Asko Uuras

Delegating tasks to ChatGPT

An empirical approach to understanding delegation between agents

School of Technology and Innovation
Master's thesis in Economics and Business Administration
Program in Information Systems

Vaasa 2023

UNIVERSITY OF VAASA
School of Technology and Innovation
Author: Asko Uuras
Title of the Thesis: Delegating tasks to ChatGPT
Degree: Master of Science in Economics and Business Administration
Program: Information Systems
Supervisor: Duong Dang
Year: 2023 **Pages:** 55

ABSTRACT :

The information system (IS) delegation theoretical framework addresses the technological advancements in artificial intelligence, which the dominant IS use theory has yet to be able to answer due to the human agent primacy. Delegation, in the context of IS, is an action between a human agent and an agentic IS artifact where the human agent does not simply use the information system but delegates a task or subtask to an agentic IS artifact. Since the launch of ChatGPT in late 2022, the large public has explored and benefited from the capabilities of this new agentic IS artifact by delegating tasks. For organizations and individuals to understand their interaction with ChatGPT, the delegation process must be addressed.

This thesis aims to answer the research question of how human agents delegate work tasks to ChatGPT. Qualitative and quantitative methods are utilized to gather relevant information for a delegation conceptual model with semi-structured interviews and a self-administered web survey. The thesis presents the delegation conceptual model, replicating and extending the existing IS delegation theory to a new agentic IS artifact. The thesis provides a practical implication by identifying the primary delegation mechanism – appraisal – for human agents who utilize ChatGPT in a work setting. Identifying the delegation mechanism provides insight into human agents' delegation process and affecting attributes. In addition, the thesis has two contributions to how human agents perceive delegation. First, survey results suggest that human agents do not acknowledge the action of delegation when interacting with ChatGPT. Second, human agents do not acknowledge all the subdimensions of delegation, such as transfer of rights and transfer of responsibilities, when interacting with ChatGPT.

To build the model, delegated tasks and delegation situations have been identified through 21 interviews and the most common task and situation through 132 survey responses. As a part of identifying the delegation mechanism, four hypotheses have been set to test the human agents' acknowledgment of subdimensions of delegation. New unvalidated dichotomous measurement items have been created, and hypotheses have been tested through the Pearson x^2 goodness of fit statistical test. According to the results presented in this thesis, delegated tasks are usually subtasks where the human agent has an incentive to produce non-critical text or text to supplement planned text. Both interviews and survey results emphasize the superiority of human agents' decision-making capabilities, ability to assess generated text, and knowledge to check facts on behalf of the ChatGPT. On the other hand, ChatGPT has been described as writing sufficient text promptly, surpassing the human agent's slow pace of text production and even potentially avoiding human agent procrastination towards text production.

KEYWORDS: artificial intelligence, chatbots, information systems science, empirical research, delegating

VAASAN YLIOPISTO
Tekniikan ja innovaatiojohtamisen akateeminen yksikkö
Tekijä: Asko Uuras
Tutkielman nimi: Delegating tasks to ChatGPT
Tutkinto: Kauppatieteiden maisteri
Oppiaine: Tietojärjestelmätiede
Työn ohjaaja: Duong Dang
Vuosi: 2023 **Sivumäärä:** 55

TIIVISTELMÄ :

Tietojärjestelmätieteen delegoinnin (IS delegation) teoreettinen viitekehys pyrkii vastaamaan tekoälyn teknologiseen kehitykseen täydentäen vallalla olevaa ihmiskeskeisen käyttäjyyden (IS use) näkökulmaa. Delegointi tietojärjestelmätieteen kontekstissa on toimijuutta harjoittavien ihmisen ja artefaktin välistä toimintaa, jossa ihminen toimijana ei vain käytä tietojärjestelmää, vaan delegoi tehtävän tai alitehtävän artefaktille. ChatGPT:n julkaisusta, loppuvuodesta 2022 lähtien, suuri yleisö on etsinyt tapoja hyödyntää tämän uuden artefaktin ominaisuuksia. Jotta organisaatiot ja yksilöt ymmärtäisivät paremmin toimintaansa ChatGPT:n kanssa, delegointia prosessina on tutkittava.

Tämän tutkielman tavoitteena on vastata tutkimuskysymykseen siitä, kuinka ihmiset delegoivat työtehtäviä ChatGPT:lle. Tutkimusaineiston keruussa on hyödynnetty kvalitatiivisia ja kvantitatiivisia menetelmiä. Aineistoa on kerätty delegaation käsitteellistä mallia varten puolistruktuuroiduilla haastatteluilla ja verkkokyselyllä. Tutkielmassa esitellään delegoinnin käsitemalli, joka toistaa ja laajentaa olemassa olevaa delegointiteoriaa uudella artefaktilla. Käytännön johtopäätöksenä tunnistetaan ensisijainen delegointimekanismi: arviointi. Johtopäätös on hyödyllinen ihmisille, jotka työskentelevät ChatGPT:n kanssa. Delegointimekanismin tunnistaminen antaa käsityksen ihmisten delegointiprosessista ja delegointiin vaikuttavista osa-alueista. Lisäksi tutkielmassa esitellään kaksi havaintoa siitä, miten ihmiset hahmottavat delegoinnin. Ensinnäkin tutkimustulokset viittaavat siihen, että ihmiset eivät tunnista delegointia vuorovaikutuksessaan ChatGPT:n kanssa. Toiseksi ihmiset eivät tunnista kaikkia delegoinnin ulottuvuuksia, kuten oikeuksien ja velvollisuuksien siirtoa, ollessaan vuorovaikutuksessa ChatGPT:n kanssa.

Käsitemallin rakentamiseksi delegoituja tehtäviä ja delegointitilanteita on kartoitettu 21 haastattelun avulla. Yleisin delegoitu tehtävä ja delegointitilanne on puolestaan tunnistettu 132 kyselyvastauksen kautta. Lisäksi osana delegointimekanismin tunnistamista on asetettu neljä hypoteesia testaamaan, kuinka ihmiset tunnustavat delegoinnin ulottuvuuksia. Hypoteesitestausta varten tutkielmassa on luotu validoimattomia dikotomisia kysymyksiä, jotka vastaavat delegoinnin ulottuvuuksia. Hypoteesit on testattu tilastollisen Pearsonin x^2 -testin avulla. Tässä tutkielmassa esitettyjen tulosten mukaan delegoidut tehtävät ovat yleensä osatehtäviä, joissa ihmisellä on kannustin tuottaa ei-kriittistä tekstiä tai tekstiä täydentämään jo suunniteltua tekstiä. Sekä haastatteluissa että kyselytuloksissa korostetaan ihmisten päätöksentekokyvyn ylivoimaisuutta, kykyä arvioida luotua tekstiä sekä osaamista tarkistaa tosiasiat ChatGPT:n puolesta. Toisaalta ChatGPT:n on kuvattu päihittävän ihmisen tekstin tuottajana kirjoittamalla riittävän laadukasta tekstiä nopeasti ja tarjoavan ihmiselle jopa mahdollisuuden välttää viivyttely tekstintuotannon aloittamisessa.

AVAINSANAT: artificial intelligence, chatbots, information systems science, empirical research, delegating

Contents

Figures

Tables

Abbreviations

AGI	**Artificial general intelligence**
AI	**Artificial intelligence**
Bof8	**Basket of Eight**
FNN	**Feedforward neural network**
GPT	**Generative pre-trained transformer**
IS	**Information system**
LLM	**Large language model**
ML	**Machine learning**
NLP	**Natural language processing**
RNN	**Recurrent neural network**

1 Introduction

Information systems (IS) use is one of “the most central constructs” and “the most widely-studied construct” in the IS discipline (Burton-Jones, 2020). IS use literature in the past has emphasized human agency and the passive nature of IS artifacts (Baird & Maruping, 2021). However, the newly launched large language model (LLM) applications such as ChatGPT are not passive but are more agentic – alongside other artificial intelligence (AI) technology (Bawack et al., 2019). Recently, the idea of agentic IS artifacts has been presented with the concept of delegation (Baird & Maruping, 2021). According to Baird & Maruping (2021), user-centered discourse does not serve research of agentic IS artifacts such as medical agents, autonomous vehicles, and chatbots.

Due to the launch of ChatGPT in November 2022, the larger public started to utilize the capabilities of agentic IS artifacts on an unforeseen scale. To demonstrate the scale, ChatGPT reached a million users only within five days (Browne, 2023) and 100 million users in two months (Hu, 2023) after its launch. Straight after the launch, economic research started to gain interest in the topic (Ritala et al., 2023). Especially the social aspects and human role have been of keen interest in early 2023. Although many researchers have raised their concerns about the social impacts of AI, there is still a lack of research about the transformation in corporate functions deployment in society and the acceptance of AI tools such as ChatGPT (Dwivedi et al., 2023; Yang & Wang, 2023). ChatGPT’s capability to write text has also brought up concerns about how the technology is utilized (Dwivedi et al., 2023). Therefore, asking how human agents delegate tasks to ChatGPT in a work setting is essential. Since ChatGPT is a relatively new application, delegation mechanisms have yet to be studied in the context of ChatGPT. Organizations and individuals must consider delegation mechanisms and attributes related to task, agent, and situation to benefit from the delegation. This thesis answers aspects above through the IS delegation theoretical framework and by presenting a conceptual model following the guidelines provided by Baird and Maruping (2021).

The thesis identifies the delegation mechanism, appraisal, as a primary practical implication for human agents utilizing ChatGPT in the work setting. The thesis also has two contributions to how human agents perceive delegation. First, survey results suggest that human agents do not acknowledge the action of delegation when interacting with ChatGPT. In survey results, most human agents replied that they do ask ChatGPT to execute tasks, but more than half of the survey respondents answered that they do not delegate tasks to ChatGPT. Second, human agents do not acknowledge all the subdimensions of delegation, such as transfer of rights and transfer of responsibilities, when interacting with ChatGPT. While 107 human agents responded that they had asked ChatGPT to execute a work task, only 23 of them responded that they do transfer responsibilities to ChatGPT for the execution of a work task, and 40 of them responded that they do transfer rights to ChatGPT for the execution of a work task.

The thesis is divided into the following sections. Section 2 covers the essential technological advances from the definition of artificial intelligence to generative pre-trained transformers for a reader to grasp the much-needed shift behind the use and delegation terminology. The section also reviews the current knowledge in the form of a literature review and identifies relevant hypotheses to identify the relevant delegation mechanism. Section 3 explains the research design and methodology applied to the study. Section 4 introduces the results of the interviews and survey. Section 5 discusses the new conceptual model, including the task attributes, delegation mechanism, agentic and situational attributes, and feedback loops. Section 6 concludes the thesis with an exhibit of practical implications and contributions.

2 Literature review and hypotheses

2.1 Technological background

Since artificial intelligence (AI) still lacks a complete and widely accepted definition (Aleksander, 2017; Berente et al., 2021; Nilsson, 2009; Kok et al., 2009; Russell & Norvig, 2010; Samoili et al., 2020), the lack of definition remains as an obstacle in the discussion about delegation. One of the earliest attempts to define intelligence occurring in machines can be traced back to 1950 when Alan Turing reframed the question "Can machines think?" into the form of a game (Turing, 1950). In the game, the interrogator asks written questions and receives written answers (Dobrev, 2012; Turing, 1950; Russell & Norvig, 2010). The game boils down to the question: Can the interrogator differentiate the written responses and tell if it is from a machine or a human (Dobrev, 2012; Kok et al., 2009; Russell & Norvig, 2010)? To pass the game, called the total Turing Test, the intelligent computer would need capabilities from many subfields of AI such as natural language processing, knowledge representation (knowledge-based systems), automated reasoning (automated planning and scheduling, and optimization), machine learning, computer vision, and robotics (Russell & Norvig, 2010; Abioye et al., 2021). Even though the total Turing test does not provide a definitive definition, it still provides an overview of what elements we should look at when discussing AI. When considering information systems literature, information systems scholars have also suggested multiple definitions for AI. For example, AI has been defined as predictive statistical models (Fernández-Loria et al., 2020), or "technologies that leverage machine-based intelligence and advanced computing capacity to mimic human 'cognitive' functions" (Li et al., 2021, p. 1603), or "in short, AI is whatever we are doing next in computing" (Berente et al., 2021, p. 1435). Within the information systems field, these definitions build on the view that while machine aims to impersonate human behavior, the subfields of AI aim to complement each other to reach even better imitation.

Machine learning is one of these AI subfields (Helm et al., 2020). In addition to artificial intelligence, ML can improve by analyzing large data sets by deriving patterns from data

and thereby learning (Helm et al., 2020; Mitchell, 1997; Sturm et al., 2021). According to Mitchell (1997, p. 2), “a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” To measure experience E, as suggested by Mitchell (1997), the machine learning algorithm has usually been exposed training data (Padmanabhan et al., 2022).

Padmanabhan et al. (2022) suggest training data may include observed labels or target variables. Learning is considered supervised if training data includes target variables; otherwise, unlabeled learning data sets fall into unsupervised learning (Padmanabhan et al., 2022). According to Cunningham et al. (2008) and James et al. (2023), supervised learning data sets aim to train machines through the target variables or true answers, which are used to predict how well the model fits the response. This is impossible in unsupervised learning since the true answer is unknown (James et al., 2023).

Before representational learning, raw data sets (e.g., pixel values) had to be transformed into a suitable form for the algorithm to learn (LeCun et al., 2015). According to LeCun et al. (2015, p. 436), in representation, a learning machine harness method that enables it to be “fed with raw data and automatically discover the representation needed for detection or classification.” Now, e.g., raw images could be given to a representational learning algorithm without manipulation on the trainer’s part. This leads to the next leap in technological advancement - deep learning. According to LeCun et al. (2015), deep learning is a learning method consisting of multiple levels of representation. For example, a multiple-level representation of an image could be that the first layer represents the boundaries of a specific element, the second the position of the element, and the third the correspondence with other objects (LeCun et al., 2015). In other words, deep learning algorithms can take a raw image, extract layers from the image, and use these layers as learning data to learn without human intervention.

The progress in machine learning, especially deep learning, has enabled machines to learn faster and use training data that does not have the correct answer. This has led to the evolution of large language models combining the capabilities of machine learning and another subfield of AI - natural language processing (Beltagy et al., 2019; Shahriar & Hayawi, 2023). Natural language processing (NLP) is "an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things" (Chowdhary, 2003, p. 51). Therefore, as the name suggests, natural language processing is a way to interact with computers with a natural human language syntax, which computers learn and process.

Bengio (2008) states that a language model is "a function, or an algorithm for learning such a function, that captures the salient statistical characteristics of the distribution of sequences of words in a natural language." A large language model (LLM) is a deep learning model that aims to generate and understand natural language (Shen et al., 2023). According to Beltagy et al. (2019), language models ELMO (Peters et al., 2018) and BERT (Devlin et al., 2019) alongside GPT (Radford et al., 2018) have shown that performance on many NLP tasks has improved through unsupervised pre-training.

As the latest development in AI, GPT is also behind the agentic IS artifact of ChatGPT. GPT is an abbreviation of words generative pre-trained transformer. The word generative refers to a generative model that generates, for example, an answer sequence when answering a question (Luo et al., 2022). As Radford et al. (2018, p. 2) describe, unsupervised pre-training is "a special case of semi-supervised learning where the goal is to find a good initialization point instead of modifying the supervised learning objective." Transformer is a feedforward neural network (Vaswani et al., 2017). In short, a feedforward neural network (FNN) has an acyclic topology graph, meaning that FNN moves information only forward, unlike recurrent neural networks (RNN), in which units (nodes) can be connected with all non-input units (Schmidhuber, 2015).

ChatGPT is an application of previously described LLM derived from AI subfields of ML and NLP. ChatGPT has been developed by OpenAI, L.L.C. (OpenAI, 2023a), and the company focuses on research and deployment of artificial general intelligence (AGI) (OpenAI, 2023b). From a legislative perspective, ChatGPT is mentioned by name in the European Parliament News article, which introduced regulation on artificial intelligence (European Parliament, 2023). European parliament categorizes ChatGPT as generative AI (2023).

2.2 Delegation

Technological advancement has moved the IS artifacts toward agentic artifacts from the original human-centered view of IS artifacts (Baird & Maruping, 2021). Therefore, according to Baird & Maruping (2021), a human agent is a human who interacts with an agentic IS artifact, and a human is more like an agent rather than a user. They discuss in their paper that agentic IS artifact, in the most expansive decision-making latitude, can be, for example, an autonomous vehicle, legal agent, or chatbot (Baird & Maruping, 2021). Therefore, in this context, the construct of delegation can be extended to ChatGPT, which is a chatbot and enhances artificial intelligence.

The delegation, in the field of technology and information systems science, has focused on the task execution between humans and IS (Bawack et al., 2019; Castelfranchi & Falcone, 1998; Fügener et al., 2022; Leyer & Schneider, 2019; Miller et al., 2011). While the economics literature (Ambrus et al., 2021; Banford et al., 2014; Holmstrom, 1978; Holmstrom & Milgrom, 1991) has focused on transferring rights and responsibilities and the right to take an action. Transferring rights and responsibilities have been activities between human superiors and human subordinates (Holmstrom, 1978). Baird and Maruping (2021) defined that delegation may or may not include transferring rights and responsibilities to another agent (human agent or agentic IS artifact) for task execution. However, the delegation mechanism varies depending on the level of transfer of rights and responsibilities (Baird & Maruping, 2021). Merriam-Webster dictionary (2023a; 2023b; 2023c) defines right as “the power or privilege to which one is justly

entitled," responsibility as "moral, legal, or mental accountability," and task as "a usually assigned piece of work often to be finished within a certain time."

According to Castelfranchi & Falcone (1998), delegation could be informally defined as follows: Agent A has a need or desire for an action that is performed by Agent B, and Agent A incorporates the action into its plan. Merriam-Webster dictionary (2023d) defines action as a fact or process of doing something, typically to achieve an aim. In other words, action is what an agent does. According to Castelfranchi and Falcone (1998), a plan consists of at least one composed action to reach a goal. For example, Agent A could add the task to its planned list of actions while Agent B is expected to deliver the task (Castelfranchi & Falcone, 1998). More formally, Holmstrom (1978) defines delegation as a "decentralized decision process" where a principal allows an agent to decide from a pool of alternatives. According to Holmstrom (1978), from the principal's perspective, the delegation problem is that the principal should discover the complete batch of alternatives to delegate. Decision-making responsibility is transferred from the principal to the agent since the agent has "superior information about the principal's decision problem" (Holmstrom, 1978, p. 1).

2.3 Hypotheses

As the IS delegation framework is applied to answer the research question, it is essential to consult the framework when formulating the hypotheses. Hypothesis testing is used primarily to identify the delegation mechanism. IS delegation framework (Baird & Maruping, 2021, p. 237-239) recognizes three main delegation mechanisms: *Appraisal, Distribution,* and *Coordination. Appraisal* may include emotional and cognitive evaluation of how the "human agent feels about the delegation to agentic IS artifact" and the potential "costs and benefits." *Distribution* as a delegation mechanism contemplates the transfer of rights and responsibilities. As the authors (p. 328) point out, the transfer may be "complete or partial", and the transfer could need a negotiation to agree on what extent rights and responsibilities are transferred from one agent to another. *Coordination* requires delegators to approach from the perspectives of appraisal and distribution,

and acknowledge the roles of supervisors or monitors. In *coordination*, both agents keep each other informed about the tasks at hand, and accountability, predictability and shared understanding are integrative conditions of this delegation mechanism. Delegation mechanisms build on each other, and *appraisal* is the foundation on which *distribution* and *coordination* build. Therefore, *appraisal* is the most elementary delegation mechanism used by human agents and agentic IS artifacts (Baird & Maruping, 2021).

The tasks mentioned by the first guideline of the IS delegation framework (Baird & Maruping, 2021) are identified through interview and survey. The literature (Ritala et al., 2023) suggest that text production is one of the most evident tasks that could be given to ChatGPT. Attributes of the agents and situations in the context of ChatGPT have also been studied by Dwivedi et al. (2023) and Ritala et al. (2023). The delegation mechanism from perspectives of task execution, transfer of rights, transfer of responsibilities, and delegation has yet to be studied, and we may only hypothesize about the outcomes. According to Baird & Maruping (2021, p. 327), the delegation mechanisms "represent the elements of the delegation process" and are mentioned in the second guideline of the IS delegation framework. Therefore, as delegation mechanisms are essential to identify and thoroughly utilize the IS delegation framework, the hypotheses presented in this thesis focus on the delegation mechanisms.

As stated in previous literature, a task is required in the action of delegation (Baird & Maruping, 2021; Bawack et al., 2019; Castelfranchi & Falcone, 1998; Fügener et al., 2022; Leyer & Schneider, 2019; Miller et al., 2011). When considering the IS delegation framework, the most fundamental delegation mechanism of appraisal builds on the evaluation of benefits and costs when delegating a task, and the first guideline of the framework requires the identification of a task (Baird & Maruping, 2021). Therefore, it is assumed that human agents ask ChatGPT to execute tasks. Let the null hypothesis be "human agents do ask ChatGPT to execute work tasks" and the alternative hypothesis "human agents do not ask ChatGPT to execute work tasks."

The second and third hypotheses identify the transfer of rights and transfer of responsibilities when human agents interact with ChatGPT. Transfer of rights and responsibilities is a part of the delegation, which entitles the agent to work with a specific task and assumes accountability for the task execution (Ambrus et al., 2021; Banford et al., 2014; Holmstrom, 1978; Holmstrom & Milgrom, 1991). According to the IS delegation framework (Baird & Maruping, 2021), transferring rights and responsibilities is necessary for the distribution and coordination delegation mechanisms. With the acknowledgment of the transfer of rights and responsibilities, human agents can utilize these delegation mechanisms. Therefore, it is crucial to test whether human agents acknowledge the transfer. For the second hypothesis, considering the transfer of rights, let the null hypothesis be "human agents do transfer rights to ChatGPT for work task execution," the alternative hypothesis is "human agents do not transfer rights to ChatGPT." For the third hypothesis, considering the transfer of responsibilities, let the null hypothesis be "human agents do transfer responsibilities to ChatGPT for work task execution," and the alternative hypothesis "human agents do not transfer responsibilities to ChatGPT for work task execution." If null hypotheses are rejected, human agents do not utilize distribution or coordination delegation mechanisms when interacting with ChatGPT.

The fourth hypothesis considers that human agents may perceive the term delegation differently. As Baird and Maruping (2021) state, they have built their IS delegation framework on three concepts: Delegation involves two agents, a task or desired outcome brings these two agents together, and rights and responsibilities for tasks or outcomes are transferred. This hypothesis tests, first and foremost, whether the human agents delegate tasks to ChatGPT. Therefore, let the null hypothesis be "human agents do delegate work tasks to ChatGPT and the alternative hypothesis "human agents do not delegate work tasks to ChatGPT."

2.4 Summary

This section reviews the technical background that has created the need to discuss the agentic IS artifacts as independent and autonomous agents. The enormous advances in machine learning and natural language processing have led to the development of agentic IS artifacts, which possess the capabilities to work with open-ended and ambiguous tasks. This section also reviews the previous literature from the delegation perspective and draws research hypotheses from the literature. Since the definition of delegation and potential delegation mechanisms are known, it is possible to study further the agentic IS artifacts such as ChatGPT.

3 Research design

Considering the research question, literature review, and applied framework - including both qualitative and quantitative research methods is appropriate. Semi-structured interviews, as a qualitative research method, provide an interactive and natural approach to the topic (Saunders et al., 2019). Semi-structured interviews will provide *tasks* that human agents see the potential to be delegated, providing one of the most inseparable parts of the IS delegation framework (Baird & Maruping, 2021). Also, interviews have been used recently as a research method by Ritala et al. (2023) to identify potential tasks. Quantitative methods must be utilized to provide a broader view of the most common tasks and delegation mechanisms used in the population (Nummenmaa, 2021). In order to test hypotheses, the measurement items need to be created for the survey. The measurement items also capture the subdimensions of the IS delegation construct, which are at least partly latent.

3.1 Measurement items creation for the construct of IS delegation

As the literature review identifies, IS delegation framework provides nuanced information about the interaction between human agents and agentic IS artifacts. However, delegation as a construct does have subdimensions of transferring rights and responsibilities, which are essential to understanding the interaction between human agents and agentic IS artifacts. Baird & Maruping (2021) consider before mentioned subdimensions of delegation, but the methodology of how to verify the subdimensions is not discussed in their framework. Therefore, hypotheses have been set to understand whether the human agents identify subdimensions in their interaction with ChatGPT. To test the hypotheses, construct measurement items need to be created. To create measurement items, it is essential to define the construct of *IS delegation* and create formative indicators to measure the underlying latent variable. Measurement item creation follow the guidance of MacKenzie et al. (2011) and Ricci et al. (2019).

The formative indicator describes the "relationship between an indicator and latent construct with which it is associated" when subdimensions of a construct define the construct (MacKenzie et al., 2011, p. 296). As MacKenzie et al. (2011) and Ricci et al. (2019) suggest, the previous literature on delegation and agentic IS in information systems and related constructs in economics literature is further reviewed.

As presented in the literature review, the information systems and economic literature suggest conceptual construct definition that: IS delegation is an *action* between human agent and agentic IS artifact (Baird & Maruping, 2021; Bawack et al., 2019; Castelfranchi & Falcone, 1998; Fügener et al., 2022; Leyer & Schneider, 2019; Miller et al., 2011), which transfers *rights* and *responsibilities* (Ambrus et al., 2021; Baird & Maruping, 2021; Banford et al., 2014; Holmstrom, 1978) to execute a *task* (Alonso & Matouschek, 2008; Holmstrom & Milgrom, 1991). Therefore, the presented property is the act of delegation. The entities where the general property applies are human agents and agentic IS artifacts.

In other words, the literature suggests that IS delegation has at least three subdimensions, which must be considered in the item creation. According to the literature, these three subdimensions are *task execution*, *transfer of rights*, and *transfer of responsibilities* since they are inseparable parts of the definition of delegation (Ambrus et al., 2021; Baird & Maruping, 2021; Holmstrom, 1978). When using MacKenzie et al. (2011, p. 301) questions to determine if the focal construct is multidimensional, it is possible to conclude that these subdimensions have distinctive characteristics, according to their dictionary definitions. Also, eliminating one of the characteristics would cause a significant restriction to the domain of the construct since characteristics define the underlying construct. Therefore, task execution and transferring rights and responsibilities are essential to IS delegation, which can be defined as a multidimensional construct. Furthermore, the focal construct of IS delegation exists at the same level as its subdimensions, and subdimensions are parts of the additive function of this focal construct. According to MacKenzie et al. (2011), the subdimensions mentioned above and subdimensions'

relations with the construct can be thought of as formative indicators since change within one subdimension behavior might be associated with other subdimension (e.g., transfer of rights might be associated with transfer of responsibilities), and change in one subdimension could be associated with a change in the focal construct.

Since information systems literature has viewed delegation as an implicit part of the interaction (Baird & Maruping, 2021), the subdimensions have received only occasional attention in the previous literature. Therefore, items must be generated to create a starting point to study the delegation's subdimensions of task execution, transferring rights and responsibilities. To create clear, simple, and precise wording, items are created in a manner that explicitly describes a situation where the respondent is conversing with ChatGPT. This wording avoids misunderstandings where the respondent has not decided to have an interaction with ChatGPT. Furthermore, wording aims to avoid misunderstandings where metaknowledge has a role when delegating a specific task. Four items were generated, following the construct definition suggested by MacKenzie et al. (2011).

Item 1: During a conversation with ChatGPT, I asked ChatGPT to execute a work task to reach a desired outcome. A *task* is a usually assigned piece of work often to be finished within a certain time. This item aims to understand have respondent asked ChatGPT to execute any work task. The task can be part of a larger task or complete work.

Item 2: During a conversation with ChatGPT, I transferred rights to ChatGPT for the execution of a work task to reach a desired outcome. The definition of *right* is the power or privilege to which one is justly entitled. This item aims to understand does respondent transfer power or privilege to ChatGPT so that ChatGPT is entitled to work with a specific task.

Item 3: During a conversation with ChatGPT, I transferred responsibilities to ChatGPT for the execution of a work task to reach a desired outcome. The definition of

responsibility is moral, legal, or mental accountability. This item aims to understand does respondent transfer moral, legal, or mental accountability to ChatGPT so that ChatGPT has been morally, legally, or mentally accountable for the task execution.

Item 4: During a conversation with ChatGPT, I delegated work tasks to ChatGPT. Delegation is an *action* between you and ChatGPT that *transfers rights and responsibilities* from you to ChatGPT for the *execution of a task*. This item aims to understand whether has respondent delegated work tasks to ChatGPT for any reason, and has respondent transferred both rights and responsibilities to ChatGPT in order to reach the desired outcome.

3.2 Semi-structured interview design

Qualitative research methods have originated in anthropology and sociology, and focus on studying human behavior (Hove & Anda, 2005). Semi-structured interview considers the guidelines of the IS delegation framework (Baird & Maruping, 2021), and questions aim to gather information about delegation, tasks given to ChatGPT, frequency of conversations with ChatGPT, situations when tasks are delegated to ChatGPT, and attributes of tasks and situations.

As the interviews aim to gather information about the common tasks and delegation situations, the interviewees must interact with ChatGPT often with various tasks and situations (Tiainen, 2014). Therefore, the interviewees have been gathered from a pool of human agents who have published either a journalistic article or a LinkedIn post about the interaction with ChatGPT. The interviewees also consider their interaction with ChatGPT frequent. Answering to the interview questions has been voluntary for the interviewees. Marshall et al. (2013) suggest collecting 20-30 interviews to reach sufficient saturation. According to Glaser & Strauss (2017), saturation .70 Cronbach alpha could be reached with 12 interviews. A total of 21 interviews were conducted in May 2023.

The interview structure follows responsible conduct of research guidelines (TENK, 2012) of a good interview. Conducted interviews were semi-structured, and the language of

the interview was selected based on the preferred language of the interviewee. Open-ended questions allow interviewees to explain and ponder their interaction with ChatGPT more deeply than in a self-administered web survey setting (Marshall et al., 2015). Interviews were recorded during a remote interview. The interview structure was explained to the interviewee, and permission to record was asked before starting the interview. Interviewees were informed that they participated in research for master's thesis. Interviewees could see notes taken during the interview on a shared screen to ensure they agree with the written notes.

Table 1 Interview questions.

Do you actively use ChatGPT?
How often you use ChatGPT?
Do you delegate work tasks to ChatGPT?
What kind of tasks do you delegate (or give) to ChatGPT?
When do you delegate or (give tasks) to ChatGPT?
What kind of tasks you DO NOT (or would not) delegate to ChatGPT?
Have you observed other tasks which could be delegated, but which you do not delegate?
Would you say that ChatGPT delegates tasks to you?

Interview questions are presented in Table 1. Interview questions were formulated based on Baird & Maruping (2021) guidelines. In their research for developing models, they describe three basic guidelines: 1) identifying and explicating the most important attributes of the tasks or desired outcomes, 2) identifying and analyzing salient delegation mechanisms according to tasks, and 3) identifying and analyzing attributes of agents according to related task and delegation mechanism. They also describe two optional or situational guidelines 4) identifying and analyzing salient situational attributes related to the task and delegation mechanisms and 5) identifying and analyzing any feedback loops resulting from the outcome of delegation (Baird & Maruping, 2021).

Questions "Do you actively use ChatGPT," "How often do you use ChatGPT?" and "Do you delegate work tasks to ChatGPT?" focus on validating the interviewee to be qualified to answer the following questions. Also, "Do you delegate work tasks to ChatGPT" aims

to gather indications of prevalence among interviewees. The question “What kind of tasks do you delegate to ChatGPT?” aims to gather information regarding the first guideline of task or desired outcome and situations when a task is delegated. The question “When do you delegate tasks to ChatGPT?” answers the situational guideline and paves the way to the conversation about the delegation mechanisms and attributes of the agents described in the second and third guidelines.

Optional guideline of attributes of the delegation situation is included in the questions “Have you observed other tasks which could be delegated but which you do not delegate?” and “What kind of tasks do you not delegate to ChatGPT?”. As Baird & Maruping (2021, p. 330) note, “the nature of the task or outcome may vary by the situation.” Therefore, differentiating possible and non-delegating tasks provided a possibility to discuss situations and reasoning why ChatGPT either is not asked or cannot be asked to execute a specific task.

The last question, “Would you say that ChatGPT delegates tasks to you?” asks the interviewee to ponder the delegation feedback loops described by Baird and Maruping (2021) in the fifth guideline. Feedback loops between agents are suggested to be studied to understand the impact of the delegation on both agents. Feedback from human agents enables ChatGPT to learn (OpenAI, 2023b), but “the outcome of the delegation episode may shape the human agent’s judgment about whether to delegate in their role as a delegator or how to execute in their role as a proxy” (Baird & Maruping, 2021, p. 331).

3.3 Self-administered web survey design

According to Mazaheri et al. (2020), a survey has been the dominant research method in information system science literature. A survey is a way of gathering response data about “characteristics, actions, perceptions, attitudes, or opinions” of a large group of individuals, groups, or organizations, and it can be delivered in multiple ways to respondents (Straub et al., 2022). The survey conducted in this thesis was a self-administered web survey, which was published through the Webropol survey service.

The objective of the survey was to explore further the interview observations about the delegation regarding the IS delegation framework and gather data for hypothesis testing via measurement scale items. The measurement item creation is presented in section 3.1.

Background questions consist of the year of birth, self-reported usage, level of education (Igbaria et al., 1995; Kuegler et al., 2015), organizational level (Igbaria et al., 1995), and questions regarding subordinates. Self-reported usage is a respondent's estimation of IS use. Self-reported usage was measured through six-position categorical scales and followed the scale created by Davis (1989). Level of education includes options: elementary, high school / vocational, bachelor's degree, master's degree, and Ph.D. degree (Igbaria et al., 1995; Kuegler et al., 2015). Organizational level has been divided into two categories: *professional* and *managerial* (supervisors and managers) (Igbaria et al., 1995). Igbaria et al. (1995) classified organizational level according to the title and hierarchical level. *Professional* describes specialists in their field, and their job titles are, e.g., analyst, designer, or programmer. *Managerial* describes administrative persons whose job titles are, e.g., manager or leader (Igbaria et al., 1995). In this survey, respondents were asked to select the managerial level if their job title suggests they are a supervisor or manager or if they have a job title that indicates managerial responsibilities, e.g., leader. Otherwise, they have asked to select a professional. While Igbaria et al. (1995) dichotomic classification does not further identify respondents in managerial positions, the question about having subordinates was included.

Baird and Maruping (2021) theorize about delegation based on tasks and delegation mechanisms. A comprehensive list of tasks and situations was gathered through interviews. These tasks and situations are the basis of the questionnaire items. These items aim to gather data about the frequency of the tasks and delegation mechanisms to support the conceptual model.

The survey was open during 1.8.-31.8.2023. A total of 132 answers were gathered. Before opening the survey, a LinkedIn poll was posted to identify persons who had tested ChatGPT. The poll asked, in Finnish, the question: “Have you tested ChatGPT?” (“Oletko testannut ChatGPT:tä?” in Finnish). The poll was open for one week. During a week, a total of 272 votes were gathered. 90% of the votes (245 votes) declared they had tested ChatGPT. After opening the survey, the researcher sent the survey link to 245 persons who had tested ChatGPT. To gather more responses, the researcher posted three times on LinkedIn about the survey. The link was open to everyone, and no contact information was gathered. Therefore, the respondents’ identity and how many of the 245 persons answered the survey are unknown. In addition, interviewees were asked to fill in the survey. The link was sent to 17 interviewees, of which 10 answered the survey.

3.4 Hypothesis testing

As mentioned earlier, IS delegation is an action between a human agent and an IS artifact. In order to utilize distribution or coordination delegation mechanisms, the human agent has to transfer the responsibility and rights to the agentic IS artifact (Baird & Maruping, 2021; Holmstrom, 1978). Hypotheses are tested through measurement scale items created in section 3.1. Each hypothesis has a related scale item. The hypothesis testing aims to support the selection of the delegation mechanism and acquire further knowledge of how delegation is perceived amongst human agents. Notably, the scale is not tested, and the coefficient of determination is not calculated.

Hypotheses were tested with Pearson x^2 goodness of fit statistical test. In order to use this statistical test, all cells must contain at least five observations (Nummenmaa, 2021). Therefore, expected values are set so that this requirement is met. All null hypotheses are to be rejected if the observed frequency differs significantly from the expected frequency with alpha < .05. Hypotheses 2-4 were calculated through answers given by those participants who had selected the option “I have asked ChatGPT to execute work tasks to reach a desired outcome.” In other words, only those respondents who replied that they do ask ChatGPT to execute tasks are used in testing the latter hypotheses.

4 Results

The results presented in this section contribute to creating the conceptual model. The results of the directed content analysis contribute to all IS delegation framework guidelines by extending the theoretical framework to the context of ChatGPT (Hsieh & Shannon, 2005). Quantitative findings of the self-administered web survey further explore the interview results and how ChatGPT is perceived, and contribute to the conceptual model through hypothesis testing.

4.1 Interviews

Interview answers were analyzed through a directed content analysis. Directed content analysis (Hsieh & Shannon, 2005) is designed to depict a phenomenon with predefined categories, clusters, or codes. As both prior research and theory frameworks exist, the directed content analysis provides a way to use existing categories mentioned in the IS delegation framework (Hsieh & Shannon, 2005). Categories are 1) task attributes, 2) delegation mechanisms, 3) agent attributes, 4) situational attributes, and 5) feedback loops. Categories follow guidelines defined by Baird and Maruping (2021). An important notion is that interview results do not aim to provide any quantitative findings since the following web survey will elaborate on the quantitative perspective.

Task and agent attributes

Interview results consider the tasks that interviewees ask ChatGPT to execute, tasks that interviewees have observed could be delegated to ChatGPT, and tasks that interviewees would not ask ChatGPT to execute. All these tasks contribute to the conceptual model created based on the IS delegation framework, describing which tasks could be delegated and narrowing the scope by providing insight into tasks perceived as inexecutable. The interviewees mentioned a total of 86 different tasks that they had asked ChatGPT to execute. A comprehensive list of tasks that the interviewees have asked ChatGPT to execute is presented in Table 2.

Table 2 Executed tasks by interviewees.

Ideation
Text production (scripts for podcasts, blog texts, social media texts)
Writing pseudocode
Reading scientific articles on behalf of the user (copy-paste and summarization)
Conceptualization
Communication clarification
Translation work
Marketing/growth/revops functions
All writing tasks, including emails (provided there is no confidential information)
Text correction and formatting
Language correction and maintenance
Planning of coaching (e.g. basics of social media)
Generating funny aspects for stand-up material
Support for studying (writing essays, classifying cases etc.)
Suggestions for titles/headlines
Email text drafts
Social media post ideas
Video and podcast ideas
Assisting in structuring written messages
Building frames and structures (e.g. text based content such as tables)
Considering things [with ChatGPT] when dealing with a certain customer persona
Sparring in idea generation
Pitching support
Customer profile creation
Project planning
Creating risk analysis
Creating sales arguments
Simulating sales situations
Creating training programs
Creating plans
Creating board meeting memos
Writing newsletters
Leadership coaching (e.g. material creation for leadership coaching event)
Receiving and giving feedback
Writing summaries
Songwriting
Movie recommendations
Freestyle rap creation
Enhancing personal learning (e.g. summaries)
Language teaching (e.g. conversation)
Philosophical conversation practice
Developing emotional intelligence (e.g. conversation)

Analyzing reports and writing follow-up suggestions
Cold emailing and cold calling (e.g. writing drafts for emails and scripts calls)
Visualization (e.g. text-based prompt for other AI tool)
Personality type identification
Mathematical calculations
Creating Excel table frames
Creating business model canvas
Creating learning tasks for students
Paid advertising content planning
PowerPoint drafting (e.g. textually providing good arguments)
Creating basic content (e.g. blog writing, project description)
Applying for start-up funding (e.g. business plan, answering questions based on short descriptions)
Search Engine Optimization (e.g. text production)
Text improvement tasks
Playful tasks (telling a joke with certain elements, writing a short story for bedtime)
Creating text structures
Revising after creating the frame (e.g. checking texts in tables)
Creating individual things (e.g. a single text phrase)
Quick license term writing
Critical thinking and idea validation
Brainstorming
Creating a text description based on a picture
Combining two things or outlining a new perspective (e.g. in Service Design; text based task)
Creating writing drafts
Creating tasks for students
Checking understanding of a concept
Personality creation, company types creation
Supplementing existing classifications
Delegating source citation formatting
Gathering the three most important points from a larger text such as a research article
Summarizing of a research (writing a summary)
Excel sheets (summary of customer Excel files for example for PowerPoint)
Summaries (from briefing texts)
LinkedIn posts
Essays for a book
Travel plans
Smoothing out a text section
Metatext production for a certain amount of text
Looking for adjectives
Service productization: What do customers value?
PowerPoint creation (e.g. slide content)
Solving the blank page problem (e.g. by drafting a text)
Defining concepts and searching for concepts
Planning and project management

The list of tasks presented in Table 2 contributes directly to the first IS delegation guideline by providing a comprehensive list of tasks in which ChatGPT has already been utilized. The listing on a higher level can be categorized into linguistic, ideation, technical, summarization, communication, management, and educational tasks. Text production, language correction, and text correction are typical examples of linguistical tasks given to ChatGPT. Understanding and producing text is involved, at least indirectly, by all categories, but as a task, text production is by far the most common task. Ideation, conceptualization, idea creation, and creative writing tasks such as "creating a text description based on a picture" can overcome challenges such as the "blank page problem," also mentioned as a task given to ChatGPT. Technical assistance includes coding and mathematical calculations, which in many cases relate to Excel applications in the form of frames and tables. Summarization includes reading a text and providing a human agent summary about the read text. The text can be, for example, a scientific article and the task to read and write "three most important points of the article." Communication tasks may include material or message writing. Management tasks include, for example, project planning and business model creation. Educational tasks consider the study support and personal learning.

Table 3 Tasks which interviewees would not ask ChatGPT to execute.

Anything related to numbers (calculations, understanding magnitudes)
Tasks requiring up-to-date information
Tasks involving language genre or text structure
Any task that requires the use of a search engine or real information
Business decisions
Handling of confidential or sensitive information
Activities related to data retrieval
Legal cases or precise legal information
Tasks requiring a personal point of view or creativity
Personal matters or tasks involving personal information
Tasks involving real-time fact checking
Analysis of legal texts and regulations
Specific or NDA-bound tasks
Tasks that exceed a certain length (half an A4 page)
Tasks involving critical or difficult documents
Tasks dealing with trade secrets

Product development
Tasks that will be made public
Tasks for which the quality of training data is uncertain
Scientific research
Writing official documents
Creation of large text corpora (especially scientific research)
Writing personal and important emails
Handling confidential contracts
Tasks where careful consideration is needed on what to publish (press releases)
Grading essays or any form of evaluation
Fact-finding tasks
Article writing
Tasks that involve challenges with GDPR
Asking about things the user is not sufficiently familiar with
Producing fact-based text

Table 3 presents tasks that interviewees would not delegate to ChatGPT. According to these tasks, ChatGPT has specific limitations, such as understanding magnitudes, inability to execute calculations consistently, and retrieving up-to-date information. Almost all interviewees mentioned tasks involving sensitive and confidential information in the context of non-delegated tasks. Interviewees also had varying levels of trust in the reliability, decision-making, creativity, text generation, and research capabilities of ChatGPT.

Table 4 Tasks which could be given to ChatGPT by interviewees.

Interpretation or description of images
Use of an AI as a search engine
Maintenance of social media presence
Optimization of timber usage (calculation)
Use of multiple AI tools in a synergistic manner, for example in a social media strategy where one AI creates a strategy and others carry out tasks accordingly
Writing of emails
Writing of a standardized report
Drafting of decision letters (academic letters)
Providing prompts for daily tasks, for example, product launches
Automatic analysis of materials (data analysis)
Generation of text (prefers to write himself, potentially asking for help in choosing better phrasing)
Dealing with business secrets
Handling of personal information or sensitive data

When considering the agent attributes related to text production tasks, the interview results indicate that human agents acknowledge ChatGPT's capability to write a considerable amount of text fast and consistently without distracting violations of language rules. Amongst the interviewees, the ChatGPT has been utilized in versatile ways, but tasks such as "assisting" and "drafting" point to the fact that ChatGPT is seen in an assisting role.

Table 4 presents potential tasks that interviewees know they could delegate or give to ChatGPT but which they do not delegate or give to ChatGPT. Some of the interview responses in Tables 2 and 3 are contradictory. For example, calculations are mentioned in both tables. The calculations are mentioned once again in Table 4. The tasks that could be given to ChatGPT represent a middle ground between tasks that are given and tasks that should not be given. In Table 4, tasks have similarities between both Tables 2 and 3.

While none of the interview questions directly engage the delegation mechanisms, the task and agent attributes provide indications about the appraisal delegation mechanism. For instance, in a single case, a salient benefit of delegation could be overcoming writer's block, and one considers the risk that the IS agent may not be able to attain the desired outcome.

Situational attributes

Table 5 presents the situations when a human agent seeks to converse with ChatGPT. These answers contain feelings and needs that human agents have identified before starting to chat and write a prompt. The main reasons for starting a conversation, according to interview results, are a need for preparation or clarity, writer's block or fear of a blank page, pressure or time constraints, tedium or repetition, transition, comfort in anonymity, familiarity, seeking inspiration or second opinion, unfamiliar territory, learning and problem solving, curiosity, and fatigue or busyness. Conversation was often

started when human agents had to start something new, for example, ideate, clarify, or write. Another situation is when the human agent already has a piece of text and needs to do repetitive or tedious tasks such as translations, refining text, or writing again. Conversations were also started when there was a need to learn new things, explore new ideas, seek inspiration, or spar with a “colleague.” Interviewees also mentioned a need to write about both familiar and unfamiliar texts with ChatGPT.

Table 5 Situations when human agents interact with ChatGPT.

Need for preparation: The feeling of a need to prepare for tasks like sales or text composition.
Need for clarity: Seeking a sparring partner to clarify thoughts or when thoughts are stuck.
Experiencing writer's block: Feeling stuck and unable to produce written content.
Pressure and time constraints: Feeling stressed or under time pressure, a need to quickly understand a concept or start a creative writing process.
Fear of blank page: Experiencing "blank page syndrome" or anxiety over starting a piece of writing from scratch.
Tedium and repetition: Experiencing a task as tedious, repetitive or boring.
Transition: Feeling a need to refine text when transitioning from another task.
Comfort in anonymity: Situations where it is easier or more comfortable to ask ChatGPT a question rather than asking another person.
Familiarity: A need to quickly produce text that is familiar or known.
Seeking inspiration: The absence of inspiration, particularly when facing a blank page or starting a new task.
Unfamiliar territory: Facing tasks outside of one's own expertise.
Learning and problem-solving: Using the AI to learn new skills, such as programming basics, or to solve problems.
Seeking a second opinion: Wanting an external perspective on a piece of writing, particularly regarding its impact on a reader.
Curiosity: Using ChatGPT to explore an idea or to start a new project.
Fatigue or busyness: Feeling tired or overwhelmed, and using ChatGPT to assist with task completion.

Delegation feedback loops

Interviewees were asked to ponder whether they would say that ChatGPT delegates tasks to a human agent. Eight interviewees said they would not say that ChatGPT delegates tasks to them directly, but it sometimes has undirect tasks to human agents that must be done to proceed with the task. Also, interviewees mentioned that sometimes the workflow may have changed due to the suggestion by ChatGPT. None of the interviewees responded that ChatGPT delegates tasks to the interviewee.

4.2 Web survey

Background questions provide insight into the demographics of the respondents. 131 respondents had typed in their year of birth. The year of birth ranged between 1960 and 2002. The average year of birth was 1985,7, while the median was 1987. Therefore, the median age amongst the respondents at the time of the survey was 36 years. The standard deviation was 11,2. Of 131 respondents, the most common degree was a bachelor's degree (46.6% and 61 responses). The second most common was a master's degree (32,1% and 42 responses), the third was a joint high school/vocational degree (13,7% and 18 responses), and the fourth PhD degree (7,6% and 10 responses). Among the respondents were no respondents with elementary degree or respondents with no degree at all. 132 respondents responded to the question about their managerial or professional status. 39.4% (52) identified themselves as managerial, while 60.6% (80) declared themselves professional. 131 respondents answered the question, "Do you have subordinates?". 31,3% (41 responses) of the respondents say they have subordinates who report to them. 68.7% (90 responses) said that they do not have subordinates. 132 answered the question regarding the use of ChatGPT. 30.3% (40 responses) of the respondents use ChatGPT less than once each week. 24,2% (32 responses) say that they use ChatGPT about once each week. As many respondents (32 responses) find that they use ChatGPT several times each week. 9 respondents (6,8%) report using ChatGPT once daily, while 13,6% (18 responses) use ChatGPT several times daily. One respondent (0.8% of total respondents) reported not using ChatGPT.

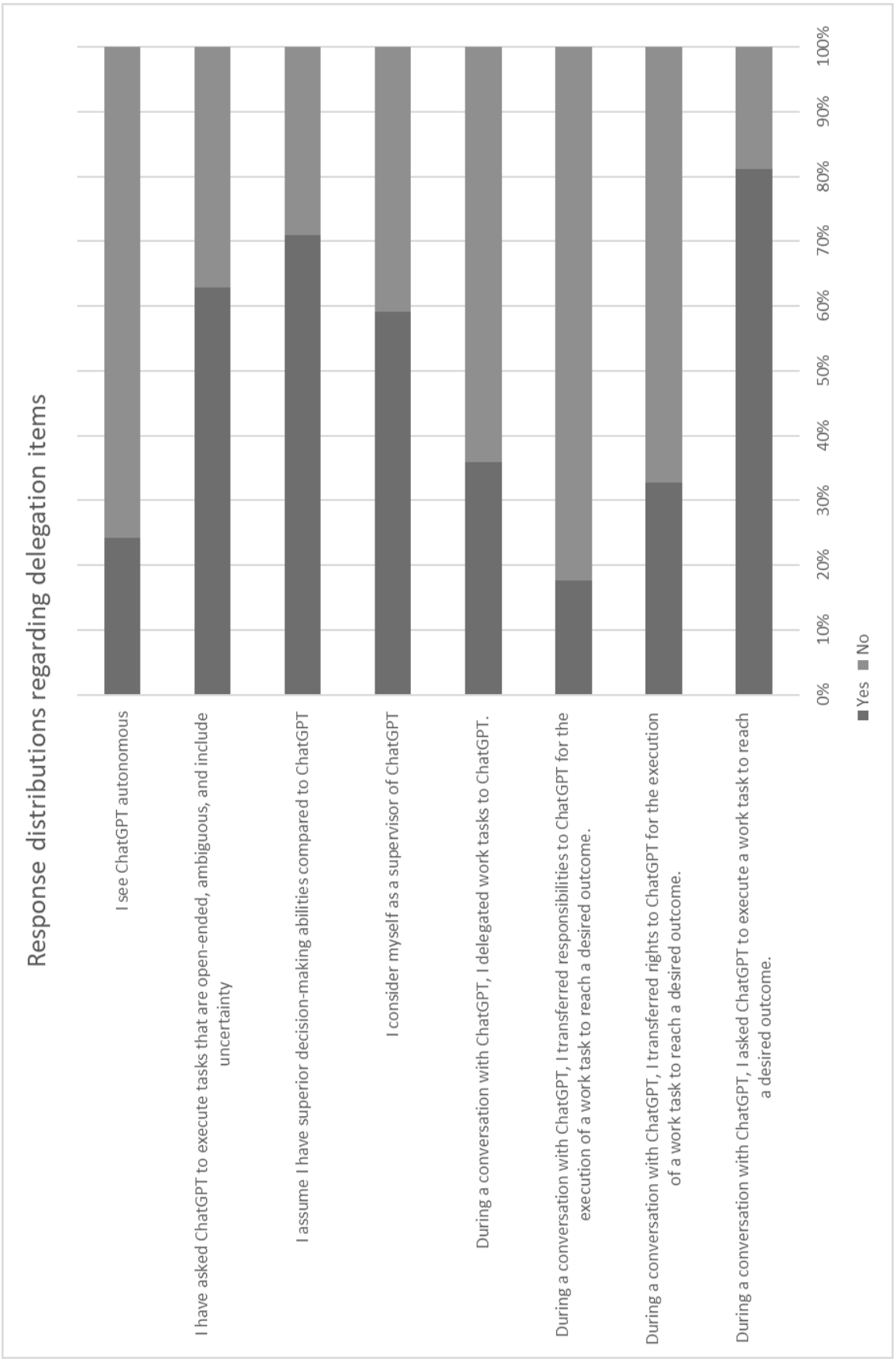

Figure 1 Response distribution regarding delegation items.

Figure 1 presents eight items regarding the delegation involving ChatGPT. The first four items are derived from Baird and Maruping's (2021, p. 320) illustration of "key differences between IS Use and IS delegation." Results reflect the respondents' perception of ChatGPT as an agentic IS artifact. The latter four items are part of a newly created measurement scale to understand better the subdimensions of delegation and the overall perception of delegation amongst the respondents.

132 respondents selected the most fitting option to the claim "I see ChatGPT autonomous." One hundred (100) respondents (75,8%) selected option "no." 32 respondents (24,2%) selected option "yes." 129 respondents selected the most fitting option to the claim: "I have asked ChatGPT to execute tasks that are open-ended, ambiguous, and include uncertainty." 81 respondents (62,8%) selected option "yes." Option "no" was selected by 48 respondents (37.2%). 132 respondents selected the most fitting option based on the claim, "I assume I have superior decision-making abilities compared to ChatGPT." 93 respondents (71%) selected option "yes". 38 respondents (29%) selected option "no". 132 respondents selected the most fitting option based on the claim, "I consider myself a supervisor of ChatGPT." 78 respondents (59,1%) selected option "yes" while 54 respondents (40,9%) selected option "no." As presented, the majority of respondents do not see ChatGPT autonomous, and assume they have superior decision-making abilities, and consider themselves as a supervisor of ChatGPT. When considering IS delegation framework and the comparison between IS use and IS delegation (Baird & Maruping, 2021, p. 320), respondents perceive ChatGPT more through IS use. This means most respondents view ChatGPT as a tool rather than an agentic IS artifact.

132 respondents selected one or more tasks they had asked ChatGPT to execute. All respondents selected a total of 1178 tasks. The most common task, "text production, e.g., essays, newsletters, etc." was selected by 101 respondents (76.5%). The task was followed by "Ideation" (90 respondents; 68,2% of total respondents), "Drafting any text" (84; 63,6%), "Writing summaries" (71; 53,8%), "Brainstorming" (69; 52,3%). Other cases gathered less than 50% of the respondents. By identifying the most common task, "text

production," the survey contributes to the conceptual model by providing a solid indication that text production could be used as a task to which the conceptual model's other attributes should be related.

129 respondents selected one or more situations in which they had asked ChatGPT to execute a task. All respondents selected a total of 739 situations. The most common situation when respondents had asked ChatGPT to execute a task was when they had to "prepare materials, e.g., presentation" (78 respondents; 60,5% of all respondents). The situation was followed by "I had to refine text" (75; 58,1%), "I wanted to get inspired" (74; 57,4%), and "I wanted to write quickly a text about a familiar topic (71; 55%), "I experienced writer's block or fear of blank page" (68; 52,7%), and "I wanted to solve a problem" (66; 51,2%). Other situations gathered less than 50% of the respondents' answers. The above-mentioned situations contribute to the conceptual model by providing indications of situational attributes.

Total of 132 respondents selected the most fitting option to the claim: "During a conversation with ChatGPT, I asked ChatGPT to execute a work task to reach a desired outcome. A task is a usually assigned piece of work often to be finished within a certain time. This question aims to understand whether you have asked ChatGPT to execute any task. The task can be a part of a larger task or a complete task." 107 respondents (81,1% of all) selected "yes." 25 respondents (18,9%) selected the option "no."

Total of 131 respondents selected a more fitting option to the claim: "During a conversation with ChatGPT, I transferred rights to ChatGPT for the execution of a work task to reach a desired outcome. The definition of *right* is the power or privilege to which one is justly entitled. This question aims to understand do you transfer your power or privilege to ChatGPT so that ChatGPT is entitled to work with that specific task." 88 respondents (67,2%) selected option "no." 43 respondents (32,8%) selected option "yes."

Total of 130 respondents selected the most fitting option to the claim: “During a conversation with ChatGPT, I transferred responsibilities to ChatGPT for the execution of a work task to reach a desired outcome. The definition of *responsibility* is moral, legal, or mental accountability. This question aims to understand do you transfer your moral, legal, or mental accountability to ChatGPT so that ChatGPT has been morally, legally, or mentally accountable for the task execution.” 107 respondents (82,3%) selected option “no.” 23 respondents (17,7%) selected option “yes.”

Total of 131 respondents selected the most fitting option to the claim: “During a conversation with ChatGPT, I delegated work tasks to ChatGPT. Delegation is an action between you and ChatGPT that transfers rights and responsibilities from you to ChatGPT for the execution of a specific task. This question aims to understand whether you have delegated work tasks to ChatGPT for any reason and do you transfer both rights and responsibilities to ChatGPT in order to reach the desired outcome”. 84 respondents (64,1%) selected option “no.” 47 respondents (35,9%) selected option “yes.”

4.3 Hypothesis testing

Figure 2 illustrates the tables used to calculate the goodness of fit. Only those respondents who replied that they had asked ChatGPT to execute a text production task are qualified to be considered in the first hypothesis testing. Note that only respondents who replied that they asked ChatGPT to execute work tasks and have selected the option that they have asked ChatGPT to execute a text production task are considered when calculating the goodness of fit. In other words, the aforementioned numbers present all the respondents, while the numbers in this section focus on a narrower segment of respondents. Also, one respondent only answered the task execution item. Therefore, task execution’s observed value is 86, but in the latter hypotheses, n=85. If one respondent had answered all items, the n in the latter hypotheses test would have been 86.

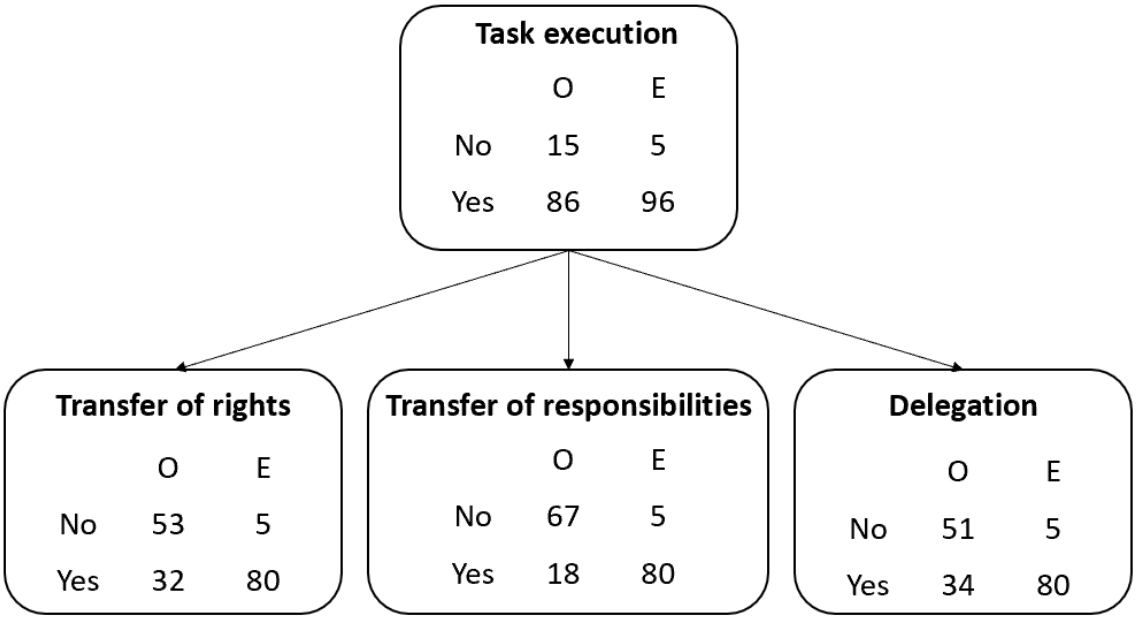

Figure 2 Observed (O) and Expected (E) values of each hypotheses.

1. Hypothesis: Asking ChatGPT to execute work tasks

For human agents who ask ChatGPT to execute work tasks, the observed value is 86, and the expected value is 96. For human agents who do not ask ChatGPT to execute work tasks, the observed value is 15, and the expected value is 5. A chi-square goodness of fit statistical test has been performed to examine the differences between observed and expected values. Differences between observed and expected values are statistically significant, X^2 (1, N=101) = 21.042, $p < .00001$. Therefore, there is evidence for rejecting the null hypothesis and accepting the alternative hypothesis that all human agents have not asked ChatGPT to execute work tasks. The result contributes to the conceptual model by identifying that in all interaction cases with ChatGPT, work tasks are not given in the form of text production.

2. Hypothesis: Transferring rights to ChatGPT

According to the literature, human agents transfer rights when they utilize the distribution or coordination delegation mechanisms. Therefore, the expected value should align with the number of respondents. Due to the limitations of the chi-square goodness of fit statistical test, the expected value is set as high as possible to 80. All cells must contain

at least five observations to calculate valid results with the statistical test. The goodness of fit statistical test was performed to test the null hypothesis of "human agents do not transfer rights to ChatGPT to reach a desired outcome." The observed value of respondents who selected the option "yes" at the item "…, I transferred rights to ChatGPT for the execution of the task to reach the desired outcome" is 32, while the expected value is 80. The observed value of respondents who selected the option "no" is 53, and the expected value is 5. Differences between observed and expected values were statistically significant, X^2 (1, N=85) = 489.6, p < .00001. Therefore, there is evidence for rejecting the null hypothesis and accepting the alternative hypothesis. The result contributes to the conceptual model by identifying that all human agents do not transfer rights when asking ChatGPT to execute a text production task in a work setting.

3. Hypothesis: Transferring responsibilities to ChatGPT

Literature suggests that human agents do transfer responsibilities to agentic IS artifacts during the action of delegation. To test the third hypothesis, do human agents transfer responsibilities to ChatGPT, the survey included the item "…I transferred responsibilities to ChatGPT for the execution of work task to reach the desired outcome". The chi-square goodness of fit test was performed to examine the difference between the observed and expected values, which reflect the respondents' views towards the item. For the respondents who selected the option "yes," the observed value is 18, and the expected value is 80. For the respondents who selected the option "no," the observed value is 67, and the expected value is 5. Differences between observed and expected values were statistically significant, X^2 (1, N=85) = 816.85, p < .00001. Therefore, there is evidence to reject the null hypothesis and accept the alternative hypothesis that human agents do not transfer responsibility to ChatGPT when asking to execute a text production task. The result contributes to the conceptual model by suggesting that the transfer of responsibilities is human agents do not transfer responsibilities in the case of work-related text production.

4. Hypothesis: Delegating work tasks to ChatGPT

To test whether human agents delegate work tasks to ChatGPT, the null hypothesis of "Human agents do delegate work tasks to ChatGPT" is set. The item for measuring the observations is "During a conversation with ChatGPT, I delegated work tasks to ChatGPT." The literature suggests that human agents delegate tasks to agentic IS artifacts, as in the previous two hypotheses. Therefore, expected values are set in the same manner as in previous hypotheses. The expected frequency for human agents who delegate is 80, and for those who do not delegate, the expected frequency is 5. The observed frequency for human agents who delegate is 34, and for those who do not 51. Differences between observed and expected values were statistically significant, X^2 (1, N=85) = 449,65, p < .00001. Therefore, there is evidence for rejecting the null hypothesis and accepting the alternative hypothesis that human agents do not delegate, by definition, text production tasks to ChatGPT in a work setting.

4.4 Summary

Interview and survey findings align and clearly depict the human agents' perception of the interaction between ChatGPT and human agents. Human agents acknowledge the capabilities of ChatGPT and are willing to ask ChatGPT to execute assisting tasks. The most common task, according to survey results, is text production. On the other hand, according to interview results, interviewed human agents preserve some of the tasks through cognitive and emotional appraisal since they would not ask ChatGPT to execute them even when it is technically possible. This section's survey and hypothesis testing results are compatible with the IS delegation framework.

5 Discussion

The conceptual model in the context of ChatGPT is presented in this section. The conceptual model builds on the interview, survey, and hypothesis testing results. The conceptual model answers the original research question by demonstrating delegation mechanism appraisal, closely related to the most common task text production. In addition, the conceptual model reviews the agent and situational attributes and feedback loops between agents.

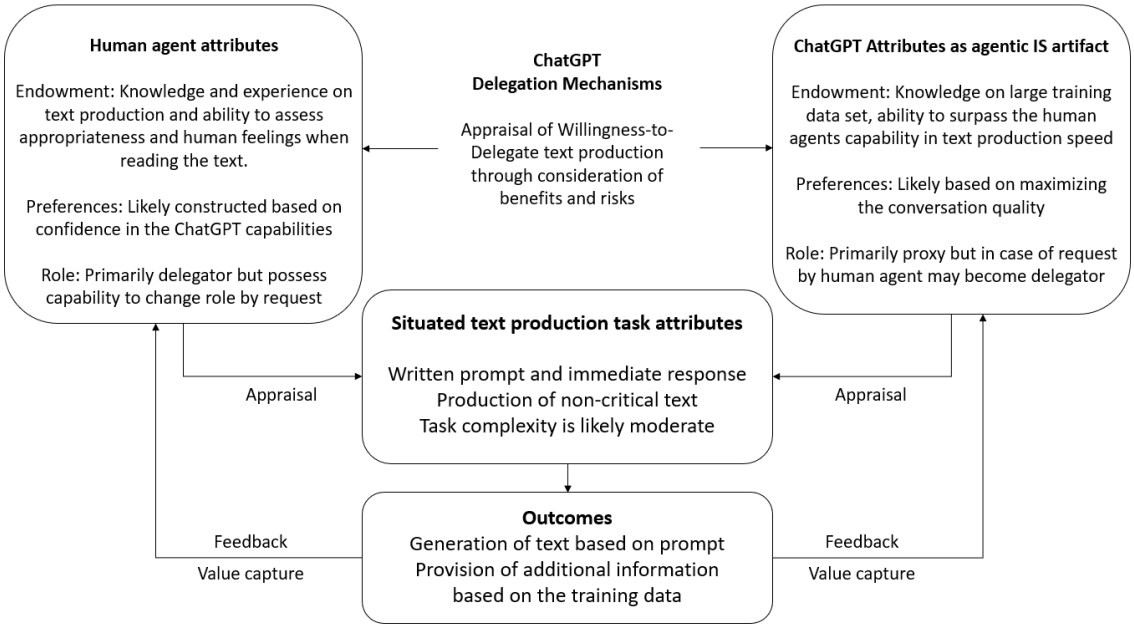

Figure 3 Conceptual model for delegation in the context of ChatGPT.

Figure 3 presents the Baird & Maruping (2021) conceptual model brought into the context of ChatGPT. According to Baird and Maruping (2021), the conceptual model applies the IS delegation framework. In this thesis, the conceptual model answers the research question by examining the appraisal process and the related attributes affecting the decision to delegate. The discussion section follows the guideline order suggested by the IS delegation framework (Baird & Maruping, 2021).

Task attributes

Baird and Maruping (2021) propose three task attributes in their IS delegation framework: action requirements, complexity, and decomposability. Action requirement types define whether the action involves thought (cognitive), whether the action exists digitally (digital), or whether the action is physical (physical). The complexity of a task could be explained through how much effort the agent needs to put into completing the task. Uncertain, dynamic, and interdependent tasks require more effort from the agent. The potential of dividing a task into smaller subtasks is called decomposability (Baird & Maruping, 2021).

According to interview results, the nature of the task (guideline #1) is usually a free-form text-based prompt in a digital environment. Delegated tasks are usually subtasks where a human agent is incentivized to produce non-critical text or text to add to the planned text. Therefore, the task is already divided into subtasks, and effort has been put into completing the task. The complexity of a text production task is primarily moderate. Tasks are straightforward text production and utilize the agent's training data. This said, the complexity could be elevated through more complex and abstract topics.

Delegation mechanism

As Baird and Maruping (2021) identified, delegation mechanisms represent the delegation process. According to interview findings, the appraisal is the most appropriate delegation mechanism for human agents, while ChatGPT acts as a proxy (guideline #2). Human agents acknowledge the potential risks and have determined boundaries of how much they trust ChatGPT’s capabilities and judgment. A distinct example of a cognitive boundary that surfaced in the interviews are legal tasks, for example, tasks containing NDA or business secret. Examples of emotionally inclined boundaries are tasks that include personal information or personal touch, such as writing an email or updating social media.

According to interview and survey results, other delegation mechanisms proposed by Baird and Maruping (2021) do not qualify as a primary delegation mechanism in the context of ChatGPT. Both distribution and coordination as a delegation mechanism require the transfer of rights and responsibilities and possibly negotiation and regulation. As presented in the results section, both hypotheses regarding the transfer of rights and responsibilities were rejected, meaning human agents do not consistently transfer rights or responsibilities to ChatGPT when asking ChatGPT to execute a text production task. Therefore, appraisal is the only viable delegation mechanism when considering the interaction between ChatGPT and human agents.

Based on the interviews, appraisal mainly includes assessing the benefits and risks and slightly less evaluation of liabilities, contingencies, coordination effort, and distribution of task execution rights. According to interview and survey results, human agents are hesitant to acknowledge the agentic nature of ChatGPT and preserve tasks for themselves through cognitive and emotional appraisal. This means that human agents do not transfer the rights and responsibilities in a magnitude that liabilities, coordination, and distribution of task execution rights would be necessary. While appraisal might be, for now, the primary delegation mechanism, the human agents may accept the subdimensions of delegation later in the context of ChatGPT. As Baird & Maruping (2021, p. 331) describe, confidence and trust towards agentic IS artifact is established over "repeated interaction."

Agent attributes

According to Baird & Maruping (2021, p. 322), human agents and agentic IS artifacts have at least three foundational, "relevant," and "pervasive" attributes (guideline #3): endowments, preferences, and roles. Endowments are an inseparable part of an agent and demonstrate an agent's "resource-based assets and capabilities" (p. 324). In delegation, these assets and capabilities of one agent are used to execute a task. Preferences are

motivations that could be dismantled into goals and decision models. Goals are cognitive desired outcomes, and decision models are “internalized representations of how choices will be ranked in choice sets” (p. 325). The roles of delegator and proxy describe the relationship status between agents after the delegation is accepted (Baird & Maruping, 2021).

According to interviews, the most salient endowments for a human agent are knowledge and experience in text production and the capability to assess the text’s correctness and reception. Both interview and survey results strongly emphasize the superiority of human agents’ decision-making capabilities, ability to assess generated text, and knowledge to check facts on behalf of the ChatGPT. On the other hand, ChatGPT endows the capability to write sufficient text promptly, surpassing human agents’ slow pace of text production and even potentially avoiding human agent procrastination towards text production. This fast text production creates a preference for human agents. For instance, the task of creating a text draft immediately is easy to delegate to the agent who has superior knowledge harnessed through training data. In this case, delegation provides human agent insights into topics drawn from training data and potentially text that could be used in a main task. As Baird and Maruping (2021) have stated, there is an asymmetry between agents, and both agents can produce text on their own if a topic is given. However, only human agents can initiate the production. While only a human agent can initiate text production through a prompt, a human agent naturally adopts the role of a delegator.

Situational attributes

According to Baird & Maruping (2021, p. 327), situational attributes (guideline #4) influence the choice of delegation through “situational incentive and complexity.” Incentives, such as increased productivity in a work setting, may encourage human agents to delegate. Complexity has three primary characteristics: stability, observability, and controllability. Stability points to how consistently certain inputs produce expected outputs.

Observability of the delegation situation refers to the transparency of the system's current state. Controllability is the characteristic of "how much regulatory power agents have over the situation" when influencing the inputs to reach the desired output (Baird & Maruping, 2021, p. 327).

Delegation situations vary from formal to informal and between important to leisure activities, even in a work context. Most interviewees expect immediate text-based outcomes as a response to the instructions given through written prompts. Immediate response is an incentive to delegate tasks since it increases productivity in a work setting (Baird & Maruping, 2021). Interview answers point that the consistency is task-relevant, but ChatGPT performs on a satisfactory level with text production tasks when considering the survey results. The immediate response of ChatGPT could also be described as an observability characteristic for human agents, while complete controllability of ChatGPT is attained through a customizable input prompt. The above-mentioned incentives can be seen in the survey result of the most common situation, "I had to prepare materials," where the incentive is to prepare materials effectively. ChatGPT also provides the possibility to ask for specific desired output and intervene if necessary.

Feedback loops

According to Baird & Maruping (2021), feedback loops are informative in dynamic and temporal situations. Feedback loops provide information to both agents regardless of the role. For example, feedback loops may inform or alter the decision of delegator whether to delegate or provide information to proxy how to improve on task execution (Baird & Maruping, 2021). Feedback loops are important to both human agents and ChatGPT since delegation situations are temporal between ChatGPT and human agents (Baird & Maruping, 2021). According to interviews, text produced by ChatGPT is seen implicitly as feedback for the prompt, and interviewees see the produced text as a direct result of their prompt. ChatGPT, as an agentic artifact, uses prompts as feedback for text production, and human agents can directly communicate through prompts.

6 Conclusion

This thesis aims to answer the research question of how human agents delegate work tasks to agentic IS artifacts and utilizes qualitative and quantitative methods to gather relevant information for the conceptual model with semi-structured interviews and self-administered web survey. The thesis provides practical implication for human agents utilizing ChatGPT and two contributions towards the delegation.

The thesis contributes to the theory of IS delegation (Baird & Maruping, 2021) by replicating and extending the conceptual model of IS delegation to a new agentic IS artifact of ChatGPT. The practical implication is identifying the delegation mechanism between human agents and ChatGPT in the work setting. Also, to answer the original research question, human agents delegate tasks through the appraisal delegation mechanism, which represents the delegation process (Baird & Maruping, 2021). Practical implication enables human agents to understand delegation as a process when working with ChatGPT by providing new empirical insight.

The thesis also has two contributions to how human agents perceive delegation. First, survey results suggest that human agents do not acknowledge the action of delegation when interacting with ChatGPT. In survey results, most human agents replied that they do ask ChatGPT to execute tasks, but more than half of the survey respondents answered that they do not delegate tasks to ChatGPT. Second, human agents do not acknowledge all the subdimensions of delegation, such as transfer of rights and transfer of responsibilities, when interacting with ChatGPT. While 107 human agents responded that they had asked ChatGPT to execute a work task, only 23 of them responded that they do transfer responsibilities for the execution of a work task, and 40 of them responded that they do transfer rights in order to execute a work task. In hypothesis testing, the proportions of respondents reasonably matched survey results when a specific task, text production, was considered.

Limitations

When considering the measurement scale creation, due to the scope of this thesis, the aim is not to create a validated measurement scale but to follow MacKenzie et al. (2011) steps to generate items to represent the construct. Due to the scale is not validated, the results should be treated as preliminary and used in future research as a starting point in scale validation. Also, scale is not validated in this thesis as a whole, but scale items are tested through hypotheses to verify the literature expectations.

Acknowledging the ontological challenges with the formative measurement models is also essential. MacKenzie et al. (2011) cite Borsboom (2005, p. 63) when they state that "latent variable theory is ontologically ambiguous". Due to the nature of this theory, several assumptions are made during the creation of the construct of IS delegation. As a formative indicator is used as a measurement in this thesis, the measurement model presented does not assume that the construct of IS delegation is a real entity but a theoretical one. Even though delegation is naturally measurable and observable action, the subdimensions of the construct are not. Therefore, it has been a requisite to create formative measurements for this thesis, which do not rely on the realist assumptions that "must always present a real, mind-independent entity" (MacKenzie et al., 2011, p. 303).

While Fügener et al. (2022) state that delegation efficiency relates to the knowledge about the delegated task, it is essential to note that items are not created based on specific tasks in this thesis. This is due to the formative measurement where this study, as part of understanding how employees delegate tasks, to understand whether the employees see the transfer of rights and responsibilities as part of the delegation.

Acknowledgments

The thesis has been subjected to an AI-assisted linguistic evaluation of the English language. Grammarly was employed to check syntactic structure, punctuation, and lexical choices.

References

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Delgado, J. M. D., Bilal, M., ... & Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering*, *44*, 103299. https://doi.org/10.1016/j.jobe.2021.103299

Aleksander, I. (2017). Partners of humans: a realistic assessment of the role of robots in the foreseeable future. *Journal of Information Technology*, *32*(1), 1-9. https://doi.org/10.1057/s41265-016-0032-4

Alonso, R., & Matouschek, N. (2008). Optimal delegation. *The Review of Economic Studies*, *75*(1), 259-293. https://doi.org/10.1111/j.1467-937X.2007.00471.x

Ambrus, A., Baranovskyi, V., & Kolb, A. (2021). A delegation-based theory of expertise. *American Economic Journal: Microeconomics*, *13*(4), 373-419. http://doi.org/10.1257/mic.20190008

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS quarterly*, 45(1). https://doi.org/10.25300/MISQ/2021/15882

Banford, C. G., Buckley, M. R., & Roberts, F. (2014). Delegation revisited: how delegation can benefit globally-minded managers. *International Journal of Physical Distribution & Logistics Management*, *44*(8/9), 646-654. https://doi.org/10.1108/IJPDLM-07-2013-0191

Bawack, R., Fosso Wamba, S., & Carillo, K. D. (2019). From IT to AI Artifact: Implications for IS Research on AI Adoption and Use. Retrieved 25.6.2023. https://aisel.aisnet.org/digit2019/6

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A pretrained language model for scientific text. https://doi.org/10.48550/arXiv.1903.10676

Bengio, Y. (2008). Neural net language models. *Scholarpedia*, *3*(1), 3881. https://doi.org/10.4249/scholarpedia.3881

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS quarterly*, *45*(3). https://doi.org/10.25300/MISQ/2021/16274

Browne, R. (8.2.2023). All you need to know about ChatGPT, the A.I. Chatbot that's got the world talking and tech giants clashing. CNBC. Retrieved 18.3.2023. https://www.cnbc.com/2023/02/08/what-is-chatgpt-viral-ai-chatbot-at-heart-of-microsoft-google-fight.html

Burton-Jones, A., Stein, M., and Mishra, A. (2020). MIS Quarterly Research Curation: IS Use, MIS Quarterly. https://www.misqresearchcurations.org/blog/2017/12/1/is-use

Borsboom, D. (2005). Measuring the mind: Conceptual issues in contemporary psychometrics. *Cambridge University Press*. https://doi.org/10.1017/CBO9780511490026

Castelfranchi, C., & Falcone, R. (1998). Towards a theory of delegation for agent-based systems. *Robotics and Autonomous systems*, 24(3-4), 141-157. https://doi.org/10.1016/S0921-8890(98)00028-1

Chowdhury, G. G. (2003). Natural language processing. Annual Review of Information Science and Technology, 37(1), 51-89. https://doi.org/10.1002/aris.1440370103

Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. In Machine learning techniques for multimedia: case studies on organization and retrieval (pp. 21-49). *Berlin, Heidelberg: Springer Berlin Heidelberg*. https://doi.org/10.1007/978-3-540-75171-7_2

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340. https://doi.org/10.2307/249008

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. https://doi.org/10.48550/arXiv.1810.04805

Dobrev, D. (2012). A definition of artificial intelligence. https://doi.org/10.48550/arXiv.1210.1568

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research,

practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

European Parliament. (14.6.2023). EU AI ACT: First regulation on artificial intelligence. Retrieved 6.9.2023. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Fernández-Loría, C., Provost, F., & Han, X. (2022). Explaining data-driven decisions made by AI systems: the counterfactual approach. *MIS Quarterly* Vol. 46 No. 3 pp. 1635-1660. https://doi.org/10.48550/arXiv.2001.07417

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive challenges in Human–Artificial Intelligence Collaboration: Investigating the path toward productive del egation. *Information Systems Research*, 33(2), 678-696. https://doi.org/10.1287/isre.2021.1079

Glaser, B. G., & Strauss, A. L. (2017). Discovery of grounded theory: Strategies for qualitative research. Routledge. https://doi.org/10.4324/9780203793206

Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., ... & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: definitions, applications, and future directions. *Current reviews in musculoskeletal medicine*, 13, 69-76. https://doi.org/10.1007/s12178-020-09600-8

Holmström, B. R. (1978). On Incentives and Control in Organizations. Stanford University. Retrieved 24.7.2023. https://www.proquest.com/openview/faa27e671ef06cabfe1d6e38f817d2a7/1?pq-origsite=gscholar&cbl=18750&diss=y

Holmstrom, B., & Milgrom, P. (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization*, 7(special_issue), 24-52. https://doi.org/10.1093/jleo/7.special_issue.24

Hove, S. E., & Anda, B. (2005, September). Experiences from conducting semi-structured interviews in empirical software engineering research. *In 11th IEEE International Software Metrics Symposium* (METRICS'05) (pp. 10-pp). IEEE. https://doi.org/10.1109/METRICS.2005.24

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Hu, C. (2.2.2023). ChatGPT sets record for fastest-growing user base – analyst note. Reuters. Retrieved 18.3.2023. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Igbaria, M., Meredith, G., & Smith, D. C. (1995). Career orientations of information systems employees in South Africa. *The Journal of Strategic Information Systems*, *4*(4), 319-340. https://doi.org/10.1016/0963-8687(95)80002-8

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Unsupervised learning. In *An Introduction to Statistical Learning: with Applications in Python* (pp. 503-556). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-38747-0_12

Kok, J. N., Boers, E. J., Kosters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence*, *1*, 270-299. Retrieved 5.9.2023. https://www.eolss.net/sample-chapters/c15/e6-44.pdf

Kuegler, M., Smolnik, S., & Kane, G. (2015). What's in IT for employees? Understanding the relationship between use and performance in enterprise social software. *The Journal of Strategic Information Systems*, *24*(2), 90-112. https://doi.org/10.1016/j.jsis.2015.04.001

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Leyer, M., & Schneider, S. (2019). Me, you or AI? How do we feel about delegation. Association for Information Systems. In *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm & Uppsala, Sweden, June 8-14, 2019. ISBN 978-1-7336325-0-8 Research Papers. Retrieved 24.7.2023. https://aisel.aisnet.org/ecis2019_rp/36

Li, J., Li, M., Wang, X., & Thatcher, J. B. (2021). Strategic Directions for AI: The Role of CIOs and Boards of Directors. *MIS quarterly*, *45*(3). https://doi.org/10.25300/MISQ/2021/16523

Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H., & Liu, T. Y. (2022). BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, *23*(6), bbac409. https://doi.org/10.1093/bib/bbac409

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 293-334. https://doi.org/10.2307/23044045

Marshall, C., Brereton, P., & Kitchenham, B. (2015, April). Tools to support systematic reviews in software engineering: a cross-domain survey using semi-structured interviews. In *Proceedings of the 19th international conference on evaluation and assessment in software engineering* (pp. 1-6). https://doi.org/10.1145/2745802.2745827

Marshall, B., Cardon, P., Poddar, A., & Fontenot, R. (2013). Does sample size matter in qualitative research? A review of qualitative interviews in IS research. *Journal of computer information systems*, *54*(1), 11-22. https://doi.org/10.1080/08874417.2013.11645667

Mazaheri, E., Lagzian, M., & Hemmat, Z. (2020). Research directions in information systems field, current status and future trends. *Australasian Journal of Information Systems*, 24. https://doi.org/10.3127/ajis.v24i0.2045

Merriam-Webster. (2023a). Right. Merriam-Webster, Incorporated. Retrieved 24.7.2023. https://www.merriam-webster.com/dictionary/right

Merriam-Webster. (2023b). Responsibility. Merriam-Webster, Incorporated. Retrieved 24.7.2023. https://www.merriam-webster.com/dictionary/responsibility

Merriam-Webster. (2023c). Task. Merriam-Webster, Incorporated. Retrieved 24.7.2023. https://www.merriam-webster.com/dictionary/task

Merriam-Webster. (2023d). Action. Merriam-Webster, Incorporated. Retrieved 24.7.2023. https://www.merriam-webster.com/dictionary/action

Miller, C. A., Shaw, T., Emfield, A., Hamell, J., deVisser, E., Parasuraman, R., & Musliner, D. (2011, September). Delegating to automation: Performance, complacency and bias effects under non-optimal conditions. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 95-99). Sage CA: Los Angeles, CA: Sage Publications. https://doi.org/10.1177/1071181311551020

Mitchell, T. M. (1997). Machine learning. ISBN: 0070428077

Nilsson, N. J. (2009). The quest for artificial intelligence. Cambridge University Press. Retrieved 5.9.2023. https://dl.acm.org/doi/abs/10.5555/1667096

Nummenmaa, L. (2021). Tilastotieteen käsikirja. Werner Söderström Ltd. ISBN: 978-952-04-0138-2

Padmanabhan, B., Sahoo, N., & Burton-Jones, A. (2022). Machine learning in information systems research. *Management Information Systems Quarterly*, *46*(1), iii-xix. Retrieved 5.9.2023. https://misq.umn.edu/misq/downloads/download/editorial/752/

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Retrieved 6.6.2023. https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

Ricci, L., Lanfranchi, J. B., Lemetayer, F., Rotonda, C., Guillemin, F., Coste, J., & Spitz, E. (2019). Qualitative methods used to generate questionnaire items: a systematic review. *Qualitative health research*, *29*(1), 149-156. https://doi.org/10.1177/1049732318783186

Ritala, P., Ruokonen, M., & Ramaul, L. (2023). Transforming boundaries: how does ChatGPT change knowledge work? *Journal of Business Strategy*. https://doi.org/10.1108/JBS-05-2023-0094

Russell, S. J. & Norvig, P. (2010). Artificial intelligence a modern approach. Pearson Education, Inc. ISBN-13: 978-0-13-604259-4

Samoili, S., López-Cobo, M., Gómez, E., De Prato, E., Martinez-Plumed, F., & Delipetrev, B. (2020). AI Watch Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence. European Comission, Joint Research Center, Publications Office. https://doi.org/10.2760/382730

Saunders, M., Lewis, P., & Thornhill, A. (2019). Research methods for business students. Pearson education. ISBN: 978-0-273-71686-0

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neuralnet-works*, *61*, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology*, 307(2), e230163. https://doi.org/10.1148/radiol.230163

Shahriar, S., & Hayawi, K. (2023). Let's have a chat! A Conversation with ChatGPT: Technology, Applications, and Limitations. https://doi.org/10.48550/arXiv.2302.13817

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information systems*, 13(1), 24. https://doi.org/10.17705/1CAIS.01324

Straub, D. W., Gefen, D., Recker, J. (25.3.2022). "Quantitative Research in Information Systems," Association for Information Systems (AISWorld) Section on IS Research, Methods, and Theories. Retrieved 19.10.2023. http://www.janrecker.com/quantitative-research-in-information-systems/

Sturm, T., Gerlach, J. P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., ... & Buxmann, P. (2021). Coordinating Human and Machine Learning for Effective Organizational Learning. *MIS quarterly*, 45(3). https://doi.org/10.25300/MISQ/2021/16543

Tiainen, T. (2014). Haastattelu tietojenkäsittelytieteen tutkimuksessa. Retrieved 19.10.2023. https://trepo.tuni.fi/bitstream/handle/10024/96052/haastat-telu_tietojenkasi___ttelytieteen_2014.pdf?sequence=1

Tutkimuseettinen neuvottelukunta (TENK). (2012). Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsittely Suomessa. Retrieved 26.6.2023. https://tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf

Turing, A. M. (1950). Computing machinery and intelligence (pp. 23-65). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6710-5_3

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information*

processing systems, *30*. Retrieved 6.6.2023. https://dl.acm.org/doi/abs/10.5555/3295222.3295349

Yang, L., & Wang, J. (2023). Factors influencing initial public acceptance of integrating the ChatGPT-type model with government services. *Kybernetes*. Retrieved 17.10.2023. https://www.emerald.com/insight/content/doi/10.1108/K-06-2023-1011/full/html

OpenAI. (2023a). Terms of use. Retrieved 6.6.2023. https://openai.com/policies/terms-of-use

OpenAI. (2023b). Planning for AGI and beyong. Retrieved 6.6.2023. https://openai.com/blog/planning-for-agi-and-beyond.

Appendix 1. Survey results: Tasks

	n	Percent
Ideation	90	68,2%
Text production e.g. essays, newsletters, etc.	101	76,5%
Writing pseudocode	17	12,9%
Studying or enhansing personal learning e.g. summarising scientific articles	50	37,9%
Conceptualization e.g. defining a concept	44	33,3%
Communication e.g. writing emails	50	37,9%
Translations and refining language	62	47,0%
Planning	50	37,9%
Drafting any text	84	63,6%
Sparring	38	28,8%
Coding	38	28,8%
Customer profile creation	10	7,6%
Project planning	25	18,9%
Creation of risk analysis	12	9,1%
Creation of sales arguments	35	26,5%
Simulating sales situations	6	4,5%
Creation of training programs	12	9,1%
Leadership coaching	8	6,1%
Receiving and giving feedback	16	12,1%
Writing summaries	71	53,8%
Language teaching	12	9,1%
Philosophical conversation	19	14,4%
Analyzing reports	23	17,4%
Visualization	17	12,9%
Personality type identification	6	4,5%
Mathematical calculations	29	22,0%
Creating excel table frames	18	13,6%
Creating business model canvas	13	9,8%
Creating learning tasks for students	12	9,1%
Paid advertising content planning	10	7,6%
Powerpoint drafting	25	18,9%
Applying for start-up funding e.g. writing business plan	8	6,1%
Search engine optimization	15	11,4%
Writing licence term	6	4,5%
Critical thinking and idea validation	26	19,7%
Brainstorming	69	52,3%
Asking to describe a picture	11	8,3%
Citation formulation	5	3,8%
Looking for adjectives	21	15,9%
Service productization	10	7,6%
Other	4	3,0%

Appendix 2. Survey results: Situations

	n	Percent
I had to prepare materials e.g. presentation	78	60,5%
I had a need for sparring partner	48	37,2%
I have experienced writer's block or fear of blank page	68	52,7%
I have experienced stress or been under time pressure	46	35,7%
I had to do tedious, repetitive or boring task	51	39,5%
I had to refine text	75	58,1%
I have been too embarrassed to ask specific question from another person e.g. during a meeting	18	14,0%
I wanted to write quickly a text about familiar topic	71	55,0%
I wanted to get inspired	74	57,4%
I faced a task which was unfamiliar to me	41	31,8%
I wanted to learn a new skill e.g. programming	20	15,5%
I wanted to solve a problem	66	51,2%
I wanted to explore a new idea e.g. for new project	48	37,2%
I felt myself tired or overwhelmed	31	24,0%
Other	4	3,1%