flexibilitet. Förstå användaren (på ett personligt plan). Lätt att tycka om.	Zamora (2017)
Engagerande personlighet. Ärliga med att de är maskiner.	Mone (2016)
Robotisk (skapar trovärdighet).	Westerman et al. (2018)

Småprat.	Bickmore och Cassell (2005)
Innehåll i konversation. Uppträdande (personlighetsmässigt). Flöde i konversationen.	Skjuve et al. (2019)
Tydligt vilken kompetens chatboten har. Samma kontext genom konversation. Möjliggöra småprat (främst vid initiala och avslutande hälsningar och för att implementera humor hos chatboten). Hantera misslyckande i dialogen. Avsluta konversationen på ett smakfullt sätt. Människoliknande konversationssätt. Utnyttja meddelandeformatet på ett bra sätt (turordning i dialog osv). Ställa frågor för att minska sökområde/styra konversationen åt rätt håll (mot rätt svar). Personlighet (som matchar syfte/ändamål).	Jain et al. (2018)
Skapa engagemang genom att tex tilltala användare med namn (empati). Småprat.	Portela och Granell-Canut (2017)
Sofistikerade ordval (mer människolik). Välkonstruerade meningar/meningsuppbyggnad (mer människolik).	Lortie och Guitton (2011)
Dynamisk (ställer följdfrågor) - mer människolik och engagerande. Statisk (ingen återkoppling på det som användaren säger) - mindre engagemang.	Schuetzler et al. (2014)
Kunna svara på sociala och emotionella hintar (mer engagerande och pålitlig - ansågs också mer människolik).	McDuff och Czerwinski (2018)
Visa passande känslor (i sammanhanget) bidrar till att agenten uppfattas mer trovärdig, kompetent och varm.	Demeure et al. (2011)
Klarar av att hantera och stanna i kontext. Transparent med sitt syfte. Persona, inklusive värderingar, egen attityd och kultur.	Neururer et al. (2018)
Negativt med för stor människolikhet.	Piccolo et al. (2018)

2.4.1. Faktorer som används i undersökningen

De faktorer som enligt tidigare studier, rapporterade i litteraturgenomgången, påverkar hur människolik en chatbot upplevs är följande:

1. *Småpratande* (Bickmore & Cassell, 2005; Jain et al., 2018; Portela & Granell-Canut, 2017)
2. *Uppträdande/sätt* (Jain et al., 2018; Skjuve et al., 2019)
3. *Håller kontext i konversationen* (Jain et al., 2018; Neururer et al., 2018)

4. *Naturligt flöde i dialogen (turordning & väntetid)* (Jain et al., 2018; Skjuve et al., 2019)
5. *Sofistikerade ordval* (Lortie & Guitton, 2011)
6. *Välkonstruerade meningar* (Lortie & Guitton, 2011)
7. *Dynamisk i dialogen (ställer följdfrågor)* (Jain et al., 2018; Schuetzler et al., 2014)
8. *Visar passande känslor (i sammanhanget)* (Demeure et al., 2011; Jain et al., 2018)
9. *Kan svara/reagera på sociala och emotionella hintar* (McDuff & Czerwinski, 2018)

I den fortsatta undersökningen kommer det ligga fokus på de faktorer som på ett lämpligt sätt kan undersökas via en enkät. Detta gör att faktorn *håller kontext i konversationen* inte kommer inkluderas i undersökningen, en kort fråga i text kan inte på ett rättvist sätt analysera hur väl en chatbot har hållit kontext från början till slut i en konversation. *Naturligt flöde i dialogen (turordning & väntetid)* har också valts bort, då frågor kring detta med stor sannolikhet hade resulterat i att deltagaren blivit tvungen att försöka föreställa sig hur det känns att vänta några tiondelssekunder mer eller mindre på svar från en chatbot. Denna faktor hade varit mer passande att testa i ett användartest där väntetiden kan simuleras och användaren kan känna efter hur det känns. Faktorn *kan svara/reagera på sociala och emotionella hintar* ansågs också problematisk att undersöka via formatet som är valt för denna studie, därför exkluderades även den. Denna avgränsning gjorde att jag använde mig av följande faktorer i undersökningen i min enkät, nämligen; *småpratande, uppträdande/sätt, sofistikerade ordval & välkonstruerade meningar, dynamisk i dialogen (ställer följdfrågor)* och *visar passande känslor (i sammanhanget)*.

3. Metod

3.1. Litteratursökning

Litteratursökningen inför denna studie genomfördes till stor del via Högskolan Kristianstads digitala sökverktyg Summon@HKR. Ett mindre antal kompletterande sökningar gjordes även direkt i den digitala databasen ACM Digital Library (som har fokus på informationsteknologi och datavetenskap). Böcker rörande AI och konversationsdesign hittades även på Universitetsbiblioteken i Göteborg, både digitalt och i de fysiska biblioteken. Sökningen genomfördes främst med engelska sökord då de flesta vetenskapliga artiklar är skrivna på engelska. De sökord och kombinationer av sökord som användes var: AI, chatbot, perception + chatbot, preference + chatbot, chatbot + humanness, natural language, natural language processing, NLP, NLP + chatbot, Turing test, HCI, uncanny valley, user experience och conversational design. Utifrån de artiklar som jag började skumma igenom hittade jag också många relevanta artiklar genom att gå igenom respektive källförteckning.

3.2. Design av undersökning

För att få en bild av hur de identifierade faktorerna upplevs från en bred grupp användare så valdes en enkät som undersökningsmetod. I enkäten inkluderas frågor där deltagarna presenterades för olika situationer som byggde på de faktorer rörande användarupplevelse och människolikt beteende i chatbotar som identifierats i litteraturgenomgången. Alla frågor och exempel som används i enkäten är skrivna för denna undersökning och helt och hållet självgenererade. I enkäten samlades också grundläggande demografisk data in. De faktorer som undersöktes i enkäten var; *småpratande*, *uppträdande/sätt*, *sofistikerade ordval*, *välkonstruerade meningar*, *dynamisk i dialogen* och *visar passande känslor (i sammanhanget)*. Dessutom undersöktes hur benägna deltagarna var att söka hjälp från en chatbot om de behöver assistans när de är ute på internet och hur personlig chatboten ska vara i sin interaktion med användaren.

3.3. Deltagare

Sammanlagt besvarade 75 personer enkäten. Deltagarna bestod av 35 kvinnor, 39 män och en person som valde att inte delge könsidentitet. Åldersspannet på deltagarna låg mellan 20-61 år och medelåldern låg på 36 år.

3.4. Material

Enkäten skrevs till en början på engelska (se Bilaga 1), detta för att skapa bredd och nå ut till alla delar av de kanaler som den skulle skickas ut i. För att presentera resultatet har en svensk översättning gjorts i efterhand. Den svenska översättningen finns bifogad i sin helhet i Bilaga 2 och det är den som det kommer refereras till framöver i resultat, analys och diskussion.

Enkäten är uppdelad i fyra delar. Den första delen syftar till att samla in demografisk data om deltagarnas ålder, kön, hur mycket tid de spenderar på internet under en dag och om de använder sig av smartphones. Det fanns möjlighet för deltagarna att välja att inte ange information som de eventuellt kan uppfatta som känslig, så som ålder och kön. I nästa del samlades data in om hur ofta och vid vilka tillfällen deltagarna har använt sig av chatbotar. Här inkluderades också en definition för vad en chatbot innebär så att deltagarna fick en bättre gemensam bild av vad begreppet står för specifikt i denna studie. I tredje delen undersöktes om deltagarna söker sig till chatbotar när de behöver hjälp på en hemsida, för att analysera till vilken grad chatbotar är ett önskvärt och använt hjälpmedel idag. Dessutom fick deltagarna svara på vilka typer av chatbotar som de har använt. Den fjärde delen av enkäten består av elva frågor där deltagarna fick tänka sig in i olika situationer och hur de skulle vilja bli bemötta av en chatbot och hur det skulle vara önskvärt att den artikulerade sig. Dessa frågor var, som nämnt ovan, framtagna specifikt för denna undersökning. Nedan följer en genomgång av de elva frågorna och vilken faktor som låg till grund för formuleringen och undersökningen i frågan.

Den första frågan syftar till att undersöka hur på vilket personligt plan användare vill bli bemötta av en chatbot, vem som ska starta konversationen och om användaren vill bli tilltalad vid namn eller inte (*småpratande*). Den andra frågan undersöker hur benägen en

användare är att fortsätta interagera med en chatbot om den inte förstår vad användaren vill, och på vilket sätt man vill att chatboten ska beskriva hur man tillsammans ska ta sig vidare från den situationen (*passande känslor för sammanhanget*). Den tredje frågan undersöker till vilken grad chatboten önskas ha en egen personlighet och till vilken grad den delar med sig av detta (*småpratande, uppträdande/sätt*). Den fjärde frågan undersöker hur detaljerad information som är önskvärd i ett svar (*sofistikerade ordval & välkonstruerade meningar*). Den femte frågan undersöker om användare efterfrågar följdfrågor som säkerställer att chatboten har uppfattat vad frågan handlar om och kan ge ett relevant svar (*dynamisk dialog*). Den sjätte frågan undersöker vilket beteende som är önskvärt i ett mer officiellt sammanhang och om användaren förväntar sig ett liknande bemötande som hen skulle få i en fysisk bank (*uppträdande/sätt*). Den sjunde frågan undersöker vilket beteende som är önskvärt i ett mer avslappnat sammanhang och om användaren förväntar sig att en chatbot som existerar i ett underhållningssammanhang har en mer underhållande personlighet (*uppträdande/sätt*). Den åttonde frågan undersöker hur tydligt användare önskar att en chatbot artikulerar vad den är kapabel till och i vilket sammanhang den verkar (*uppträdande/sätt*). Dessutom kan den ge indikationer på vilken sorts beteende eller personlighet som är att önskvärt hos en chatbot. Den nionde frågan undersöker vilken typ av språkbruk som är önskvärt i en chatbot, i relation till hur välkonstruerade meningarna är (*sofistikerade ordval & välkonstruerade meningar*). Den tionde frågan undersöker hur önskvärt det är att en chatbot kan läsa av situationer där en användare är frustrerad, och till vilken grad den ska ha ett ursäktande beteende (*uppträdande/sätt, visar passande känslor (för sammanhanget)*). Den elfte frågan undersöker hur önskvärt det är att en chatbot ställer relevanta följdfrågor för att användaren ska kunna säkerställa att man får svar på rätt fråga och i slutändan får ett lyckat resultat av sin dialog (*dynamisk dialog*).

För att undersöka reliabiliteten av deltagarnas svar genomfördes ett reliabilitetstest. Testet inkluderade de identifierade faktorerna från litteraturgenomgången (*småpratande, uppträdande/sätt, sofistikerade ordval, välkonstruerade meningar, dynamisk i dialogen och visar passande känslor (i sammanhanget)*) och resultatet visade ett värde av 0.59 baserat på Cronbachs Alpha (Cronbach, 1951). Detta värde indikerar att enkäten hade en god intern reliabilitet.

3.5. Procedur

Enkäten skapades och genomfördes på engelska med hjälp av verktyget Emarketeer. Valet att använda en enkät på engelska gjordes för att skapa möjligheter att nå ut till en större grupp deltagare. Enkäten skickades ut dels i kanaler där det fanns en stor representation av tekkunniga deltagare (en IT-avdelning på ett större företag i Göteborg), dels i kanaler där det fanns människor som använder chatbotar (anställda och kunder på företag som jobbar med chatbotar) och enkäten delades även inom kanaler på sociala medier (LinkedIn och Facebook) för att sammantaget nå ut till en bred bas av användare med olika mycket erfarenhet av chatbotar. Enkäten låg ute under tio dagar och sammanlagt kom det in 75 fullständiga svar.

För att sprida enkäten användes dels ett mail med inbjudan till att delta som skickades ut till IT-avdelningen på ÅF i Göteborg. Till övriga målgrupper skickades en länk ut via mail, samma länk publicerades också på sociala medier. När deltagarna gick in på länken välkomnades de av ett meddelande som förklarade syftet till enkäten, att den skulle ta cirka fem minuter att genomföra och att alla svar skulle behandlas med full anonymitet (för fullständigt meddelande, se Bilaga 2). Under de tio dagarna skickades en påminnelse ut efter en vecka med förhoppningen att samla in några extra svar.

När insamlingen av data avslutades efter tio dagar exporterades datan dels till ett pdf-dokument där alla frågor och svar visualiserades i diagram (denna visualisering skapades automatiskt i Emarketeer). Datat exporterades också till ett kalkylark där alla svar sparades i ett flöde så att det gick att avläsa varje separat deltagares svar. Datat analyserades i relation till forskningsfrågan och med syfte att undersöka vilka faktorer som kunde indikeras påverka användarupplevelsen positivt och om de var kopplade till faktorer som setts påverka hur människoliknande beteende chatboten upplevs ha. För att göra detta analyserades först de frågor som berörde samma faktor i grupper tillsammans. Efter detta undersöktes kopplingar mellan hur vana internetanvändare deltagarna var och hur detta påverkade resultaten. Datat analyserades både genom att titta på svaren utifrån grafer som skapades kring hur deltagarna sammanlagt svarade på varje enskild fråga, men också genom att studera varje deltagares individuella svar och flödet genom hela undersökningen.

3.6. Etiska överväganden

Vetenskapsrådet (2002) har artikulera fyra forskningsetiska principer; *informationskravet*, *samtyckeskravet*, *konfidentialitetskravet* och *nyttjandekravet*. *Informationskravet* innebär att forskaren ska informera de som berörs av undersökningen om forskningens syfte. *Samtyckeskravet* innebär att deltagare i en undersökning ska ha rätt att bestämma själva över sin medverkan. *Konfidentialitetskravet* innebär att uppgifter om alla deltagare ska hållas konfidentiella och att de ska sparas där utomstående inte kan komma åt informationen. *Nyttjandekravet* innebär att datan som samlas in endast får användas i forskningssyfte.

Potentiella deltagare till enkäten i denna studie informerades om vad datan skulle användas till; en c-uppsats på kandidatnivå som behandlar förväntningar och uppfattningar om chatbotar och rekommendationer för hur man som designer kan tänka när man designar chatbotar. Det var frivilligt att genomföra enkäten och deltagarna kunde närsomhelst avbryta sitt deltagande i undersökningen. Undersökningen i detta examensarbete har genomförts med deltagarnas fulla anonymitet vilket betyder att ingen data som kan kopplas till individer har samlats in eller sparats. En av deltagarna hade av misstag skrivit in sitt förnamn i textrutan där man ombads ange sin ålder. För att säkerställa anonymitet raderades denna information när datan exporterades från tjänsten Emarketeer. Uppgifterna som kom in via undersökningen används endast i forskningssyfte i denna studie.

4. Resultat och analys

Av deltagarna som besvarade enkäten angav alla förutom en (98,7%) att de använder sig av en smartphone. I tabellen nedan syns en sammanställning av hur mycket tid deltagarna tillbringar på internet och hur vana de är att använda chatbotar.

Tabell 2. Sammanfattning av deltagarnas vana vid att använda internet och chatbotar.

Fråga	Antal svar	Svarsfrekvens
Antal timmar jag spenderar på internet varje dag:		
0–1	2	2,7%
2–5	36	48%
6–10	30	40%
11–20	7	9,3%
Hur stor erfarenhet har du av att interagera med chatbotar?		
Ingen	8	10,7%
Några gånger per år	36	48%
Några gånger i månaden	15	20%
Några gånger i veckan	10	13,3%
Varje dag	6	8%

Procentuellt var det störst andel av gruppen som använder internet 11–20 timmar per dag som också använder chatbotar dagligen, 28,5%. I gruppen som använder internet 6–10 timmar per dag använde 3,3% chatbotar dagligen, i gruppen som använde internet 2–5 timmar varje dag var det 8% som använde chatbotar dagligen. I gruppen som använde internet 0–1 timme per dag var det ingen som använde chatbotar dagligen. Detta kan peka på att det är större chans att en användare som är ute mycket på internet också använder chatbotar oftare, dock är det så pass stor skillnad på storleken av dessa grupper att det är svårt att dra några generella slutsatser.

I efterföljande fråga undersöktes i vilken situation deltagarna har använt chatbotar. Här var det endast sju stycken som angav att de inte använt en chatbot i någon av de situationer som angavs, vilket indikerar att det faktiskt är sju personer (9,3%) från gruppen som aldrig använt en chatbot, och inte åtta (som angavs i föregående fråga). De fem vanligaste situationerna som chatbotar användes i var; *kundtjänst* (64%), *för att leta efter information* (40%), *för att spela musik* (25,3%), *enkäter/utvärderingar* (21,3%) och *för att fråga efter vädret* (18,7%). En sammanställning av situationer där chatbotar användes ses nedan.

Tabell 3. Sammanställning av i vilka situationer deltagarna har använt chatbotar.

Fråga	Antal svar	Svarsfrekvens
<i>Jag har använt en chatbot i följande situationer: Du kan ange mer än ett alternativ.</i>		
För att leta efter information.	30	40%
För att boka möten.	2	2,7%
För att fråga efter vädret.	14	18,7%
För att spela musik.	19	25,3%
För att ha någon att prata med.	3	4%
För att få inspiration när jag handlar på internet.	4	5,3%
Kundtjänst.	48	64%
Enkäter/utvärderingar.	16	21,3%
I jobbrelaterade situationer.	9	12%
Ingen av ovanstående.	7	9,3%
Om andra, vänligen ange när.	11	14,7%

Andra situationer som deltagarna angav att de använt chatbotar i var; för att släcka lampor och hantera smarta funktioner i hemmet, för att implementera dem på hemsidor, i Microsoft Office-produkter, vid experiment för att skapa en egen chatbot, för skojs skull & för att se om det verkligen är användbart, marknadsföring, smarta hem-funktioner, för att hantera alarm och andra grundläggande vardagsuppgifter, kundservice och för jobb. De chatbotar eller virtuella assistenter som var mest använda var webbaserade chatbotar (57,3% av deltagarna hade använt dessa), Apples Siri (49,3%), chatbotar som skapats i Facebooks Messenger (25,3%) och Googles virtuella assistent (24%).

Två av frågorna undersökte till vilken grad deltagarna sökte sig till chatbotar om de behövde hjälp på en hemsida, eller inte hittade vad de letade efter. 66,7% angav att de frågar chatbotar om hjälp ibland (om de finns tillgängliga), 25,3% använde sig aldrig av chatbotar vid dessa tillfällen och 8% angav att de ber chatbotar om hjälp hela tiden. När deltagarna sedan fick ange på vilket sätt de föredrog att be om hjälp på hemsidor (här var det möjligt att välja mer än ett alternativ) var de tre vanligaste alternativen att maila kundservice (69,3%), använda ett kontaktformulär (64%) och att använda en chatbot (52%). Det var inte lika många som angav att det använder sig av chatbotar när de blev ombedda att ange vilka olika alternativ de använder sig av om de behöver hjälp på en hemsida (52%) som det var när frågan var hur frekvent de ber om hjälp av chatbotar när de behöver hjälp på en hemsida (sammanlagt 74,7% angav i föregående fråga att de använder sig av en chatbot för att få hjälp ibland eller hela tiden). Detta kan bero på att

de inte tänkte på de andra alternativen när de endast blev tillfrågade om hur ofta de använder chatbotar, och när de fick ange flera andra alternativ så ansåg de att dessa alternativ var att föredra, dock går det inte att avläsa på ett säkert sätt vad denna diskrepans beror på.

Ovanstående data samlades in för att få en övergripande förståelse över hur vana deltagarna var att använda internet och smartphones överlag och chatbotar generellt. I följande avsnitt av enkäten bads deltagarna att tänka sig in i olika situationer och ange hur de föredrog att bli tilltalade av en chatbot i de olika fall som presenterades. Detta gjordes utifrån de faktorer som identifierats från litteraturgenomgången.

4.1. Småpratande

Chatbotens småpratande har undersökts utifrån om den delar med sig av information som inte är relevant i förhållande till uppgiften men också hur informell den är eller om den tilltalar användaren på ett personligt sätt, till exempel genom att tilltala vid namn. Detta gjordes främst i fråga 1 och 3.

Tabell 4. Sammanställning av svar på de frågor som undersökte användarens inställning till småprat i interaktionen med en chatbot.

Fråga	Antal svar	Svarsfrekvens
1. Jag föredrar följande hälsning från en chatbot:		
Att den säger hallå.	11	14,7%
Att den säger hallå, följt av mitt namn.	7	9,3%
Att den säger hallå, följt av mitt namn och frågar vad jag vill ha hjälp med.	27	36%
Det ska vara mitt initiativ att starta en konversation.	30	40%
3. Du har bett en chatbot om hjälp att hitta resmål till din nästa affärsresa. Vilken uppdatering föredrar du?		
"Snart tillbaka."	3	4%
"Söker efter resmål..."	49	65,3%
"Går igenom alla spännande alternativ."	8	10,7%
Jag vill inte ha en uppdatering, bara resultatet när chatboten är färdig med sökningen.	6	8%
"Åh, en resa! *klappar händerna* Får jag följa med? Tänk på det medan jag söker upp de bästa alternativen för ditt resmål."	7	9,3%
"Svävar iväg, drömmer om det perfekta resmålet, snart tillbaka..."	2	2,7%

I den första frågan angav en stor grupp av deltagarna att de inte ville att chatboten skulle säga något innan de gjorde det (40%). Samtidigt var det en relativt stor grupp som ville bli hälsade både vid namn och att chatboten skulle fråga vad man behövde hjälp med. Det är inte helt klart vilken del av detta svarsalternativ som lockade mest, om det var det personliga tilltalet eller att direkt få frågan om vad man behöver hjälp med, dock var det en förhållandevis liten grupp som föredrog att chatboten endast hälsade med användarens namn vilket kan antyda att det är frågan om vad man behöver hjälp med som är viktigast i sammanhanget.

I ett senare skede av interaktionen, när användaren får hjälp av chatboten att söka efter resmål och väntar på ett svar, var det önskvärt att få en uppdatering som visade att chatboten tänkte, men de flesta föredrog att den ändå höll sig kort, koncis och inte försökte skapa kontakt kring något som inte var direkt kopplat till uppgiften. De svar som gav indikation på att chatboten ville småprata eller interagera genom att säga saker som inte var direkt relevanta för uppgiften var inte populära hos de flesta deltagare.

Sammantaget indikerar detta att deltagarna inte har ett behov av att småprata med chatbotar. Detta kan vara situationsbaserat (beroende på om chatboten verkar som en service/tjänsteassistent eller om den används i sociala sammanhang), vilket kan utforskas djupare i framtida forskning.

4.2. Uppträdande/sätt

Fem av frågorna i enkäten kan bidra till analysen kring hur deltagarna upplever chatbotens uppträdande eller sätt att bete sig. För fråga 3, 6 och 7 var detta det primära fokuset för undersökningen, men även i fråga 8 och 10 kan man avläsa beteenden som är relaterade till chatbotens uppträdande och sätt att bete sig. Uppträdandet kan ses dels i om chatboten använder humor eller en ton som antyder att den har en personlig åsikt om ämnet men också genom hur avslappnat eller officiellt den uttrycker sig gentemot användaren.

Tabell 5. Sammanställning av svar på de frågor som undersökte användarens inställning till olika typer av uppträdande/sätt hos chatboten.

Fråga	Antal svar	Svarsfrekvens
3. Du har bett en chatbot om hjälp att hitta resmål till din nästa affärsresa. Vilken uppdatering föredrar du?		
"Snart tillbaka."	3	4%
"Söker efter resmål..."	49	65,3%
"Går igenom alla spännande alternativ."	8	10,7%

Jag vill inte ha en uppdatering, bara resultatet när chatboten är färdig med sökningen.	6	8%
"Åh, en resa! *klappar händerna* Får jag följa med? Tänk på det medan jag söker upp de bästa alternativen för ditt resmål."	7	9,3%
"Svävar iväg, drömmer om det perfekta resmålet, snart tillbaka..."	2	2,7%
6. Du interagerar med en chatbot som tillhör din bank. Vilket öppningsmeddelande föredrar du?		
"Yo, var är pengarna?"	3	4%
"Hallå, hur kan jag hjälpa dig idag?"	31	41,3%
"Hej! Vill du spara eller spendera idag?"	6	8%
"God morgon. Jag är din bankassistentbot. Hur kan jag hjälpa dig idag?"	35	46,7%
7. Du interagerar med en chatbot som hjälper dig spela musik. Vilket öppningsmeddelande föredrar du?		
"Vänligen ange genre, artist eller sångtitel för att få lämpliga alternativ."	6	8%
"Berätta för mig vad du vill lyssna på så kan jag hjälpa dig att hitta det!"	40	53,5%
"Hej! Kul att se dig igen, vilken genre känner du för att lyssna på idag?"	22	29,3%
"Groovy, du är tillbaka! Jag har lite härliga låtar till dig, är du redo att släppa loss?"	7	9,3%
8. Du interagerar för första gången på jobbet med en chatbot för att boka möten. Vilket öppningsmeddelande föredrar du?		
"Hej, jag är din möteskompis! Vad har du på hjärtat? :)"	4	5,3%
"Hallå! Kul att se dig, låt mig veta hur jag kan hjälpa dig."	11	14,7%
"Vill du boka ett rum eller hantera din kalender?"	24	32%
"Välkommen, vill du se tillgängliga tider i din kalender?"	14	18,7%
"Välkommen, jag är en bot som kan hjälpa dig att hantera din kalender och boka mötesrum. Vill du att jag föreslår lämpliga tider för ett möte?"	22	29,3%
10. Du är frustrerad för att en chatbot inte kan förstå dig och du har uttryckt din frustration till chatboten. Vilket svar föredrar du?		
"Jag förstår inte. Vänligen försök igen."	16	21,3%
"Jag är ledsen, det verkar som att jag inte hänger med, är du vänlig och omformulerar och försöker igen?"	27	36%

"Jag är verkligen ledsen, jag vill verkligen hjälpa dig men jag förstår helt enkelt inte. Vänligen omformulera eller försök med enklare ämnesord."	25	33,3%
"Jag är hemskt ledsen, det här är inte hur jag ville att det skulle bli heller..! Jag gör mitt bästa för att hjälpa dig, var snäll och ha tålamod så ska jag göra allt för att gottgöra dig."	7	9,3%

I alla frågor syns en klar preferens mot de svarsalternativ där chatboten håller sig till ämnet på ett någorlunda kort och koncist sätt, och med en officiell eller professionell ton. I fråga 3 valde den största delen av deltagarna (sammanlagt 77,3%) svarsalternativ som var kortfattade och on point. 8 av de procenten uppgav till och med att de inte ville ha någon uppdatering medan de väntade utan endast svaret när det var färdigt att levereras.

Även i ett sammanhang som är mer privat och i ett underhållande syfte, att lyssna på musik (fråga 7), föredrog en majoritet av deltagarna svar som var neutrala eller mer kortfattade. Storleksmässigt representerar gruppen som föredrog det minst personliga alternativet och de som föredrog det mest personliga tilltalet (med humoristisk ton) ungefär lika många deltagare, dessa båda grupper är också de minsta med 8% resp 9,3% av deltagarna som föredrog dessa svar. Den största delen av deltagarna föredrog istället ett svar som är kort och koncist och det verkar inte heller spela någon roll om tonen matchar den mer avslappnade och underhållande uppgiften, att lyssna på musik.

I fråga 8 fanns det två alternativ där chatboten ville etablera en mer kompisliknande relation eller med en liten ton av småpratande eller personligt igenkännande minst populära och det var endast 20% som sammanlagt föredrog dessa alternativ. Även här föredrog deltagarna tydliga och officiella alternativ. Detta återkommer i fråga 10 där den största delen av deltagarna angett att de föredrar svar där chatboten uttrycker att den är ledsen för att den inte förstår, och där den ger ett bud på vad man kan förändra i sin fråga för att hjälpa den att förstå, till exempel att använda ämnesord eller att omformulera frågan. Ett visst mått av ursäktande var önskvärt, så länge chatboten inte uttrycker sig för yvigt eller emotionellt.

Sammanfattningsvis föredrog en majoritet av deltagarna ett uppträdande där chatboten agerade officiellt; genom att den var tydlig, koncis och konverserade med en officiell ton. Resultaten visar på att det är viktigare för chatboten att vara tydlig och transparent än att försöka ha en egen personlighet eller underhålla.

4.3. Sofistikerade ordval och välkonstruerade meningar

Två frågor i enkäten var specifikt inriktade på att undersöka hur chatboten artikulerade sig i fråga om sofistikerade ordval och välkonstruerade meningar, fråga 4 och fråga 9. Svarsalternativen är formulerade utifrån mer eller mindre korrekt punktuation, ordval och meningslängd.

Tabell 6. Sammanställning av svar på de frågor som undersökte användarens inställning till sofistikerade ordval och välkonstruerade meningar i chatbotens dialog.

Fråga	Antal svar	Svarsfrekvens
4. Du har frågat en chatbot när nästa buss till en tågstation går. Vilket svar föredrar du?		
"12:15"	7	9,3%
"Buss 11 mot *slutstation* går 12:15."	31	41,3%
"Buss 11 går 12:15, buss 24 går 12:18 och buss 11 går 12:25."	29	38,7%
"Det finns fem olika bussar som går till tågstationen från din hållplats, nummer 11, nummer 24, nummer 4, nummer 13 och nummer 121. Buss 11 går 12:15, buss 24 går 12:18, buss 4 går 12:26, buss 13 går 12:30 och buss 121 går 12:31."	8	10,7%
9. Du har bokat en resa med en chatbot. Vilket svar föredrar du?		
"Resa bokad avgång imorgon 9:05 ankomst 12:08 i Stockholm."	9	12%
"Din resa är bokad. Avgång imorgon 9:05 och ankomst 12:08 i Stockholm."	23	30,7%
"Din resa är bokad. Ditt tåg avgår från stationen imorgon 9:05 och ankommer till Stockholm 12:08."	43	57,3%

Gemensamt för båda dessa frågor är att deltagarna verkar föredra när chatboten samtalar på ett korrekt sätt och med meningsbyggnad och ordval som är välkonstruerade och gärna med sofistikerade ordval. I fråga 9 föredrog en majoritet av deltagarna (57,3%) det alternativ där svaret var mest välartikulerat, officiellt och mer sofistikerat, även om detta också blev det alternativ där meningen blev längst. Detta kan indikera att de svar som anges som är mest lika hur människor samtalar eller skriver i en dialog med andra människor i en officiell situation är att föredra. Dock behöver detta undersökas vidare, för att se om det är just själva sofistikationen och konstruktionen av meningarna som är det viktiga eller att de liknar en mänsklig konversation som är av officiell karaktär.

4.4. Dynamisk dialog

En dynamisk dialog i detta sammanhanget syftar på en dialog där chatboten klarar av att ställa följdfrågor för att säkerställa att den levererar rätt svar i rätt situation och att den klarar av att hålla kontext genom ett samtal. Främst undersöktes i denna studie om användaren uppskattade relevanta följdfrågor i ett antal olika situationer och detta gjordes i fråga 5 och 11.

Tabell 7. Sammanställning av svar på de frågor som undersökte användarens inställning till hur dynamisk dialogen med chatboten var, främst i förhållande till om chatboten ställer relevanta följdfrågor.

Fråga	Antal svar	Svarsfrekvens
5. Du har frågat en chatbot om morgondagens väder. Vilket svar föredrar du?		
"I vilken stad?"	4	5,3%
"15 grader."	0	0%
"15 grader och soligt med lite moln."	12	16%
"I Göteborg kommer det vara 15 grader och soligt med lite moln imorgon."	59	78,7%
11. Du ber en chatbot att hjälpa dig välja mellan två hårtrimmers. Vilket svar föredrar du?		
"Köp det andra alternativet."	5	6,7%
"Köp det andra alternativet, den har fått högst betyg i kundrecensioner."	35	46,7%
"Behöver du en trimmer för kropp, skägg eller hår?"	32	42,7%
"Du köpte en trimmer för kropp för fem månader sedan, är den här för skägg eller hår?"	3	4%

I fråga 5, var det endast 5,3% av deltagarna som föredrog att få en följdfråga som säkerställer att chatboten visste vilken stad som användaren vill veta vädret i innan den gav ett svar. Den största gruppen (78,7%) föredrog ett utförligt svar där chatboten angav vilken stad som den kollade vädret i och sedan angav vädret, där kommer alltså en bekräftelse på att de har förstått varandra i svaret, utan en följdfråga mellan. I fråga 11 var det en större grupp (42,7%) som föredrog att få en följdfråga där de kan tydliggöra vad de är intresserade av. Det kan bero på att det inte är lika tydligt i de andra alternativen hur väl chatboten har förstått frågan. I det mest populära alternativet får deltagarna rådet att köpa det andra alternativet för att det har fått högst betyg i kundrecensioner. Alltså får de en bekräftelse på varför de ska välja det alternativet, men det är egentligen inte så tydligt utifrån vilka specifika preferenser eller funktioner som chatboten har gjort den rekommendationen. Trots detta föredrar nästan hälften av deltagarna detta alternativ, vilket visar på att det är ett fullgott råd för många. Det gör också att det är svårt att säga att det finns en generell väg att gå. Alternativen är olika men båda tilltalar en relativt stor grupp användare. Det som går att se är att det kan vara en god idé att chatboten är tydlig med vad den svarar på, sedan kan detta ske genom att ställa följdfrågor eller genom att den upprepar frågan i svaret för att säkerställa att båda parter är medvetna om vad som

händer i interaktionen. Det krävs dock ytterligare undersökningar för att kunna avgöra vad denna skillnad beror på och i vilka sammanhang relevanta följdfrågor och dynamiska konversationer är önskvärda. Vad gäller kontext har detta inte undersökts till någon större grad i denna undersökning, dock kan man se att deltagarna inte föredrog det alternativ där chatboten minns tidigare inköp och ställer följdfrågor utifrån detta ("Du köpte en trimmer för kropp för fem månader sedan, är den här för skägg eller hår?").

4.5. Passande känslor för sammanhanget

Två frågor i enkäten undersökte vad som händer när en chatbot inte förstår var användaren vill, fråga 2 och fråga 10. Svartalternativen är formulerade för att visa olika sätt chatboten kan guida användaren i hur hen kan ta sig framåt för att hjälpa chatboten att förstå och även för att undersöka hur känslig chatboten ska vara för frustration, hur den reagerar på en emotionell nivå.

Tabell 8. Sammanställning av svar på de frågor som undersökte användarens inställning till om och hur chatboten visar passande känslor för sammanhanget.

Fråga	Antal svar	Svarsfrekvens
2. Jag föredrar följande svar när en chatbot inte förstår mig:		
"Jag är ledsen, jag förstår inte."	10	13,3%
"Jag är ledsen, jag förstår inte. Kan du kanske hjälpa mig genom att ställa din fråga på ett annat sätt?"	28	37,3%
"Jag är ledsen, jag förstår inte. Vill du prata med en människa istället?"	32	42,7%
"Jag vet inte vad du menar, försök igen!"	1	1,3%
"Jag har sökt internet på internet efter "[din föregående fråga/påstående]"..."	4	5,3%
10. Du är frustrerad för att en chatbot inte kan förstå dig och du har uttryckt din frustration till chatboten. Vilket svar föredrar du?		
"Jag förstår inte. Vänligen försök igen."	16	21,3%
"Jag är ledsen, det verkar som att jag inte hänger med, är du vänlig och omformulerar och försöker igen?"	27	36%
"Jag är verkligen ledsen, jag vill verkligen hjälpa dig men jag förstår helt enkelt inte. Vänligen omformulera eller försök med enklare ämnesord."	25	33,3%
"Jag är hemskt ledsen, det här är inte hur jag ville att det skulle bli heller..! Jag gör mitt bästa för att hjälpa dig, var snäll och ha tålamod så ska jag göra allt för att gottgöra dig."	7	9,3%

I fråga 2 kan man se att i båda de svarsalternativ som föredrogs av en stor grupp av deltagarna uttrycker chatboten att den ber om ursäkt och den ger också ett förslag på hur användaren kan gå vidare för att lösa situationen. Även i fråga 10 föredrog många av deltagarna ett svar där chatboten bad om ursäkt och gav ett förslag på hur de kan gå vidare i dialogen (sammanlagt 69,3%).

Gemensamt för båda frågorna är att deltagarna uppskattar ett svar där chatboten ber om ursäkt för besväret och ger bud på hur användaren kan gå vidare för att förhoppningsvis få ett bättre resultat. Ett visst mått av emotionell intelligens verkar vara önskvärt, dock var det inte populärt om chatboten drar ut på ursäkterna och uppvisar stora känslor.

4.6. Jämförelse mellan olika grupper

Om man jämför de grupper av deltagarna som sagt att det inte har någon som helst erfarenhet av att interagera med chatbotar med de som använder chatbotar varje dag finns det en del intressanta iakttagelser att nämna. Alla utom en i gruppen som använde chatbotar varje dag (83,3%) tyckte att det borde vara deras initiativ att starta en konversation med en chatbot, de ville inte bli hälsade av chatboten. I den oerfarna gruppen var det tvärtom bara en deltagare som inte ville att chatboten skulle säga något innan (12,5%), de andra föredrog någon typ av hälsning eller öppningsfras.

När en chatbot inte förstår föredrog 66,6% av de vana chatbotanvändarna att få alternativet att prata med en människa istället. Ingen föredrog ett svar där chatboten guidade användaren till att omformulera eller ställa frågan på annat sätt. I den oerfarna gruppen var det större spridning på svaren, dock var det även här 50% som ville få alternativet att kunna prata med en människa istället. Svaren antyder också att det inte är lika viktigt för den erfarna gruppen att chatboten har emotionell intelligens när det kommer till att få bekräftelse på sina känslor vid frustration. Den oerfarna gruppen uppskattade att få en tydligare ursäkt med mer simulerade känslor bakom, detta var inte önskvärt på samma sätt hos den erfarna gruppen.

De erfarna chatbotanvändarna hade större benägenhet att välja kortfattade och precist artikulerade svar. De var också mer överens inom gruppen och valde oftare samma svar än vad den oerfarna gruppen gjorde. En annan skillnad som uppmärksammades var att den erfarna gruppen inte verkade tycka att det var lika viktigt att få artikulerat tydligt i dialogen med chatboten vad den är, vad den ska och kan göra osv, det var vanligare att den oerfarna gruppen ville ha denna typen av information.

Den erfarna gruppen verkade mer benägna att uppskatta att chatboten ställer följdfrågor för att garantera relevant hjälp. Den oerfarna gruppen valde oftare ett svarsalternativ där de fick ett direkt råd.

Ovanstående iakttagelser kan hänga samman med att den erfarna gruppen har en större förståelse för vad chatboten faktiskt klarar av och vad den är och gör "på riktigt", oerfarna användare har en tendens att ha högre förväntningar på chatbotar och de tror även att de klarar av mer komplexa uppgifter och kan förstå, eller snarare simulera förståelse och känslor, på ett högre plan (Luger & Sellen, 2016). Den erfarna gruppen

som inte förväntar sig detta har då heller inget behov av att chatboten simulerar att den känner en genom att hälsa med namn, att den simulerar emotionell intelligens, och de förväntar sig inte heller att den förstår allt utan uppskattar följdfrågor och förtydliganden så att båda parter är med på vad som händer i dialogen.

Inga utmärkande skillnader kring preferenser sågs mellan könen. Åldersmässigt var det inte heller många utmärkande skillnader. En iakttagelse är dock att ingen av de deltagare som var över 40 år föredrog de alternativ där chatboten hade en humoristisk ton eller uttryckte en egen personlighet. Överlag var det tydligare att det fanns stora skillnader på preferenser inom alla grupper snarare än att det fanns skillnader mellan grupper.

4.7. Designriktlinje för design av chatbotar

Designriktlinjer är ett sätt att formulera resultat inom designforskning (Dix, 2010; Dourish, 2004; Faraon, 2018). Inom utbildningsforskning har van den Akker (1999) formulerat en mall för hur designriktlinjer kan formuleras, översatt till svenska ser den ut såhär: *“Om du vill designa X [för ändamålet/funktionen Y i kontext av Z], så är du rekommenderad att ge designen egenskaperna A, B och C [fokus på substantiv], och att göra det via förfarande K, L och M [fokus på process], på grund av argument P, Q & R.”*. Den designriktlinje som kan artikuleras idag utifrån denna studie lyder som följer:

- Om du vill designa en chatbot som möjliggör för en positiv användarupplevelse så rekommenderas du att ge chatboten en persona som är tydlig, transparent med vad den är och som håller en officiell och professionell ton. Detta kan göras genom att använda ett sofistikerat språk med välformulerade meningar och ett förfarande där chatboten har en officiell ton och svarar tydligt och transparent i en dynamisk dialog. Möjlighet att reagera med viss emotionell respons ska ges vid de tillfällen då dialogen misslyckas och användaren kan bli frustrerad.

Denna riktlinje bygger på att användares preferenser på vad en positiv dialog är skiljer sig markant från person till person och att en majoritet av deltagarna i denna studie uppskattade när en chatbot betar sig och konverserar på ett officiellt eller formellt sätt. Detta har kunnat ses genomgående i resultaten i denna studie, som också har bekräftat resultat i tidigare studier (Demeure et al., 2011; Lortie & Guitton, 2011; McDuff & Czerwinski, 2018). Humor och liknande beteende skapar också höga förväntningar hos användaren som många gånger inte kan infrias i vidare interaktion med chatboten (Luger & Sellen, 2016). I de fall där chatboten har en mer avancerad intelligens och kan läsa av och lära sig från tidigare interaktioner kan den också anpassa sig mer efter varje individuell användare för att interagera med ett språk och en personlighet som matchar användarens förväntningar. Detta kan då vara positivt och det är också en av de stora möjligheterna med interaktion som sker via NLP och chatbotar eller andra virtuella assistenter, att interaktionen kan anpassas till en högre grad efter varje användare för att skapa en upplevelse som passar varje individ. Dock behöver fortsatt hänsyn tas till att utforska om chatbotens personlighet ska ha mänskliga drag eller om den ska byggas upp med egna karaktäristika.

4.8. Sammanfattning

Som nämnts ovan har de faktorer som undersöks i enkäten valts ut för att de har setts påverka hur människoliknande en chatbots beteende upplevs av användare. Utifrån detta har datan analyserats för att se om det går att koppla slutsatser kring hur attraktivt ett människoliknande beteende är för användarupplevelsen av chatbotar.

På flera håll i undersökningen kan man se att det är relativt stora grupper som har skilda preferenser när det kommer till interaktionen med chatbotar, till och med motpoler. Detta gör att det är svårt att ge generella riktlinjer för designandet av chatbotar i nuläget. Människor har personliga preferenser kring hur de vill interagera med andra människor, och i förlängningen med AI och chatbotar när detta sker genom samma form av dialog som i en mänsklig konversation. En rekommendation är därför att skapa chatbotar som kan personaliseras och som användare med olika preferenser kan interagera med på sina villkor och efter sina egna preferenser (McDuff & Czerwinski, 2018). När människor interagerar med varandra så anpassar man sig ofta till den andra parten, man känner av om hen känner sig obekvämt och anpassar ton och tilltal på ett sätt så att konversationen känns behaglig för båda parter. För att klara av detta krävs dock en komplex intelligens och kompetens hos chatboten (Lortie & Guitton, 2011; Schuetzler et al., 2014; Vinciarelli et al., 2015; Warwick & Shah, 2016).

Sammantaget finns det ett par slutsatser som kan dras av resultaten från denna studie. Genomgående har deltagarna föredragit de svar och alternativ som varken är för statiska eller för personliga. Det är också en stor grupp som har angett att de föredrar att inte chatboten startar en konversation innan de själva har sagt eller frågat något. En analys är att det beror mycket på tycke och smak, och att olika människor uppfattar och föredrar olika drag hos en chatbot. Detta är något som stöds av tidigare forskning, bland andra McDuff och Czerwinski (2018), Mone (2016) och Lortie och Guitton (2011).

En generell tolkning som kan göras utifrån resultaten i denna studie är att en stor del av användare föredrar ett beteende och ett språkbruk som inte är alltför enkelt men inte heller för extremt eller långrandigt. I förhållande till människolikhet är det svårt att dra generella slutsatser. De faktorer som har setts påverka om en chatbot upplevs som människolik eller inte, och som har varit passande att undersöka via en enkät, är, som nämnts tidigare; *småpratande, uppträdande/sätt, sofistikerade ordval & välkonstruerade meningar, dynamisk i dialogen /ställer följdfrågor*) och *visar passande känslor (i sammanhanget)*. Även om deltagarna överlag uppskattade ett sofistikerat språk med välkonstruerade meningar, en dynamisk dialog och ett visst mått av känslor som speglar den aktuella situationen, så är det svårt att avgöra om detta sammanlagt faktiskt utgör ett typiskt mänskligt beteende. Dels existerar det faktum att ingen människa är den andra lik, och dessutom finns det frågetecken kring om dessa faktorer rakt av representerar ett generellt mänskligt beteende. Det är inte ovanligt att människor när de skriver sms eller andra korta meddelanden genom digitala medier använder sig av mycket förkortningar, slang och även utelämnar korrekt punktuation, till viss del i utbyte mot emojis. Trots detta så är alltså sofistikerade ord och välkonstruerade meningar en av de faktorer som har blivit angivna som människoliknande. Den typen av mänskligt beteende ses kanske oftare idag i officiella sammanhang och är inte lika vanligt i vardagsspråk. Att deltagarna i

denna studie även angett att det är önskvärt med ett visst mått av emotionell intelligens och emotionell respons, men däremot inte hade ett behov av att småprata, pekar också på att det inte är en konsekvent bild av mänskligt beteende som efterfrågas. Däremot kan man se många av dessa karaktäristika i människor i offentliga sammanhang, i en officiell roll eller en serviceroll. En hypotes man kan skapa utifrån detta är att en chatbot upplevs som en servicetjänst och därför kan ha en fördel av att anpassa sitt beteende efter en liknande roll gentemot användaren. Chatbotar som befinner sig i olika kontexter och med olika inriktning ska möjligtvis inte bete sig på samma sätt (Neururer et al., 2018), detta är dock något som behöver undersökas vidare.

I många fall finns det en stor bredd på de svar som kommit in i de olika frågorna i denna studie. Det är inte heller ovanligt att de svar som skiljer sig mest från varandra båda är representerade av relativt stora grupper av deltagarna. Det indikerar att det finns stora skillnader på preferenser genom grupperna och att det är problematiskt att skapa en användarupplevelse som passar en stor grupp människor när man använder ett så bekant och individualiserat förfaringssätt som språk och tal.

5. Diskussion

Syftet med detta examensarbete var att undersöka och analysera diskussionen inom forskningen och den diskrepans som observerats i resultaten från den vetenskapliga litteraturen, med fokus på om det är önskvärt med mänskligt beteende i chatbotar. Detta för att ge indikationer på vilka faktorer som bidrar till en positiv användarupplevelse i chatbotar i förhållande till hur människolika de uppfattas, och formulera en designriktlinje som kan stötta framtida arbete med design av chatbotar. Frågeställningen som vägledde arbetet lyder som följer; *Vilka faktorer som gör att chatbotar upplevs som människolika kan användas för att skapa en positiv användarupplevelse vid interaktion mellan användare och chatbot?* För att besvara frågeställningen genomfördes en litteratursökning för att identifiera faktorer och sedan tillämpades en enkät för att undersöka hur dessa faktorer påverkar deltagarnas upplevelse av chatbotar. Undersökningen baserades på fem faktorer, nämligen; *småpratande, uppträdande/sätt, sofistikerade ordval & välkonstruerade meningar, dynamisk i dialogen (ställer följdfrågor)* och *visar passande känslor (i sammanhanget)*.

5.1. Faktorer kopplade till människolikhet

I följande avsnitt diskuteras resultaten från undersökningen i förhållande till frågeställningen och studier från litteraturgenomgången.

5.1.1. Småpratande

I vissa fall har användare förväntat sig att chatboten ska engagera sig med småprat, speciellt vid inledande och avslutande hälsningar (Jain et al., 2018). Portela och Granell-Canut (2017) såg också i sin studie att småpratande leder till ett ökat förtroende hos användaren. Dock finns en skillnad mellan människor som är introverta och extroverta, där extroverta känner ett större förtroende för chatboten om den småpratade, medan

introverta inte upplever någon skillnad i sitt förtroende mot en chatbot (Bickmore & Cassell, 2005).

Resultatet från studien i detta examensarbete visar på att småpratande inte nödvändigtvis är önskvärt då så många som 40% av deltagarna inte ville att chatboten skulle tilltala dem innan de har startat en konversation. Dessutom syns en övervägande preferens för de svarsalternativ där chatboten formulerar sig kort, koncist, sakligt och inom vad som är relevant för uppgiften. Slutsatsen från denna studie är alltså att denna faktor som indikerar mänskligt beteende inte är nödvändig för att skapa en positiv användarupplevelse vid interaktion med en chatbot.

Hur man pratar med en chatbot eller andra virtuella assistenter kan påverkas av flera olika faktorer. Som nämnt tidigare använder fler och fler människor olika meddelandeappar, och både sms och sociala medier har påverkat stort hur människor kommunicerar genom skriftspråk (Cougnon, Maskens, Roekhaut, & Fairon, 2017). Denna typ av beteendeförändringar är ett exempel på externa faktorer som har påverkat och som med stor sannolikhet kommer fortsätta påverka hur människor interagerar med varandra, men också med framtida teknik.

Allteftersom utvecklingen av virtuella assistenter och annan AI fortsätter, kommer det också finnas fler möjligheter för människor att komma i kontakt med denna typen av system. I vissa fall använder man system utan att vara helt medveten om det (som tidigare exempel där man använder en email där inkorgen automatiskt rensas från skräppost), men med andra system bjuder man istället väldigt medvetet in en ny typ av teknik i sitt hem. Google Home och Amazons Alexa är bra exempel på detta. I studien i detta examensarbete visar resultatet att användare inte föredrar att chatbotar småpratar. I de fall där en användare bjuder in virtuella assistenter till sitt hem skulle det kunna innebära en skillnad. Dels har man medvetet bjudit in assistenten i sitt hem, vilket skulle kunna innebära en annan attityd mot assistenten. Det finns också en annan skillnad som kan påverka på vilket sätt en användare vill prata med sin virtuella assistent. En chatbot eller en virtuell assistent som existerar i ett specifikt sammanhang, där de till exempel ska hjälpa en användare att uppnå ett mål inom en specifik tjänst eller hitta rätt på en specifik hemsida, agerar som ett hjälpmedel för att få hjälp med en uppgift. Användaren har dock i första hand sökt upp tjänsten eller produkten, och det är där fokuset ligger. Alltså är tjänsten det viktiga, och användaren stöter på chatboten eller den virtuella assistenten någonstans under resans gång. I fallet med, till exempel, Google Home och Amazons Alexa är fokuset istället omvänt. Användaren har skaffat sig en virtuell assistent i första hand. I många fall görs detta för att en användare vill få hjälp med olika uppgifter, men det är inte en specifik uppgift som har föranlett den första interaktionen med assistenten. Detta skulle kunna innebära att användaren, på grund av den större investeringen och intentionen med införskaffandet av assistenten, i förlängningen vill lägga mer tid på att skapa en relation med assistenten, att den ska lära känna användaren för att på så sätt optimera interaktionen och den gemensamma förståelsen. Utifrån olika typer av miljö och sociala sammanhang skulle man alltså kunna utforska om önskan att småprata med en virtuell assistent, eller chatbot, är starkare. Beroende på användningsområde kan det också finnas en skillnad i användarens preferenser. Luger och Sellen (2016) har sett en

skillnad vad gäller chatbotar, där fokus ligger på att lösa en uppgift, och konversationsagenter (conversational agents) där fokus ligger på att samtala med en AI.

5.1.2. Uppträdande/sätt

Vissa tidigare studier har argumenterat för att en chatbot ska kunna uppvisa olika beteende och sätt i olika situationer, beroende på om sammanhanget är mer officiellt eller om det handlar om underhållning (Jain et al., 2018), samt att det är positivt om chatboten har en personlighet (Jain et al., 2018; Liao et al., 2018; Mone, 2016; Skjuve et al., 2019). Humor är en annan faktor som har setts kunna påverka användarens relation till chatbotar på ett positivt sätt (Jain et al., 2018), men det finns också argument för att det är riskabelt att inkludera humor då det skapar höga förväntningar som inte kan infrias (Luger & Sellen, 2016).

Resultaten från studien i detta examensarbete visar på att det är viktigare att chatboten är tydlig och transparent än att den har en egen personlighet eller vill underhålla. När chatboten fick en möjlighet att uttrycka åsikter eller känslor som inte var relevanta för uppgiften var dessa svar i många fall de minst populära bland deltagarna. I relation till denna faktor finns det en diskrepans i förhållande till tidigare studier i litteraturgenomgången ovan. Det kan inte fastställas att ett människoliknande beteende är önskvärt men det kan inte heller uteslutas att det är önskvärt att en chatbot har en konsekvent och autentisk personlighet (Jain et al., 2018; Neururer et al., 2018). Det kan också vara värdefullt att inte endast titta på beteenden mellan människor när man vill designa för interaktion mellan människa och maskin, då det finns studier som pekat på att människor inte pratar på samma sätt med en chatbot som med andra människor (McDuff & Czerwinski, 2018; Mou & Xu, 2017; Luger & Sellen, 2016; Demeure et al., 2011). Mou & Xu (2017) såg bland annat att människor var mer öppna, extroverta och samvetsgranna när de talade med andra människor, gentemot i de fall där de pratade med en chatbot.

Som nämnt tidigare skulle en väg framåt mot positiva användarupplevelser vid interaktion med chatbotar kunna vara att utforska hur man kan designa chatbotar som anpassar sitt beteende efter varje specifik användare. Park et al. (2015) har i en studie skapat en modell som genom att bedöma vad och hur människor skriver på sociala medier kan förutspå och mäta personlighetsdrag. Genom att använda en liknande modell skulle också en chatbot kunna analysera användarens språk för att läsa av deras personlighet, och på så sätt anpassa sitt eget beteende efter användarens preferenser.

Personlighet och chatbotars uppträdande diskuteras vidare i avsnitt 5.2 *Personlighet och "den kusliga dalen"*.

5.1.3. Sofistikerade ordval och välkonstruerade meningar

Sofistikerade ordval och välkonstruerade meningar är faktorer som användare uppskattar i chatbotar, och som även påverkar hur människolika de uppfattas (Lortie & Guitton, 2011).

Resultaten från studien i detta examensarbete stöttar resultaten från tidigare studier, då de visar att deltagarna föredrar de svar som är välformulerade och sofistikerade. Även

i de fall där det innebar att svaren blir längre kunde man se en preferens för de meningar som var välkonstruerade. Det finns en antydning till att de svar som deltagarna föredrog även var de svar som liknar de svar man kan förvänta sig från en människa i en officiell roll eller situation. Dessa svar har ett korrekt språkbruk. Deltagarna föredrog inte vardagsspråk med förkortningar och slang som används vid till exempel sms eller andra digitala meddelandeformer. Sammanfattningsvis uppfattade deltagarna denna faktor positivt, det är däremot inte lika tydligt om själva faktorn faktiskt visar upp ett generellt mänskligt beteende eller om det är ett beteende som deltagarna känner igen mer från formella sammanhang eller andra situationer när de förväntar sig att få hjälp av en expert eller person inom ett serviceyrke.

I många digitala medier, som sms och chattar, används alltså ett språkbruk där det bland annat förekommer mycket förkortningar och inkorrekt punktuation. Detta språkbruk kan då argumenteras vara en faktor som borde visa på människoliknande beteende, men det var inte det som deltagarna föredrog i denna studie. I förlängningen kan det innebära att användare föredrar sofistikerade ordval och välkonstruerade meningar, men att det faktiskt inte är på grund av att det påminner dem om mänskligt beteende, utan snarare tvärtom (eller i alla fall en viss typ av mänskligt beteende). En förklaring till detta skulle kunna vara att människor inte agerar på samma sätt i sin interaktion med en chatbot som de gör med andra människor (Luger & Sellen, 2016; McDuff & Czerwinski, 2018; Mou & Xu, 2017; Demeure et al., 2011).

5.1.4. Dynamisk dialog

Att ha en dynamisk dialog innebär inte bara att kunna ställa relevanta följdfrågor, men också att kunna hålla kontext genom en dialog och kunna komma ihåg och koppla ihop information som sagts tidigare med nya uttalanden eller frågor (Jain et al., 2018; Neururer et al., 2018; Schuetzler et al., 2014). I studien från detta examensarbete finns det inga tecken som visar på att det är önskvärt att chatboten kan hålla information från flera månader tillbaka i minnet. En hypotes man kan dra från detta är att det är önskvärt att chatboten kan hålla kontext på samma nivå som en annan människa man interagerar med, men inte som en entitet som har koll på allt man gör till en sådan grad att det blir obehagligt och man känner sig övervakad. Dock behövs det mer data för att kunna verifiera eller falsifiera denna hypotes.

Främst undersöktes i detta examensarbete faktorn *dynamisk dialog* genom att analysera till vilken grad deltagarna uppskattade relevanta följdfrågor. Resultaten visar att det är önskvärt att chatboten håller en dynamisk dialog. Däremot är det inte lika tydligt om det spelar någon roll ifall detta görs genom att chatboten ställer följdfrågor eller om det fungerar lika bra att chatboten tydliggör om den har förstått eller inte på sättet den svarar. Detta kan till exempel ske genom att chatboten upprepar frågan i svaret för att säkerställa att användaren och chatboten har en gemensam förståelse för vad som händer i dialogen, samt att parterna har förstått fråga och svar rätt.

Siri, Apples virtuella assistent, använder sig av en liknande metod. När användaren talar in sin fråga skrivs den samtidigt ut på skärmen och på så sätt blir det tydligt om Siri har förstått frågan rätt. Dessutom visar Siri resultaten på skärmen, så användaren får både

visuell feedback och svar via röst. Detta har visat sig höja förtroendefaktorn, till skillnad mot till exempel Amazons Alexa och Google Assistant där användaren får förlita sig på endast feedback från röstmeddelanden (Bickmore et al., 2018). Utifrån detta finns det etiska överväganden som behöver göras vid design av chatbotar och andra virtuella assistenter. Ett exempel, utifrån ämnet ovan och att en dynamisk dialog kan skapa förtroende, är att om faktorer som höjer förtroende för chatboten/den virtuella assistenten används vid design av desamma så behöver man också se till att chatboten verkligen äger rätt sorts kompetens och har en gedigen kunskap inom det aktuella området. Eller att den åtminstone kan avgöra om ett ämne som den inte har kunskap om berörs, och då kan kommunicera detta på ett tydligt sätt. För att visa på vilka konsekvenser detta kan ge kan vi ta ett exempel från Bickmore et al. (2018). I sin studie såg Bickmore et al. (2018) att användare, som nämnt ovan, uppvisade ett större förtroende för Apples Siri, än Alexa och Google Home. Deltagarna i studien fick fråga medicinska frågor till de olika assistenterna och efter att de fått svar fick de ange hur de skulle agerat i den fiktiva situationen, utifrån den information de fått från de virtuella assistenterna. Utifrån dessa svar räknades det ut hur stor skada deltagarna kunde ha åsamkats genom att följa råden från de virtuella assistenterna. Alexa misslyckades med uppgiften i de flesta fall (drygt 90%) men det var jämförelsevis en liten del där svaren ledde till skada för användaren (1,5%). Siri å andra sidan misslyckades endast med 22,5% av uppgifterna, däremot kunde 29,9% av fallen leda till skada och 20,9% av fallen leda till dödsfall (Bickmore et al., 2018). Att kunna designa chatbotar som bygger förtroende är en sak. Att kunna designa chatbotar med en förståelse för hur de kommer påverka människorna som använder dem på flera olika nivåer är en annan sak. Inom området för AI är detta en utmaning och det behöver inkluderas människor med kompetens inom social vetenskap och kognitionsvetenskap för att kunna förutspå och förebygga vilka implikationer den nya tekniken kan ge i de situationer där den ska användas (Hagendorff & Wezel, 2019).

5.1.5. Passande känslor för sammanhanget

När en chatbot reagerar på ett emotionellt sätt som är relevant och passande i ett visst sammanhang kan det skapa ett mer trovärdigt intryck och en uppfattning att chatboten är mer kompetent och varm i sitt bemötande (Demeure et al., 2011).

Att en chatbot har passande känslor för sammanhanget har i studien som gjorts till detta examensarbete undersökts genom hur den agerar när den inte förstår eller när dialogen misslyckas på något sätt. Resultaten visar att användare uppskattar att chatboten ber om ursäkt när något blir fel, att den är ödmjuk och visar andra vägar framåt, antingen genom att guida användaren i hur hen kan omformulera sig eller genom att erbjuda hjälp från till exempel en människa istället. Ett visst mått av emotionell respons var positivt, dock var det inte många deltagare som föredrog svar med stora känslorvall i. Resultaten överensstämmer med tidigare studier, då de reaktioner som var mest populära i studien från detta examensarbete också är de som kan argumenteras vara passande för sammanhanget. Chatboten ska fungera som en assistent i en servicesituation och då är det rimligt att den inte agerar alltför emotionellt.

5.2. Personlighet och ”den kusliga dalen”

Ett flertal tidigare studier har undersökt på vilka sätt man kan avgöra hur människolika chatbotar och andra virtuella agenter är i sitt beteende och hur man kan skapa människoliknande chatbotar (Bickmore & Cassell, 2005; Lortie & Guitton, 2011; Mone, 2016; Neururer et al., 2018; Paikari & van der Hoek, 2018; Schuetzler et al., 2014; Skjuve et al., 2019). Det behövs dock mer forskning kring om detta faktiskt är önskvärt för att skapa en positiv användarupplevelse (Neururer et al., 2018; Piccolo et al., 2018; Skjuve et al., 2019).

Skjuve et al. (2019) menar att chatbotar inte har kommit tillräckligt långt i sin utveckling för att de ska kunna vara så människoliknande att de frammanar obehagskänslor, likt ”den kusliga dalen”. Å andra sidan finns det exempel på chatbotar som har skapat obehag eller olustkänslor hos användare på grund av att de antingen har betett sig för mycket som en maskin (Mone, 2016), eller uppvisat emotionella uttryck som inte stämmer överens med de uttryck som användare förväntar sig utifrån deras förståelse för vad chatboten är kapabel till (McDuff & Czerwinski, 2018). I ovan nämnda studier finns det indikationer som visar på att det redan nu kan uppstå obehagskänslor då människor har lätt för att känna av om något inte riktigt stämmer, och reagera negativt i sådana situationer (McDuff & Czerwinski, 2018). I teorin om ”den kusliga dalen” menar Mori (1970, 2012) att robotar som är för människoliknande skapar ett obehag. Det finns indikationer som visar på att det eventuellt kan uppstå liknande reaktioner om chatbotar är för människolika i sitt beteende (McDuff & Czerwinski, 2018; Mone, 2016; Strait et al., 2015). Det kan bero på att användare skapar höga förväntningar vid interaktion med en chatbot som simulerar en personlighet, vilket gör att användaren tror att chatboten ska kunna hantera olika situationer på ett, i förhållande till människor, mer naturligt beteende (Luger & Sellen, 2016). När sedan chatboten inte reagerar eller beter sig enligt konventionella mönster som människor är vana vid från dialoger med andra människor, upplevs detta som att något inte stämmer (McDuff & Czerwinski, 2018).

När det kommer till chatbotar, och annan AI, finns det heller inget som säger att personlighet behöver innebära att chatboten ska ha en personlighet som är lik en människas (Mone, 2016; Piccolo et al., 2018). Det behöver utredas och artikuleras tydligare vad *personlighet* egentligen betyder i chatbotsammanhang. För att undvika att chatbotar frammanar ”den kusliga dalen” kan en lösning vara att skapa chatbotar som inte har beteenden eller personligheter som är för människolika. I tidigare studier har man kunnat se att användare drar sig till och uppskattar chatbotar med personlighet (Jain et al., 2018; Liao et al., 2018; Mone, 2016). Det finns dock en risk att automatiskt anta att det innebär att chatbotar med personligheter som speglar människolika beteenden är det som är önskvärt (McDuff & Czerwinski, 2018; Mone, 2016). Beroende på sammanhang kan det vara positivt att chatbotar har olika beteenden eller personligheter (Jain et al., 2018).

Att användare skulle dra sig till chatbotar som har en personlighet är inget som resultaten från studien i detta examensarbete har kunnat bekräfta. Resultaten har istället pekat på att deltagarna inte uppskattar när en chatbot uppvisar prov på humor eller uttrycker sig som om den har egna tankar om ämnet i fråga. I de fall där chatboten avvek

det minsta från det som kunde anses vara mest relevant, att lösa uppgiften eller besvara frågan från användaren, kan man se att det endast var ett fåtal deltagare som uppskattade den interaktionen. Dock har alla chatbotar i de frågor som är inkluderade i denna studie någon slags uppgift att assistera användaren med, till exempel att hitta resealternativ eller spela upp musik. Om man till exempel istället tittar på en chatbot som skapas för att agera som ett socialt stöd eller kompis kan det bli mer aktuellt och även önskvärt att chatboten har en utmärkande personlighet som kan göra att den uppfattas mer människolik, eller åtminstone att den uttrycker sig på ett sätt som särskiljer den från andra chatbotar, så att den kan argumenteras ha en egen personlighet. Då kommer vi dock behöva återvända till frågan om en för människolik chatbot kan upplevas som obehaglig. Att veta om och hur chatbotens personlighet kan frammana en "kuslig dal" kommer vara viktigt för att förstå hur chatbotens personlighet kan byggas upp. I dessa fall kan det alltså vara ännu viktigare att artikulera tydligt vad en personlighet innebär hos en chatbot. I och med att resultaten från studien i detta examensarbete visar på en så pass spretig bild av deltagarnas preferenser är det svårt att specificera specifika mänskliga faktorer som kan bidra till en positiv användarupplevelse. Om man vill designa chatbotar som tilltalar en bred grupp användare behöver man snarare börja med att utforska hur man kan designa en chatbot som kan anpassa sitt beteende efter den användare som den interagerar med i stunden.

Skjuve et al. (2019) argumenterar att chatbotar inte har kommit såpass långt i utvecklingen att de kan misstas för människor, och därför är det inte heller aktuellt att de kan frammana "den kusliga dalen" och skapa obehag. Frågan är om det går att utgå endast utifrån den tekniska utvecklingen och kompetensen i dagsläget, eller om det inte i själva verket är så att förväntningarna och den kompetens som användare tror finns i chatbotarna som spelar störst roll. Om användare tror att chatbotar har lika hög kompetens och intelligens som människor (de höga förväntningar många användare har på chatbotar indikerar att de ofta tror att chatbotar har en hög intelligens (Jain et al., 2018; Luger & Sellen, 2016) så blir resultatet detsamma. Användaren kan känna obehag eller att något är fel när chatboten inte beter sig enligt förväntningarna.

5.3. Begränsningar i studien

Studien har genomförts i en begränsad miljö och en stor del av kanalerna som enkäten skickades ut till innefattade personer med relativt stor vana av internet och teknik, även om man kan se i resultaten att det har olika stor erfarenhet av chatbotar. Det skulle vara intressant att göra en studie som tydligare undersöker skillnaderna i uppfattningen av chatbotarna i relation till deras vana och förståelse av teknik överlag. De användare som är mer teknikvana och har en djupare förståelse för hur chatbotar fungerar rent tekniskt har också större tålamod och mer realistiska förväntningar på chatbotarnas kompetens än användare som inte har samma förståelse (Luger & Sellen, 2016). Det hade varit intressant att fortsätta studien för att se om det går att se relationer mellan denna förståelse och på vilket sätt användaren uppskattar människoliknande beteende i chatbotar.

Frågorna som ställs i enkäten är upplagda efter exempel på situationer. Detta för att försöka få en förståelse för hur folk uppfattar de olika faktorer som påverkar en chatbots beteende och användarupplevelsen, utan att faktiskt säga rakt ut vilka faktorer det är som

undersöks. Valet gjordes för att inte påverka eller styra användaren att tänka utanför de presenterade scenarierna, då detta skulle kunnat påverka genom att de tolkar faktorerna på olika sätt. Däremot finns det även i detta fall rum för tolkningar. För att kunna avläsa datan mer korrekt rekommenderas en ytterligare studie som designas tvärtom, det vill säga med ett direkt fokus på de olika faktorerna och hur användaren uppfattar och tolkar dem i kontexten kring chatbotar.

5.4. Implikationer för UX-design

När användandet av chatbotar fortsätter att spridas kommer det skapa nya utmaningar för designers. Tidigare har det varit vanligt att designa grafiska användargränssnitt, nu börjar istället design av konversationer att begära ett större fokus från designers (Luger & Sellen, 2016). När gränssnitt som styrs med hjälp av natural language blir vanligare kommer större fokus behöva läggas på design av konversationer, kanske till och med i högre grad än på design av grafiska gränssnitt (Følstad & Brandtzaeg, 2017). En gren inom UX-design som fokuserar på detta är *konversationsdesign* (eng. conversational design eller conversational informatics) (Nishida, 2015). Konversationsdesign syftar till att undersöka och designa interaktion som sker via språkliga konversationer, vilket även kan inkludera sociala interaktioner mellan användare och deras omgivning (Nishida, 2015). Utvecklingen från grafiska gränssnitt till röststyrda gränssnitt av röst innebär också att designers kan behöva ändra designperspektiv. Istället för att designen har en *förklarande* uppgift går trenden till att den får en mer *tolkande* uppgift. Det kommer bli ännu viktigare för designers att verkligen förstå användarens process, intention och mål, så att man kan hjälpa dem att komma åt innehåll via konversationer istället för grafiska användargränssnitt (Følstad & Brandtzaeg, 2017).

I övergången från grafiska till röststyrda och andra “naturliga” användargränssnitt (som exempelvis de som implementeras i chatbotar och konversationsdesign) kommer det med största sannolikhet att bli mer komplext att designa en användarupplevelse som är positiv för en bred grupp användare. En användare som sitter framför en hemsida eller ett program har till exempel ett visst mått av valfrihet samtidigt som hen är låst till att ta sig fram genom flödet via de funktioner som finns tillgängliga (såsom knappar, inputfält, länkar och dylikt). Detta leder till att en stor del av användarna sannolikt kommer fram till sitt mål på liknande sätt som de har fått lära sig för att kunna just navigera i den digitala världen. Det finns helt enkelt inte tillräckligt många olika alternativ att klara av uppgiften på. Även om olika användare kommer att hantera uppgiften utifrån sina personliga synsätt så finns det begränsningar i form av det aktuella grafiska gränssnittet. Om användaren däremot via tal och skrift interagerar med ett gränssnitt på ett sätt som liknar en mänsklig konversation har användaren en helt annan referensram från det verkliga livet att reagera på. Om hen säger eller frågar något till en chatbot och responsen blir helt skild från det användaren skulle kunna förvänta sig om hen frågade samma sak till en annan människa så uppstår en krock och det finns större risk att användaren upplever att interaktionen inte fungerar som den ska. Om interaktionen inte är autentisk kan den lätt uppfattas som falsk (Neururer et al., 2018). Människor har också mycket olika preferenser kring hur man vill samtala med andra människor och vad man förväntar sig, vilket i förlängningen innebär

att det antagligen finns lika många olika förväntningar på hur en interaktion med en chatbot bör se ut som det finns människor (Lortie & Guitton, 2011). Chatbotar öppnar upp för att användaren ska kunna interagera med teknik på ett mer personligt och individualiserat sätt, vilket också skapar fler olika sorters förväntningar och möjligtvis också fler vägar att nå samma mål. Dessa förväntningar och olika vägar mot målen behöver förutses och förberedas för redan i designen av chatboten om den ska kunna tilltala en bredare grupp användare. Alternativen kan vara att antingen skapa chatbotar som kan lära känna användaren och anpassa sig efter hans preferenser för interaktion, dialog och samtalston (McDuff & Czerwinski, 2018) eller att vara transparent och tydlig med vilka åtgärder som chatboten klarar av att reagera på, hur man pratar med den, eller till och med endast ge förskrivna alternativ som sedan användaren får välja mellan genom att trycka på knappar eller liknande (Jain et al., 2018). Det senare alternativet distanserar sig dock från konversationsdesign och närmar sig istället ett mer traditionellt grafiskt användargränssnitt.

5.5. Konsekvenser för individer och samhället

Att skapa positiva användarupplevelser i konversationer med AI är inte bara relevant för chatbotar. AI växer och integreras i fler produkter och områden i samhället (Hecht, 2018; McNeal & Newyear, 2013). Detta är en utveckling som med största sannolikhet kommer fortsätta framöver, då ett av målen med AI är att ersätta, komplettera och förstärka i princip alla uppgifter som människor utför idag (Makridakis, 2017). Om detta blir fallet, kan design av interaktionen mellan användare och AI vara avgörande både för att användare ska kunna förstå hur de interagerar med olika AI-system, och för att de ska *vilja* interagera med dessa system, produkter och tjänster.

Ju smartare produkter omkring oss blir, desto mer kommer det finnas behov eller önsknings om att kunna kommunicera med dessa produkter på ett smidigt sätt. Att sätta en skärm på till exempel robotdammsugaren tar plats och kräver att användare sätter sig in i och lär sig hur de olika funktionerna ska hanteras. Om det istället går att kommunicera med robotdammsugaren via tal reducerar detta inlärningskurvan och skapar andra förutsättningar för hur användare kan interagera med produkter i sin omgivning. Det finns inte begränsningar i utrymme på samma sätt som det finns om man ska lägga upp funktioner på en skärm och mer komplexa uppgifter kan anropas på ett intuitivt sätt. Dock är det viktigt att skapa gränssnitt med ett fokus på tillgänglighet så att människor med fysiska eller sensoriska begränsningar inte exkluderas från att använda dessa tjänster och produkter.

Hur man interagerar med chatbotar kan överföras till hur interaktion med annan AI kan se ut, vilket gör resultaten av denna studie relevant för flera områden. Funktioner som kan hanteras via röst öppnar upp för möjligheter att kontrollera smarta funktioner runt omkring oss på ett intuitivt sätt. Att skapa en positiv användarupplevelse av en interaktion som sker via dialoger med AI är därför intressant för olika områden och branscher, både övergripande och på detaljnivå. Till exempel kan man redan se att AI hjälper kunder att beställa mat via skärmar på snabbmatsrestauranger. Om denna interaktion kunde hanteras via röst istället skulle man kunna undvika spridandet av bakterier som samlas på dessa

ytor. På en annan abstraktionsnivå kan AI hjälpa till att analysera stora mängder data och på så sätt avlasta monotona uppgifter inom olika yrkesroller och därigenom öppna upp för att människor istället kan ägna sig åt mer kreativa och tillfredsställande uppgifter. Om smarta produkter runt omkring oss också kan tala med oss öppnar det upp för att designa nya användarupplevelser för användarna. Att få hjälp med grundläggande uppgifter i vardagen via kommunikation med virtuella assistenter kan också hjälpa människor att frilägga tid för viktigare och roligare aktiviteter. Att bli avlastad kan i förlängningen innebära mindre stress i vardagen.

Några av de trender som spås dominera inom AI under 2019 är bland annat *autonomisering*, *robotisk processautomation (RPA)* och *NLP/AI-assistenter* (Bisnode, 2019). Gemensamt för dessa tre trender är att de innebär att AI kan automatisera både mindre och större delar av det arbete som människan genomför idag. I första hand kan AI med mekanisk intelligens ta över monotona och repetitiva uppgifter (Huang & Rust, 2017) vilket öppnar upp för att människor istället kan frigöras för mer kreativa och komplexa uppgifter. Om det blir vanligare att AI tar över delar av våra arbetsuppgifter kommer det också finnas behov av att kunna samarbeta och interagera med dessa system på ett bra sätt. För att detta ska fungera behöver de ha användarvänliga gränssnitt. En positiv användarupplevelse, som också ger användaren en förståelse för vad systemet klarar av att göra, kommer innebära att människor faktiskt använder systemet eller produkten. Först då kommer deras fulla potential att utnyttjas. Om interaktionen med dessa AI och andra virtuella assistenter blir vanligare så behöver det utredas vilka aspekter av en sådan interaktion som kan upplevas som obehagliga, och användare behöver kunna lita på att de kan hantera dialog och interaktion på ett säkert sätt. Redan idag existerar många assistenter och tjänster som använder AI (Siri, självkörande bilar, chatbotar som kundtjänst m.m.) och det förutspås att AI kommer att ta över fler och fler uppgifter och jobb från människan framöver (Bisnode, 2019; Huang & Rust, 2017; Makridakis, 2017). I förlängningen skulle design av konversation och kommunikation med AI kunna stötta och förändra människors syn på AI, samt hur människor känner inför att interagera och lämna över kontroll till dessa system och robotar. En av utmaningarna inom fältet för AI är just att deras framgång är knuten till hur accepterade de är i samhället, vilket innebär att de bara kommer användas i de fall där människor känner sig bekväma med att byta ut människor mot AI (Hagendorff & Wezel, 2019). Kanske kan design av AI:ns personlighet och de konversationer som kan föras med den även lugna människors oro och rädsla för vad som ska hända när tekniken fortsätter göra framsteg. Om detta är positivt för att det möjliggör ett givande samarbete mellan människa och maskin eller om det öppnar upp för att göra AI smartare och mer förmögna att manipulera och kontrollera människan är inte något som denna studie lägger något värde i. De olika möjligheterna visar dock på vikten av att designers behöver fortsätta vara medvetna om de designbeslut de tar och vilka konsekvenser dessa beslut kan få i det långa loppet.

6. Slutsatser och framtida forskning

Syftet med detta examensarbete var att undersöka och analysera diskussionen inom forskningen och den diskrepans som observerats i resultaten från den vetenskapliga

litteraturen, med fokus på om det är önskvärt med mänskligt beteende i chatbotar. Frågeställningen som vägledde arbetet lyder som följer; *Vilka faktorer som gör att chatbotar upplevs som människolika kan användas för att skapa en positiv användarupplevelse vid interaktion mellan användare och chatbot?*

De faktorer som i resultatet från denna undersökning helt eller delvis upplevdes positiva av deltagarna var:

- *Sofistikerade ordval och välkonstruerade meningar.* Det var tydligt att deltagarna överlag föredrog svar som byggdes upp med sofistikerade ordval och välkonstruerade meningar, även i de fall där det innebar att svaren blev längre.
- *Dynamisk dialog.* Denna faktor har i denna undersökning främst undersökts genom att se till vilken grad deltagarna uppskattade *relevanta följdfrågor*. Deltagarna har visat en preferens för svar där chatboten är tydlig med vad den svarar på, antingen genom att den ställer relevanta följdfrågor eller att den upprepar frågan när den ger sitt svar.
- *Vad gäller faktorn Uppträdande/sätt* så kan man i resultaten från studien i detta examensarbete se att de flesta deltagare föredrog de svar där chatboten var tydlig, koncis och svarade med en officiell ton. Resultaten visar på att det är viktigare att chatboten är tydlig och transparent än att den försöker ha en egen personlighet eller underhålla användaren. Alltså kan man här se att ett *officiellt/formellt* uppträdande/sätt sågs som positivt i interaktionen med chatbotar.
- *Passande känslor för sammanhanget.* När interaktionen mellan chatbot och användare misslyckades på något sätt, uppskattar deltagarna ett svar där chatboten ber om ursäkt för besväret och ger bud på hur användaren kan gå vidare för att förhoppningsvis få ett bättre resultat. Ett visst mått av emotionell intelligens verkar vara önskvärt, dock är det inte populärt om chatboten drar ut på ursäkterna och uppvisar stora känslor. Detta indikerar att faktorn *Passande känslor för sammanhanget* upplevs som positiv av deltagarna.

De slutsatser som kan dras från denna studie är följande; (1) Det finns en stor bredd i människors preferenser kring hur de vill interagera med chatbotar och AI; (2) chatbotar är inte uteslutna från att kunna frammana “den kusliga dalen”, där de skapar ett obehag eller en olustkänsla i användaren vid interaktion; (3) det behöver artikuleras vilka faktorer som kan och bör bygga upp en chatbots personlighet och vad som särskiljer dessa från en mänsklig personlighet; (4) allteftersom utvecklingen rör sig bort från grafiska användargränssnitt mot konversationsdesign kan detta innebära nya utmaningar och en ny approach för designers, där de fäster större vikt vid att designa dialoger, konversationer och personligheter för att skapa en positiv användarupplevelse.

Ytterligare en faktor som kan påverka om mänskligt beteende upplevs som positivt i interaktionen mellan chatbot och människa är teorin om “den kusliga dalen”. För att tydliggöra huruvida teorin om “den kusliga dalen” kan appliceras på chatbotar kan framtida forskning fokusera på om det är chatbotens faktiska intelligens och

kompetens som avgör om det uppstår en "kuslig dal" eller om det i själva verket är användarnas *förväntningar på* eller *upplevelse av* chatbotens intelligens och kompetens som är avgörande för om man kan se en koppling till teorin.

Som nämnts i tidigare stycken existerar det i näringsliv och i vetenskapliga studier diskussioner som behandlar hur chatbotar kan utvecklas för att närma sig mänskliga konversationer eller göras mer människolika i sitt framförande (Akma et al., 2018; Jain et al., 2018; Luger & Sellen, 2016; McDuff & Czerwinski, 2018; Mone, 2016; Neururer et al., 2018; Piccolo et al., 2018; Skjuve et al., 2019; Westerman et al., 2018), dock finns det ingen konsensus om detta är rätt väg framåt i förhållande till användarnas upplevelse av interaktionen med chatbotar (Neururer et al., 2018; Piccolo et al., 2018; Skjuve et al., 2019). Slutsatserna som har kunnat dras utifrån resultaten i detta examensarbete visar på att det inte är självklart att det är positivt med ett människoliknande beteende i en chatbot.

Framtida forskning bör vidare undersöka huruvida användarupplevelsen påverkas positivt eller negativt vid interaktion med människoliknande chatbotar samt på vilket sätt sammanhanget kan vara avgörande för hur chatboten ska bete sig. Det är speciellt intressant att hitta faktorer som hjälper till att artikulera hur en chatbots personlighet kan och bör särskilja sig från en mänsklig personlighet, för att optimera användarupplevelsen och för att inte frammana "den kusliga dalen". Dessutom är det viktigt att ha i åtanke att det inte nödvändigtvis är svart och vitt om det är positivt med ett människoliknande beteende i chatbotar. I relation till just design av positiva användarupplevelser mellan chatbot och människa kanske frågan snarare behöver vara; *Hur kan en chatbot anpassa sitt beteende efter användarens preferenser för att skapa en positiv användarupplevelse?* Människor har personliga preferenser kring hur de vill att en chatbot ska bete sig och hur de vill tala med den och bli tilltalad tillbaka. Därför borde man i framtida forskning undersöka djupare hur man kan skapa chatbotar som kan personaliseras och som användare med olika preferenser kan interagera med på sina villkor och efter sina egna preferenser. Samt, från ett mer övergripande perspektiv; *Bör en chatbot/AI efterlikna människoliknande beteenden eller kan det vara bättre, för att i framtiden skapa en harmonisk relation med AI, att det förblir tydligt att de är maskiner och att de inte själva kan agera på människoliknande sätt?* Den senare frågan är komplex och mångfacetterad. Om man vill att AI ska kunna hjälpa människor i mer komplicerade sammanhang kan det också krävas att de förstår mänskligt beteende för att kunna fortsätta att utvecklas och hjälpa människor med dessa komplicerade uppgifter. Till exempel i medicinska sammanhang, där de skulle kunna läsa av en människas olika signaler för att avgöra vilket tillstånd hen är i och på vilket sätt hen behöver hjälp. Men hur stor förståelse kan AI få av en människa och mänskligt beteende, utan att faktiskt kunna börja "känna" av dessa känslor eller inkorporera liknande beteenden i sin egen existens? Innebär en fullständig förståelse av människan att AI kan utveckla ett medvetande (eng. consciousness)? Om AI ska kunna ta över mer komplexa arbetsuppgifter från människan kan det kräva att de har ett medvetande och även ett samvete (Meissner, 2019). Å andra sidan skulle medvetna AI kunna bli svåra att kontrollera och eventuellt skulle de också börja utföra uppgifter på ett sätt som går mer i linje med deras mål och önskningar, snarare än att det skulle gynna människan.

I det långa loppet är det viktigt att undersöka hur konversationsdesign och design av en chatbots personlighet kan påverka människans framtida relation till AI, samt om det är klokast att utgå ifrån en optimistisk syn på AI och utveckla det så långt det går, med en utopi i målsiktet, eller om det är säkrast att inta ett mer pessimistiskt och dystopiskt förhållningssätt så att vi inte hamnar i en framtid där superintelligent AI styr över människosläktet.

7. Källförteckning

- AbuShawar, B., & Atwell, E. (2016). Usefulness, localizability, humanness, and language-benefit: additional evaluation criteria for natural language dialogue systems. *International Journal of Speech Technology*, 16(2), 373-383.
- Akma, N., Hafiz, M., Zainal, A., Fairuz, M., & Adnan, Z. (2018). Review of Chatbots Design Techniques. *International Journal of Computer Applications*, 181(8), 7-10.
- Bickmore, T., & Cassell, J. (2005). Social Dialogue with Embodied Conversational Agents. In J. C. J. van Kuppevelt, L. Dybkjær, & N. O. Bernsen (Eds.), *Advances in Natural Multimodal Dialogue Systems* (pp. 23-54). Dordrecht: Springer Netherlands.
- Bickmore, T., Trinh, H., Olafsson, S., O'Leary, T., Asadi, R., Rickles, N., & Cruz, R. (2018). Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An observational study of Siri, Alexa, and Google Assistant. *Journal of Medical Internet Research*, 20(9), e11510 (11513 pages).
- Bisnode. (2019). *9 AI-trender som boomar 2019*. Retrieved from Solna:
- Bostrom, N. (2017). *Superintelligens: Vägar, faror, strategier*. Lidingö: Fri Tanke förlag.
- Braga, A., & Logan, R. (2017). The Emperor of Strong AI Has No Clothes: Limits to Artificial Intelligence. *Information*, 8(4), 156-177.
- Cancel, D. (2019). *Conversational Marketing and Sales - how to grow leads, shorten sales cycles, and improve your customer's experience with real-time conversations*: John Wiley and Sons, Incorporated.
- Candello, H., Pinhanez, C., & Figueiredo, F. (2017). *Typefaces and the Perception of Humanness in Natural Language Chatbots*. Paper presented at the CHI 2017, 6-11 maj, Denver, CO, USA.
- Chakrabarti, C., & Luger, G. F. (2015). Artificial conversations for customer service chatter bots: Architecture, algorithms and evaluation metrics. *Expert Systems with Applications*, 42(2015), 6878-6897.
- Cougnon, L.-A., Maskens, L., Roekhaut, S., & Fairon, C. (2017). Social media, spontaneous writing and dictation, Spelling variation. *Journal of French Language Studies*, 27(2017), 309-327.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817.
- Demeure, V., Niewiadomski, R., & Pelachaud, C. (2011). How is believability of a virtual agent related to warmth, competence, personification, and embodiment? *Presence*, 20(5), 431-448.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm Aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126.
- Dix, A. (2010). Human-computer interaction: A stable discipline, a nascent science, and the growth of the long tail. *Interacting with Computers*, 22(1), 13-27.
- Dourish, P. (2004). What we talk about when we talk about context. *Personal Ubiquitous Computing*, 8(1), 19-30.
- Faraon, M. (2018). *Co-creating democracy: Conceptualizing co-creative media to facilitate democratic engagement in society*. Stockholms Universitet, Stockholm.

- Følstad, A., & Brandtzaeg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38-42.
- Gams, M., Gu, I. Y., Härmä, A., Muñoz, A., & Tam, V. (2019). Artificial intelligence and ambient intelligence. *Journal of Ambient Intelligence and Smart Environments*, 11(1), 71-86.
- Govers, F. X. (2018). *Artificial Intelligence for Robotics*.
- Hagendorff, T., & Wezel, K. (2019). 15 challenges for AI: or what AI (currently) can't do. *AI & Society*, 3(2019), 1-11. doi:<https://doi-org.ezproxy.hkr.se/10.1007/s00146-019-00886-y>
- Hecht, J. (2018). Managing expectations of artificial intelligence. *Nature Outlook - Digital Revolution*, 563(7733), 141-143.
- Huang, M., & Rust, R. (2017). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.
- Jain, M., Kumar, P., Kota, R., & Patel, S. N. (2018). *Evaluating and informing the design of chatbots*. Paper presented at the DIS 2018, Session 18: Interacting with Conversational Agents, 9-13 juni, Hong Kong.
- Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). *The rise of bots: A survey of conversational interfaces, patterns and paradigms*. Paper presented at the DIS 2017, 10-14 juni, Edinburgh, United Kingdom.
- Liao, Q. V., Hussain, M., Chandar, P., Davis, M., Khazaen, Y., Crasso, M. P., . . . Geyer, W. (2018). *All work and no play? Conversations with a Question-and-Answer chatbot in the wild*. Paper presented at the CHI 2018, 21-26 april, Montreal, QC, Canada.
- Lortie, C. L., & Guitton, M. J. (2011). Judgment of the Humanness of an Interlocutor Is in the Eye of the Beholder. *PLoS ONE*, 6(9), e25085.
- Luger, E., & Sellen, A. (2016). *"Like Having a Really bad PA": The Gulf between User Expectation and Experience of Conversational Agents*. Paper presented at the #chi4good, CHI 2016, 7-12 maj, San Jose, CA, USA.
- Lynch, P. J., & Horton, S. (2016). *Web style guide: foundation of user experience design* (4 ed.). Yale: Yale University.
- Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90(2017), 46-60.
- McDuff, D., & Czerwinski, M. (2018). Designing Emotionally Sentient Agents. *Communications of the ACM*, 61(12), 74-83.
- McNeal, M. L., & Newyear, D. (2013). Introducing Chatbots in Libraries. *Library Technology Reports*, 49(8), 5-10.
- Meissner, G. (2019). Artificial intelligence: consciousness and conscience. *AI & Society*, 2(2019), 1-11.
- Mone, G. (2016). The edge of the uncanny. *Communications of the ACM*, 59(9), 17-19.
- Mori, M. (1970). Bukimi no tani [The uncanny valley]. *Energy*, 7(4), 33-35.
- Mori, M. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine*, 19(2), 98-100.
- Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72(2017), 432-440.
- Neururer, M., Schlögl, S., Brinkschulte, L., & Groth, A. (2018). Perceptions on authenticity in Chat Bots. *Multimodal Technologies and Interaction*, 2(60), 1-19.
- Nishida, T. (2015). Conversational informatics: Toward cultivating wisdom from conversational interaction. *Procedia Computer Science*, 60(2015), 7-16.
- Norman, D. (2013). *The design of everyday things*. Cambridge: the MIT Press.

- O'Neill, K. (2019). Should A Bot Have to Tell You It's a Bot? Retrieved from <https://medium.com/s/story/should-a-bot-have-to-tell-you-its-a-bot-e9fa29f0b9d4>
- Paikari, E., & van der Hoek, A. (2018). *A framework for understanding chatbots and their future*. Paper presented at the Proceedings of 11th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE'18), 27 maj, Göteborg, Sverige.
- Panetta, K. (2017). Gartner Top Strategic Predictions for 2018 and Beyond. Retrieved from <https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/>
- Park, G., Schwarts, A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., & Stillwell, D. J. (2015). Automatic Personality Assessment Through Social Media Language. *Journal of Personality and Social Psychology*, 108(6), 934-952.
- Pereira, J., & Díaz, Ó. (2019). Using Health Chatbots for Behavior Change: A Mapping Study. *Journal of Medical Systems*, 43(5), 1-13.
- Piccolo, L., Mensio, M., & Alani, H. (2018). *Chasing the chatbots: Directions for Interaction and Design Research*. Paper presented at the CONVERSATIONS 2018, 5th International Conference on Internet Science, 26 oktober, St. Petersburg, Russia.
- Polson, N., & Scott, J. (2018). *AIQ: Hur artificiell intelligens fungerar*. Riga: Daidalos AB.
- Portela, M., & Granell-Canut, C. (2017). *A new friend in our smartphone? Observing interactions with chatbots in the search of emotional engagement*. Paper presented at the Proceedings of Interacción '17, 25-27 september, Cancun, Mexico.
- Schuetzler, R., Grimes, M., Giboney, J. S., & Buckman, J. (2014). *Facilitating Natural Conversational Agent Interactions: Lessons from a Deception Experiment*. Paper presented at the Thirty Fifth International Conference on Information Systems, 14-17 december, Auckland, Nya Zeeland.
- Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, 58(2016), 278-295.
- Skjuve, M., Haugstveit, I. M., Følstad, A., & Brandtzaeg, P. B. (2019). Help! Is my chatbot falling into the uncanny valley? An empirical study of user experience in human-chatbot interaction. *Human Technology*, 15(1), 30-54.
- Statista. (2019). Number of mobile phone messaging app users worldwide from 2016 to 2021 (in billions). Retrieved from <https://www.statista.com/statistics/483255/number-of-mobile-messaging-users-worldwide/>
- Strait, M., Vujovic, L., Floerke, V., Scheutz, M., & Urry, H. (2015). *Too Much Humanness for Human-Robot Interaction: Exposure to Highly Humanlike Robots Elicits Aversive Responding in Observers*. Paper presented at the CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 18-23 april, Seoul, Republic of Korea.
- Tegmark, M. (2017). *Liv 3.0*. Stockholm: Volante.
- Thompson, C. (2018). May A.I help you? *The New York Times Magazine, The tech and design issue*. Retrieved from <https://www.nytimes.com/interactive/2018/11/14/magazine/tech-design-ai-chatbot.html> [2018-11-18]
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433-460.

- van den Akker, J. J. H. (1999). Principles and methods of development research. In J. J. H. van den Akker, R. Branch, K. Gustafson, N. M. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training*. Dordrecht: Kluwer Academic Publishers.
- Vetenskapsrådet. (2002). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*. Retrieved from Stockholm:
- Vinciarelli, A., Esposito, A., André, E., Bonin, F., Chetouani, M., Cohn, J. F., . . . Salah, A. A. (2015). Open Challenges in Modelling, Analysis and Synthesis of Human Behaviour in Human–Human and Human–Machine Interactions. *Cognitive Computation*, 7(4), 397-413.
- Waddell, K. (2017). Chatbots Have Entered the Uncanny Valley. *The Atlantic*. Retrieved from <https://www.theatlantic.com/technology/archive/2017/04/uncanny-valley-digital-assistants/523806/> [2017-04-21]
- Warwick, K., & Shah, H. (2016). Passing the Turing Test Does Not Mean the End of Humanity. *Cognitive Computation*, 8(3), 409-419.
- Weizenbaum, J. (1966). ELIZA - A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- Westerman, D., Cross, A. C., & Lindmark, P. G. (2018). I Believe in a Thing Called Bot: Perceptions of the Humanness of “Chatbots”. *Communication Studies*, 00(00), 1-18.
- Yang, Y., Ma, X., & Fung, P. (2017). *Perceived emotional intelligence in virtual agents*. Paper presented at the CHI’17 Extended Abstracts, 6-11 maj, Denver, CO, USA.
- Zamora, J. (2017). *I’m Sorry, Dave, I’m Afraid I Can’t Do That: Chatbot Perception and Expectations*. Paper presented at the HAI 2017, 17-20 oktober, Bielefeld, Tyskland.

8. Bilaga 1 – Enkäten på originalspråk (engelska)

Perceptions and expectations of chatbots - toward design guidelines

My name is Nina Svenningsson and I'm studying UX-design at Högskolan Kristianstad, in the program Digital Design. I'm currently writing my bachelor thesis in cooperation with inUse/ÅF. The subject of my thesis is to analyze the user experience of chatbots and how a group of different factors can affect this experience. The definition of a chatbot in this survey is a virtual personal assistant that you interact with through written or spoken language, one example is Apple's Siri and another one is the bots that you can chat with in order to get help on various websites.

The results will be used in creating design guidelines for the creation of chatbots. I would be most grateful if you could take the time to help me with answering this survey which will take approximately 5 minutes.

Your responses are anonymous.

Thank you for your participation!

Best regards / Nina

My age is...

I am a...

- Woman
- Man
- Other option
- Do not wish to disclose

Number of hours I spend on the internet every day:

- 0-1
- 2-5
- 6-10
- 11-20

Do you use a smartphone?

- Yes
- No

How much experience do you have with interacting with chatbots?

The definition of a chatbot in this survey is a virtual personal assistant that you interact with through written or spoken language.

- None
- A few times a year
- A few times a month
- A few times a week

- Every day

I have used a chatbot in the following situations:

You can check more than one option.

- Searching for information
- Booking meetings
- Asking for the weather
- Playing music
- To have someone to talk to
- To get inspiration when shopping online
- Customer service
- Surveys/evaluations
- Job related situations
- None of the above
- If other, please specify

If you need help when visiting a website, do you ask for help in the chat (if there is one)?

- Never
- Sometimes
- All the time

What do you do when you need help on a website?

You can check more than one option.

- Use a chatbot
- Use a contact form
- Use e-mail to contact customer service
- Use telephone to contact customer service
- Use social media to get in contact with customer service (Facebook, Twitter etc)
- If other, please specify

Which of the following assistants have you used?

You can check more than one option.

- Apples' Siri
- Amazons' Alexa
- Googles' Virtual Assistant
- Microsofts' Cortana
- Chatbots inside of Facebooks Messenger
- Others (web based chatbots)
- None of the above

Try to evaluate the following scenarios when you interact with a chatbot. Try to imagine how the different scenarios would make you feel.

1. I prefer the following greeting by a chatbot:

- By saying hello.
- By saying hello, followed by my name.
- By saying hello, followed by my name and asking me what I want help with.
- It should be my initiative to start a conversation.

2. I prefer the following response when a chatbot does not understand me:

- "I am sorry, I do not understand."
- "I am sorry, I do not understand. Maybe you can help me by asking your question in a different way?"
- "I am sorry, I do not understand. Do you want to talk with a human instead?"
- "I do not know what you mean, try again!"
- "I have searched the internet for "[your previous question/statement]"..."

3. You have asked a chatbot to help you look for destinations for your next business trip. Which update message do you prefer?

- "Back soon."
- "Searching for destinations..."
- "Going through all the exciting options."
- I don't want an update, only the result when the chatbot has finished the search.
- "Oh, a trip! *clapping hands* Can I come with? Think about it while I look up the best options for your destination."
- "Drifting away, dreaming about the perfect destination, back soon..."

4. You ask a chatbot when the next bus to a train station leaves. Which response do you prefer?

- "12:15"
- "Bus 11 towards *destination* leaves at 12:15."
- "Bus 11 leaves at 12:15, Bus 24 leaves at 12:18 and Bus 11 leaves at 12:25."
- "There are five different buses leaving to the train station from your bus stop, number 11, number 24, number 4, number 13 and number 121. Bus 11 leaves at 12:15, bus 24 leaves at 12:18, bus 4 leaves at 12:26, bus 13 leaves at 12:30 and bus 121 leaves at 12:31."

5. You have asked a chatbot for the weather tomorrow. Which response do you prefer?

- "In what city?"
- "15 degrees celsius."
- "15 degrees celsius and sunny with some clouds."
- "In Gothenburg it will be 15 degrees celsius and sunny with some clouds tomorrow."

6. You are interacting with a chatbot that is connected to your bank. Which opening message do you prefer?

- "Yo, where's the money?"
- "Hello, how can I help you today?"
- "Hi! Are you looking to save or spend today?"
- "Good morning. I am your banking assistant bot. How can I be of service today?"

7. You are interacting with a bot that helps you play music. Which opening message do you prefer?

- "Please state genre, artist or song title in order to receive appropriate choices."
- "Tell me what you want to listen to, and I can help you find it!"
- "Hi! Good to see you again, which genre do you feel like listening to today?"
- "Groovy, you're back! I have some sweet tunes for you, you ready to get down?"

8. You are interacting for the first time with a chatbot at work to book a meeting. Which opening message do you prefer?

- "Hi, I am your meeting buddy! What is on your mind? :)"
- "Hello! Nice to see you here, let me know how I can help you."
- "Do you want to book a room or manage your calendar?"
- "Welcome, would you like to see available times in your calendar?"
- "Welcome, I am a bot that can help you to manage your calendar and book a meeting room. Do you want me to suggest appropriate times for a meeting?"

9. You have booked a trip with a chatbot. Which response do you prefer?

- "Trip booked leaving tomorrow 9:05am arriving 12:08am in Stockholm."
- "Your trip is booked. Leaving tomorrow at 9:05am and arriving 12:08am in Stockholm."
- "Your trip is booked. Your train departs from the station tomorrow at 9:05am and will be arriving at 12:08 in Stockholm."

10. You are frustrated that a chatbot can not understand what you want and have expressed the frustration to the chatbot. Which response do you prefer?

- "I don't understand. Please try again."
- "I am sorry, it seems I am not following, could you please rephrase and try again?"
- "I am so sorry, I really want to help you but I just can not understand. Please rephrase or try simple keywords instead."
- "I am terribly sorry, this is not how I wanted it to end up either..! I am doing my best to help you, please bare with me, I will do everything in my power to make it up to you."

11. You ask a chatbot to help you choose between two hair trimmers. Which response do you prefer?

- "Buy the second option."

- "Buy the second option, it is rated highest by customer reviews."
- "Do you need a trimmer for body, beard or for hair?"
- "You bought a trimmer for body five months ago, is this one for beard or hair?"

9. Bilaga 2 – Enkäten översatt till svenska

Uppfattningar och förväntningar på chatbotar - mot designriktlinjer

Mitt namn är Nina Svenningsson och jag studerar UX-design på Högskolan Kristianstad, på programmet Digital design. Jag skriver för tillfället min c-uppsats i samarbete med inUse/ÅF. Min uppsats syftar till att analysera användarupplevelsen av chatbotar och hur en grupp olika faktorer kan påverka denna upplevelse. Definitionen av en chatbot i denna enkät är en virtuell personlig assistent som man interagerar med via skrift eller talspråk, ett exempel är Apples Siri och ett annat är de botar som man kan chatta med för att få hjälp på olika webbsidor.

Resultaten kommer användas för att skapa designriktlinjer för skapandet av chatbotar. Jag är tacksam om du har möjlighet att hjälpa mig att svara på denna enkät som tar uppskattningsvis 5 minuter.

Dina svar är anonyma.

Tack för ditt deltagande!

Vänliga hälsningar / Nina

Min ålder är...

Jag är en...

- Kvinna
- Man
- Annat alternativ
- Vill inte delge

Antal timmar jag spenderar på internet varje dag:

- 0-1
- 2-5
- 6-10
- 11-20

Använder du en smartphone?

- Yes
- No

Hur stor erfarenhet har du av att interagera med chatbotar?

Definitionen av en chatbot i denna undersökning är en virtuell personlig assistent som man interagerar med via skrift eller genom tal.

- Ingen
- Några gånger per år
- Några gånger i månaden
- Några gånger i veckan

- Varje dag

Jag har använt en chatbot i följande situationer:

Du kan ange mer än ett alternativ.

- För att leta efter information
- För att boka möten
- För att fråga efter vädret
- För att spela musik
- För att ha någon att prata med
- För att få inspiration när jag handlar på internet
- Kundtjänst
- Enkäter/utvärderingar
- I jobbrelaterade situationer
- Ingen av ovanstående
- Om andra, vänligen ange när

Om du behöver hjälp när du besöker en webbsida, frågar du om hjälp i chatten (om det finns en)?

- Aldrig
- Ibland
- Hela tiden

Vad gör du om du behöver hjälp på en webbsida?

Du kan ange mer än ett alternativ.

- Använder en chatbot
- Använder ett kontaktformulär
- Använder email för att kontakta kundtjänst
- Använder telefon för att kontakta kundtjänst
- Använder sociala medier för att kontakta kundtjänst (Facebook, Twitter etc)
- Om annat, vänligen ange vad

Vilken av följande assistenter har du använt?

Du kan ange mer än ett alternativ.

- Apples Siri
- Amazons Alexa
- Googles Virtual Assistant
- Microsofts Cortana
- Chatbots inuti Facebooks Messenger
- Andra (webbaserade chatbotar)
- Ingen av ovanstående

Försök att utvärdera följande scenarion kring att interagera med en chatbot. Försök att föreställa dig hur de olika scenarion skulle få dig att känna.

1. Jag föredrar följande hälsning från en chatbot:

- Att den säger hallå.
- Att den säger hallå, följt av mitt namn.
- Att den säger hallå, följt av mitt namn och frågar vad jag vill ha hjälp med.
- Det ska vara mitt initiativ att starta en konversation.

2. Jag föredrar följande svar när en chatbot inte förstår mig:

- "Jag är ledsen, jag förstår inte."
- "Jag är ledsen, jag förstår inte. Kan du kanske hjälpa mig genom att ställa din fråga på ett annat sätt?"
- "Jag är ledsen, jag förstår inte. Vill du prata med en människa istället?"
- "Jag vet inte vad du menar, försök igen!"
- "Jag har sökt internet på internet efter "[din föregående fråga/påstående]"..."

3. Du har bett en chatbot om hjälp att hitta resmål till din nästa affärsresa.

Vilken uppdatering föredrar du?

- "Snart tillbaka."
- "Söker efter resmål..."
- "Går igenom alla spännande alternativ."
- Jag vill inte ha en uppdatering, bara resultatet när chatboten är färdig med sökningen.
- "Åh, en resa! *klappar händerna* Får jag följa med? Tänk på det medan jag söker upp de bästa alternativen för ditt resmål."
- "Svävar iväg, drömmer om det perfekta resmålet, snart tillbaka..."

4. Du har frågat en chatbot när nästa buss till en tågstation går. Vilket svar

föredrar du?

- "12:15"
- "Buss 11 mot *slutstation* går 12:15."
- "Buss 11 går 12:15, buss 24 går 12:18 och buss 11 går 12:25."
- "Det finns fem olika bussar som går till tågstationen från din hållplats, nummer 11, nummer 24, nummer 4, nummer 13 och nummer 121. Buss 11 går 12:15, buss 24 går 12:18, buss 4 går 12:26, buss 13 går 12:30 och buss 121 går 12:31."

5. Du har frågat en chatbot om morgondagens väder. Vilket svar föredrar du?

- "I vilken stad?"
- "15 grader."
- "15 grader och soligt med lite moln."
- "I Göteborg kommer det vara 15 grader och soligt med lite moln imorgon."

6. Du interagerar med en chatbot som tillhör din bank. Vilket öppningsmeddelande föredrar du?

- "Yo, var är pengarna?"
- "Hallå, hur kan jag hjälpa dig idag?"
- "Hej! Vill du spara eller spendera idag?"
- "God morgon. Jag är din bankassisstentbot. Hur kan jag hjälpa dig idag?"

7. Du interagerar med en chatbot som hjälper dig spela musik. Vilket öppningsmeddelande föredrar du?

- "Vänligen ange genre, artist eller sångtitel för att få lämpliga alternativ."
- "Berätta för mig vad du vill lyssna på så kan jag hjälpa dig att hitta det!"
- "Hej! Kul att se dig igen, vilken genre känner du för att lyssna på idag?"
- "Groovy, du är tillbaka! Jag har lite härliga låtar till dig, är du redo att släppa loss?"

8. Du interagerar för första gången på jobbet med en chatbot för att boka möten. Vilket öppningsmeddelande föredrar du?

- "Hej, jag är din möteskompis! Vad har du på hjärtat? :)"
- "Hallå! Kul att se dig, låt mig veta hur jag kan hjälpa dig."
- "Vill du boka ett rum eller hantera din kalender?"
- "Välkommen, vill du se tillgängliga tider i din kalender?"
- "Välkommen, jag är en bot som kan hjälpa dig att hantera din kalender och boka mötesrum. Vill du att jag föreslår lämpliga tider för ett möte?"

9. Du har bokat en resa med en chatbot. Vilket svar föredrar du?

- "Resa bokad avgång imorgon 9:05 ankomst 12:08 i Stockholm."
- "Din resa är bokad. Avgång imorgon 9:05 och ankomst 12:08 i Stockholm."
- "Din resa är bokad. Ditt tåg avgår från stationen imorgon 9:05 och ankommer till Stockholm 12:08."

10. Du är frustrerad för att en chatbot inte kan förstå dig och du har uttryckt din frustration till chatboten. Vilket svar föredrar du?

- "Jag förstår inte. Vänligen försök igen."
- "Jag är ledsen, det verkar som att jag inte hänger med, är du vänlig och omformulerar och försöker igen?"
- "Jag är verkligen ledsen, jag vill verkligen hjälpa dig men jag förstår helt enkelt inte. Vänligen omformulera eller försök med enklare ämnesord."
- "Jag är hemskt ledsen, det här är inte hur jag ville att det skulle bli heller..! Jag gör mitt bästa för att hjälpa dig, var snäll och ha tålamod så ska jag göra allt för att gottgöra dig."

11. Du ber en chatbot att hjälpa dig välja mellan två hårtrimmers. Vilket svar föredrar du?

- "Köp det andra alternativet."
- "Köp det andra alternativet, den har fått högst betyg i kundrecensioner."

- "Behöver du en trimmer för kropp, skägg eller hår?"
- "Du köpte en trimmer för kropp för fem månader sedan, är den här för skägg eller hår?"