SCHOOL OF
ECONOMICS AND
MANAGEMENT

LUND UNIVERSITY

# AI and Strategic Decision Making Processes

*An empirical study on how Large Language Models can be utilized in Strategic Decision Making Processes*

by

Louise Hed Dahl & Johan Holmqvist

May 2023

Master's Program in International Strategic Management

BUSN09 – Degree Project 2023

**Supervisor:**

Ulf Ramberg

# Abstract

**Title:** AI and Strategic Decision Making Processes - An empirical study on how Large Language Models can be utilized in Strategic Decision Making Processes

**Seminar date:** June 2, 2023

**Authors:** Louise Hed Dahl, Johan Holmqvist

**Supervisor:** Ulf Ramberg

**Course:** BUSN09 – Degree Project in Strategic Management

**Keywords:** strategic decision making processes, artificial intelligence, large language models, ChatGPT, data privacy

**Purpose:** The purpose of this study is to contribute with an understanding of what role AI has, and might have, in strategic decision making processes.

**Theoretical framework:** The literature review of the study builds upon research within Resource Based Theory, Strategic Decision Making Theory and is supported by technical reports on Large Language Models.

**Methodology:** A qualitative single case study with abductive reasoning was chosen to fulfill the purpose of the study.

**Empirical foundation:** The empirical data was collected from six semi-structured interviews with professionals in knowledge-intensive firms from various consultancy and professional service firms.

**Conclusion:** The empirical results showed that LLMs are utilized mainly in the early stages of the decision making process due to the novelty of the technology and it being encountered with several challenges, hindering a wide application of LLMs in firms. It also indicates that there is an increased demand in seeking consulting for strategic solutions supported by LLMs, either in internal processes or as a product.

## Acknowledgements

To begin with, we would like to thank our supervisor, Ulf Ramberg, for challenging our ideas and supporting the creation of this thesis. We would also like to express our gratitude to our participants who made this case study possible by setting aside valuable time to contribute to this research. Additionally, we would like to give thanks to those who have assisted us with coming in contact with our participants.

A special thanks is directed to Robin, who has pedagogically and generously guided us through the vast jungle of Artificial Intelligence.

Lastly but not least, we would like to thank our friends and families who have supported us with encouragement throughout the entire project.

Lund, May 26th, 2023

Louise Hed Dahl　　　　Johan Holmqvist

## Abbreviations

**VUCA** - Volatile, Uncertain, Complex, Ambiguous

**AI** - Artificial Intelligence

**LLM** - Large Language Models

**ANI** - Artificial Narrow Intelligence

**AGI** - Artificial General Intelligence

**ASI** - Artificial Super Intelligence

**KIF** - Knowledge-Intensive Firms

**PSF** - Professional Service Firms

**RPA** - Robotic Process Automation

# Table of contents


## 1. Introduction

1.1. Background ........ 1
1.2. Problematization ........ 3
1.3. Purpose and Research Question ........ 5

## 2. Literature Review

2.1. Resource Based Theory ........ 6
2.2. Strategic Decision Making Theory ........ 9
2.2.1. The Strategic Decision Making Process 10
2.3. Interplay of Humans and AI in Decision Making ........ 14
2.3.1. Three Approaches to Decision Making 15
2.3.2. Ten Characteristics Determining the Approach 15
2.4. AI History and Overview ........ 17
2.4.1. The Structure Under ChatGPT-3 18
2.5. Novel Large Language Models ........ 20
2.5.1. GPT-4 Technical Report 20
2.5.2. Microsoft Testing ChatGPT-4 21
2.5.3. Open Letters and AI Legislation 22
2.6. Summary and Conceptual Model ........ 23

## 3. Metodology

3.1. Research Design ........ 26
3.2. Method of Data Collection ........ 27
3.2.1. Literature Collection 27
3.2.2. Primary Data Collection 27
3.3. Data Analysis ........ 30
3.4. Validity and Reliability ........ 31

## 4. Empirical Findings and Analysis

4.1. Case Description ........ 33
4.2. Interest and Potential Use ........ 33
4.2.1. Increased Interest Among Clients 33
4.2.2. Using AI Strategically 35
4.2.3. Analyzing the Situation of Clients 36
4.3. Generating Solutions ........ 38
4.3.1. Looking for Solutions 38
4.3.2. Creating Solutions 39

4.4. Artificial Intelligence and Business Insights 40
4.5. AI Challenges 41
4.5.1. Data Privacy 42
4.5.2. Technological Capability 43
4.5.3. Other Challenges 44
**5. Discussion**
5.1. Accessible LLMs as a Stimuli 47
5.1.1. Stimuli 47
5.1.2. Decision Recognition 47
5.1.3. Diagnosis 49
5.2. Development 51
5.2.1. Search 51
5.2.2. Design 52
5.3. AI Challenges 53
5.3.1. Data Privacy 53
5.3.2. Technological Capabilities 54
**6. Conclusion**
6.1. Limitations and Future Research 57
**References 59**
**Appendices 63**
Appendix A - Interview guide 63
Appendix B - Visual Map of Literature collection 64
Appendix C - Research Proposal 65

# List of Table

Table 1 - List of Interviewees 30

## List of Figure

Figure 1 - Conceptual Model 24

# 1. Introduction

## 1.1. Background

"[...] more recently, a lot of companies have been asking me, we see ChatGPT, we see these generative AI technologies, what does it mean for us?" (Epsilon)

Making sound decisions is what arguably any firm aims for in any situation, ranging from everyday minor decisions to the impactful strategic decisions. When stakes are high, making a sound decision for a firm can be of vital importance. This need for optimal decision making has motivated researchers to theorize and present step-by-step models to guide decision makers in their process. Typically, a rational decision making model consists of a sequential process where each stage of the process is followed linearly (Heracleous, 1994). Some researchers question the assumptions behind such models (Heracleous, 1994) and other researchers emphasize the impact human individuals have on any decision making as we are bound by our limit of rationality (Hambrick & Mason, 1984). Thus, finding any optimal way of making decisions will be limited by the individual decision makers themselves in some shape or form.

In our world which is considered to be volatile, uncertain, complex and ambiguous (VUCA), responding to this environment is becoming more challenging for firms than ever according to Taskan, Junça-Silva and Caetano (2022). They explain that recent events such as the Covid-19 pandemic and financial crises are examples of situations where things have changed quickly and profoundly in society. For these reasons, the operation of navigating through a VUCA-world requires firms to handle large amounts of data (Taskan et al. 2022). Using them to make strategic decisions can prove to be important competitive advantages for firms (Taskan et al. 2022).

A topic that has faced tremendous increase in attention recently is Artificial Intelligence (AI) technology, and in particular the Large Language Model (LLM) called ChatGPT. However, it is important to note that AI is not a new technology, in fact, Alan Turing published a paper as early as 1950 which discussed computer intelligence (The Alan Turing Institute, 2019). Moreover, AI is an umbrella which covers many types of technologies, but this study will address AI as an LLM due to its increase in attention and its promising potential. Much of the attention raised can be attributed to a particular LLM, namely ChatGPT. This generative pre-trained transformer version 3, called ChatGPT-3, was released to the public in November 2022 by the company OpenAI (OpenAI, 2022) and showed more serious potential than its predecessor ChatGPT-2, thus attracting much attention. Looking at Google Trends for example, the search term “chatgpt” was valued 0 or less than 1 before it was made public in November 2022, and was in April 2023 ranging in value between 89 to 100 (Google Trends, 2023). It is important to note that there are several companies and institutions that have created their own large language models, even though OpenAI’s models can be argued to be the most popular ones with the smartest capabilities in many aspects (OpenAI, 2023) However that might change due to competition from other companies such as Google, with their PaLM2 which has proven better multilingual capabilities (Google, 2023) or open source alternatives.

As mentioned, the pace of development is remarkably fast. At the same time as ChatGPT-3 was released to the public, OpenAI was done with its successor ChatGPT-4 and started to test it for six months, to realize it by March 2023. When comparing the two versions, which only have six months between their respective release date, there are significant differences and improvements to be found. These improvements have been shocking for many, showing both major improvements, unexpected surprises but also some stagnated areas of development. In general, AI can be classified into three categories: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI) (Joshi, 2019). Currently, the technology is viewed to be ANI and the following AGI is considered to lie far ahead in the future (Berruti, Nel & Whiteman, 2020). What makes AGI different from ANI, if it was to be realized, would be its capability to have a human intelligence, such as intuition or similar. Microsoft tested the unrestricted GPT-4 version during the six months prior to its release and stated that the GPT-4 version reasonably showed signs of an early state AGI system (Bubeck, Chandrasekaran, Eldan, Gehrke, Horvitz, Kamar, Lee, Lee, Li, Lundberg, Nori, Palangi, Ribeiro & Zhang, 2023).

As defined by Alvesson (2004), knowledge-intensive firms’ (KIFs) intellectual resources are central to their business models, and are typically further divided into R&D or professional service firms (PSFs). Consultancies, together with law and accountancy firms, are some examples of PSFs and are different from R&D firms as the typical PSF interacts directly with the market, i.e. its clients, and deals with intangible rather than tangible products (Alvesson, 2004). Furthermore, PSFs are typically homogenous as they require formal association and certification to name a few dimensions, but there are exceptions (Alvesson, 2004). In particular, management consultancies are such an exception as they do not have a regulated entry, standardized certification or homogeneity while being a PSF (Alvesson, 2004). Derived from research by Alvesson (1995 cited in Alvesson, 2004), Deetz (1997 cited in Alvesson, 2004) and Løwendahl (1997 cited in Alvesson, 2004), he summarizes characteristics of KIFs into seven points. Firstly, knowledge-based work is done by individuals who are highly qualified and who are using their intellectual skills in their work. Secondly, there is a high degree of autonomy and organizational hierarchy is de-emphasized. Thirdly, such firms tend to have adaptable organizational forms. The fourth point includes a need for communication concerning problem-solving and coordination, and the fifth point concerns specific client services. Sixthly, there is an asymmetry of information and power and lastly, the qualitative assessments can be subjective and uncertain. In all, the management consultancy firm is a KIF, and the characteristics defined by (Alvesson, 2004) are applicable to describe the case firms of this study.

Despite its quite short time of being accessible to the public, ChatGPT has already created controversy and debates in many parts of society, and the education system is one example. A current question for educators is whether schools and universities will be forced to adapt to the presence of ChatGPT, or find ways to detect and prohibit their students from using it in their work (Johansson, 2023). Representatives from the department of informatics at Lund University have expressed a need to incorporate such tools in education while also stating that the tool cannot solve everything (Johansson, 2023). Either way universities decide to tackle this issue, the technology is here and it is widely accessible. This example highlights a need for other parts of society to address this transformative technology in one way or another, and firms are no exception. In all, the technology of AI is becoming more accessible and could be a potential means for firms in their strategic decision making processes in a more VUCA-world.

## 1.2. Problematization

Firms play a crucial role in society and their existence is dependent on the value they can provide to society. As Chandler (2021) describes, the firm is the most efficient organization that can combine valuable and scarce resources on a scale for creating value. He further explains how society and firms are not separate but interwoven, meaning that their interests are aligned and cannot exist without each other. By having this role, firms have great power over how society progresses and overall quality of life (Chandler, 2021). All the while, society puts pressure on firms to stay competitive when consumers exercise their power as they decide which firms they are willing to pay (Chandler, 2021).

This pressure on firms requires them to perform and stay competitive. In order to perform, a firm must make strategic decisions, some more important than others and some more frequent than others, all depending on various factors that are specific for the particular firm. In turn, what can affect firms' strategies are its resources and those can be physical, human or organizational capital (Barney, 1991). Studying how firms respond to groundbreaking technology, such as the emerging powerful AI in form of LLMs, is therefore relevant and of interest as it will likely have a prominent effect on firms' resources, and thus in turn society and many individuals' lives one way or the other. More specifically, studying consulting firms poses a possibility to get an insight into what issues and questions many other individual firms might have as they turn to consultancy firms for strategic advice and solutions. Other parts of society, such as the education system, have already faced the question of *how* to adapt to the new reality with new AI, and not *if* it should do so. It could be expected that firms will have to do the same thing sooner rather than later, if they are not already doing it.

One way to stay competitive is to interpret and utilize big data (Taskan et al. 2022), and AI is able to do this task much more efficiently than compared to humans (Kumar, 2023). Large Language Models, such as ChatGPT-4, have started to show human level performance on various professional and academic benchmarks, however still less capable than humans in many real world scenarios (OpenAI, 2023). Inversely, this implies that there are real world scenarios where this newfound potential power could be useful. As the new LLMs are a novel technology, which has become accessible to the public recently, it has not yet been widely researched empirically to the best of our knowledge what LLMs relation currently have or could have to the strategy of firms.

### 1.3. Purpose and Research Question

The purpose of this case study is to contribute with an understanding of what role AI has, and might have, in strategic decision making processes. In order to contribute to this area of decision making literature, the following research questions will be answered:

*How can Large Language Models be utilized in strategic decision making processes?*
*What challenges of using Large Language Models are perceived by professionals in knowledge-intensive firms?*

# 2. Literature Review

This chapter provides an overview of literature within decision making and AI which will vary in generality and specificity depending on its relevance for answering the research questions of *How can Large Language Models be utilized in strategic decision making processes?* and *What challenges of using Large Language Models are perceived by professionals in knowledge-intensive firms?* Since AI may be viewed as a resource for firms in their strategic decision making, the literature review will start with the Resource Based View (RBV) which assumes that firm resources are heterogeneous instead of homogeneous (Barney, 1991). Its connection to strategic decision making will be highlighted to ensure that the fundamental theories of which the thesis is built upon is evident to the reader, as RBV relates to decision making. Subsequently, how strategic decision making is a dynamic capability will be elaborated on. Afterwards, strategic decision making is described in a general way by breaking it down through the lens of behavioral theory and comparing human versus computer in decision making. To adequately address the research questions, it is imperative to consider the most recent and relevant literature on AI, which is a rapidly evolving field (Shao, Zhao, Yuan, Ding & Wang, 2022). This thesis will incorporate relevant literature that encompasses fundamental concepts of AI, more specifically LLM, and provides insights into its potential future developments. Finally, a conceptual model will be presented to illustrate the key concepts of relationships, including AI's part.

## 2.1. Resource Based Theory

The Resource-Based View (RBV), discussed in Barney's (1991) well-known article 'Firm Resources and Sustained Competitive Advantage', challenges assumptions that other researchers within strategic research, such as Porter, make about firms' resources and their competitive advantage. Such models, like Porter's (1980 cited in Barney, 1991) 'Five Forces', assume that firms operating within an industry have resources that are identical in the sense of how they are used strategically and that if any differences between these firms' resources would appear, they would soon fade away due to market forces being in place (Barney, 1991). Barney (1991) further explains that such assumptions can serve well when having an external perspective on firms' performances, but that the RBV is more appropriate when making an internal analysis. As this study examines strategic decision making processes, conducting an internal analysis is deemed appropriate. Instead of viewing firms' resources to be homogeneous and mobile, Barney (1991) accepts two other assumptions which view firms' resources to be heterogeneous and immobile. This means that a firm attains a competitive advantage whenever it can carry out a strategy which creates value, that a competitor, current or potential, cannot carry out at the same time (Barney, 1991). Moreover, Barney (1991) specifies that when this value creating strategy's benefits cannot be duplicated by a competitor, it is a sustained competitive advantage. However, he notes that a sustained competitive advantage is not guaranteed to last for eternity, but can be disrupted by 'Schumpeterian Shocks' which can change a firm's resource from being a strength into a weakness due to adapting to a new business environment as these shocks revolutionize industries structurally (Barney, 1991). In all, Barney (1991) identifies four attributes of a resource in order to be a sustained competitive advantage: valuable, rare, imperfectly imitable and non-substitutional, which make up the VRIN-criteria.

Human capital is one of three categories of resources, and Barney (1991) defines it as "the training, experience, judgment, intelligence, relationships, and insight of individual managers and workers in a firm." (p. 101). Much of the core characteristics of a knowledge-intensive firm could be attributed to this category when comparing Barney's (1991) human capital to Alvesson's (2004) characteristics of a KIF. The second category, physical capital, is defined by Barney (1991) to encompass a firm's physical technology, equipment and raw materials. He also mentions computers and machines as part of a firm's physical capital and since they usually can be purchased across markets, he claims that they can never become a sustained competitive advantage. However, information processing systems that are embedded in the strategic decision making process of a firm could potentially be a sustained competitive advantage as they are unlikely to be imitable due to being socially complex (Barney, 1991). Lastly, the third category is named organizational capital, and includes elements such as formal and informal planning as well as informal relations within the firm and between the firm and its environment (Barney, 1991).

As a final remark in his article, Barney (1991) points out that managers, or a managerial team, can make up a sustained competitive advantage without controlling resources that are considered being VRIN. He emphasizes the managerial analysis and the ability to understand and describe the economic potential of a firm's endowments. This implies that the knowledge-intensive firms, including consulting firms helping clients with strategic issues which this study is founded upon, will hold a similar potential to reach a sustained competitive advantage on the same grounds as Barney (1991) argues for the managerial analysis and abilities. Applying the

RBV perspective which emphasizes resource heterogeneity between firms, both in regard to having them and moving them between firms (Barney 1991), is deemed appropriate since this study is focused on the knowledge-intensive firm in the form of consultancies as they typically hold heterogeneous resources (Alvesson, 2004).

Eisenhardt and Martin (2000) build upon previous RBV and contribute to this field of research by making additional points on the interpretation of RBV including dynamic capabilities and market conditions. Eisenhardt and Martin (2000) examine dynamic capabilities which they mean broadly speaking is RBV. They define dynamic capabilities as a set of specific and identifiable organizational processes and explicitly lift strategic decision making as one of them, which can be executed in more or and less effective ways. Such a process creates value for the firm by utilizing the resources in value-creating strategies. Strategic decision making is a dynamic capability in the sense that managers use their expertise from various areas such as business, functional and personal to make decisions that drive the firm (Eisenhardt & Martin, 2000). They also lift an example by Fredrickson (1984 cited in Eisenhardt & Martin, 2000) from a slowly evolving industry, the paint industry, which demonstrates a linear decision making as more effective (Eisenhardt & Martin, 2000). The process resembles the rational decision-making model mentioned earlier (Heracleous, 1994), where a sequence of problem solving steps like gathering data, development of alternatives, analysis of the alternatives and decision (Eisenhardt & Martin, 2000).

In contrast to the market which the paint industry would be found in, dynamic markets or high velocity markets are defined as markets with blurred boundaries (Eisenhardt & Martin, 2000). The authors state that there is ambiguity of what business model is successful, as well as ambiguity and shifts exists among the market players such as suppliers, buyers, competitors and complementors. Additionally, change is non-linear, less predictable and cannot be modeled in probabilities (Eisenhardt & Martin, 2000). However, they mean that the effective dynamic capabilities are reversed in the way that they are simpler in the dynamic markets compared to those in moderately dynamic markets, which are considered complicated. There are routines in the dynamic markets that consist of a few rules that specify boundary conditions (Eisenhardt & Martin, 2000). The dynamic capabilities rely on new knowledge, which is rapidly gained, for specific situations (Eisenhardt & Martin, 2000). To allow for emergent adoption, routines are purposefully simple without being completely unstructured and in order to account for changing information, the routines are iterative and cognitive mindful, not linear and mindless

(Eisenhardt & Martin, 2000), which in some ways could be compared to what they earlier stated about the slowly evolving paint industry. The effective routines are thereby adaptive to changing circumstances and come with the price of unstable processes and unpredictable outcomes (Eisenhardt & Martin, 2000). Eisenhardt and Martin (2000) suggest that routines as dynamic capabilities are found within the aggregated and existing knowledge in the moderately dynamic markets, and these routines could include analysis from “existing knowledge and rules of thumb” (2000, p.1116) which is then implemented.

Additionally, Eisenhardt and Martin (2000) view dynamic capabilities to not be “tautological, vague and endlessly recursive” (p. 1116), but rather well-known and well-studied processes which have been studied separately from RBV, such as strategic decision making. They further describe that the value of dynamic capabilities is derived from their capability to change “the resource base: create, integrate, recombine, and release resources” (2000, p. 1116). Changing the resource base can be argued to be in line with Barney’s (1991) statement about information processing systems that are embedded in the strategic decision making of a firm, potentially reaching a sustained competitive advantage and the managerial ability.

Moreover, Eisenhardt and Martin (2000) state that dynamic capabilities should be seen as strategic and organizational processes where strategic decision making is mentioned as an example. These processes are value-creating for firms in the dynamic markets by leveraging resources to create strategies that bring value (Eisenhardt & Martin, 2000). They also explain that RBV overvalues the strategic logic of leverage when change should be emphasized, aiming for a series of temporary competitive advantages rather than directly gaining a competitive advantage in the long-term, which is comparable to Barney’s (1991) sustained competitive advantage. Eisenhardt and Martin (2000) argue that the capabilities are more common in some ways than what RBV suggests, in particular that they can be substituted, seen as homogenous or achieve the same result but in different ways.

## 2.2. Strategic Decision Making Theory

Mintzberg, Raisinghani and Théorêt (1976) define strategic decisions as decisions which are important in regards to what actions are taken, what resources are committed or what example it makes. Eisenhardt and Zbaracki (1992) concur with the definition of Mintzberg et al. (1976), phrasing it as “infrequent decisions made by the top leaders of an organization that critically affect organizational health and survival” (1992, p. 17). Holding to the same view as Eisenhardt and Zbaracki (1992) and Mintzberg et al. (1976), we argue for the strategic decision making processes to involve others than only top management, such as KIFs and more specifically consultants who work directly with their market (Alvesson, 2004) by assisting top management in their strategic problems and provide recommendations for solutions.

In a literature review on strategic decision making, Eisenhardt and Zbaracki (1992) go through the most central elements of this research area, such as rationality and bounded rationality, but also garbage can as well as politics and power. The garbage can together with politics and power will not be subject for further elaboration in this study as the research questions are not focused on interpersonal relationships but rather the relationship between humans and technology. Eisenhardt and Zbaracki (1992) conclude in their review that bounded rationality is viewed as an uncontroversial assumption within decision making research due to three main reasons. Firstly, they find support in the empirical research that the rational model is colored by cognitive limits as decision makers tend to satisfice rather than optimize. Secondly, they see that there are certain fundamental phases that a majority of decisions follow, more precisely identifying the problem, development and selection. However, they emphasize that decisions go through these phases repeatedly in different manners. Thirdly and lastly, they found in the empirical research that the route for the decision is usually influenced by the complexity of the issue together with any potential disagreement between the decision makers. One of the articles reviewed by Eisenhardt and Zbaracki (1992) is by Mintzberg et al. (1976), who made a field study on 25 strategic decision processes and this study is used by Eisenhardt and Zbaracki (1992) to support the claim that rational models are wrongfully assuming that phases of decision making are sequential.

### 2.2.1. The Strategic Decision Making Process

In their study, Mintzberg et al. (1976) argue that there is a basic structure beneath the "unstructured" processes. They define the unstructured decision process to not have any predetermined responses as it has not been encountered before. Furthermore, they explain that the basic structure consists of a number of elements. According to Mintzberg et al. (1976), the case studies imply that the strategic decision making process is defined by complexity, novelty, and open-endedness as organizations typically start this process with a limited comprehension of the decision's context, as well as the means to arrive at a solution, its potential appearance and what it would be evaluated by. Moreover, the decision process is described to be iterative and discontinuous and to contain dynamic factors together with challenging stages over longer periods of time (Mintzberg et al. 1976). The dynamic factors have been left out consciously as they are found to be irrelevant for this study. An important distinction Mintzberg et al. (1976) make is that this strategic decision process occurs under ambiguous circumstances, where there is barely any predefined or readily apparent information available compared to decision making under uncertainty, where alternatives are given but with unknown consequences.

Additionally, Mintzberg et al. (1976) categorize decisions by stimuli, solution and process. They view stimuli to put decisions on a continuum with two ends of extremes, where one end consists of opportunity decisions that are initiated from a secure situation, more specifically when there is an opportunity to improve something that already exists. They view the other extreme as crisis decisions that emerge from situations which require urgent action and in between these two extremes are the so-called problem decisions. What can move a decision process along this continuum is either managerial action or passivity, which they exemplify by describing that "[...] an ignored opportunity can later emerge as a problem or even a crisis" (Mintzberg et al. 1976, p. 251).

Mintzberg et al. (1976) further classify decisions by solution in the four ways of fully-developed, ready-made, custom-made, and modified, where the latter is a combination of the ready-made and custom-made solution. The fully-developed refers to a solution being presented when the decision process begins, the ready-made is a fully-developed solution which appears during the process and the custom-made is refined for the particular decision (Mintzberg et al.1976). The process, which is between a stimulus and reaching a solution, is by far the most complex one out of the three categories (Mintzberg et al. 1976).

The process consists of the three phases of identification, development and selection, and the phases are described through seven central routines (Mintzberg et al. 1976). The identification phase has two routines, decision recognition and diagnosis where the first mentioned refers to recognizing opportunities, problems and crises which marks the stimuli point of decisional activity (Mintzberg et al. 1976). Decision makers receive a large quantity of ambiguous data in which they need to identify both problems and opportunities (Sayles, 1964; Mintzberg, 1973 cited in Mintzberg et al. 1976). Opportunity decisions often emerge from a single stimulus in the form of an idea while crisis decisions often come from a single stimulus that appears suddenly and unequivocally (Mintzberg et al. 1976). In a crisis decision, immediate action is necessary and the situation is compared to a fire or going bankrupt by Mintzberg et al. (1976).

In contrast, problem decisions typically emerge from multiple stimuli and the decision makers presumably want to read the situation before action is taken (Mintzberg et al. 1976). Problem and crisis differ as crisis decisions have a more urgent and obvious nature since the crisis situation cannot be read to the same extent as the problem (Mintzberg et al. 1976).

Additionally, Mintzberg et al. (1976) lift the phenomenon of matching, where action from managers is more likely when a problem can be matched with an opportunity. In lack of either, if there is an opportunity with no fitting problem, or inversely a problem with no apparent solution, the decision maker may be reluctant to act (Mintzberg et al. 1976). They describe the determining factor of action as the relationship between the cumulative amplitude of the stimuli and the action threshold. However, the action threshold is not a definite level but rather shifts depending on the individual manager (Radomsky, 1967 cited in Mintzberg et al. 1976). The amplitude is further described as factors, such as the influence the source has, the interest for the decision maker, perceived payoff by taking action, associated uncertainty and success probability (Mintzberg et al. 1976). Each of these factors are cumulated where their frequency, pattern, clarity and consistency determine if the stimulus together gets an amplified effect or decays over time (Mintzberg et al. 1976).

The diagnosis routine aims to comprehend the stimuli by determining cause and effect relationships for the decision situation (Mintzberg et al. 1976). The diagnosis routine occurs after the stimuli has passed the threshold and this routine is not necessarily formal or explicit, where an example of formal diagnosis would be to request consultants to analyze a new issue (Mintzberg et al. 1976). However, Mintzberg et al. (1976) note that some form of diagnosis is needed for strategic decision making since the decision maker faces a novel situation with an array of data which is partially ordered. Their research shows that the first step in diagnosis is to use existing information channels as well as new ones, and it also suggests that the formal diagnosis is more prominent in the mild problem range in the opportunity, problem and crisis continuum. Furthermore, they suggest that opportunities may not require much investigation since there might not be much to correct, rather only to improve. In contrast, intense problems or crises might entail time and cognitive pressures which discourage the formal diagnosis, such as reaching out for consulting services (Mintzberg et al. 1976).

Moving over to the development phase, Mintzberg et al. (1976) label two routines as search and design. In terms of decision making resource consumption, it dominates both the previously reviewed identification phase as well as the selection phase (Mintzberg et al. 1976). The search routine entails finding the solutions that are ready-made, and the design routine is used to develop custom-made or modify ready-made solutions. They highlight the importance of this distinction since the development phase will only be of a lighter nature when the solution already is fully developed, in other scenarios the development phase will be extensive. It is also worth noting that the research of Mintzberg et al. (1976) shows that the two routines may happen simultaneously. Regarding search, there are four types of behavior and generally search begins from a convenient starting point until a more active search is needed (Mintzberg et al. 1976). The first type is memory search where the firm's existing memory is scanned from both individuals and documents while the second type refers to a passive search where the firm waits for alternatives to appear (Mintzberg et al. 1976). The third type of behavior is trap search where the firm lets others know that they are searching for solutions, thus activating search generators (Soelberg, 1967 cited in Mintzberg et al. 1976). Lastly, the fourth type refers to active search where a firm directly seeks alternatives, ranging from a wide to narrow scanning (Mintzberg et al. 1976).

At the stage where search is not possible or a state of maturity is reached, the firm may turn to the form of a design for a custom-made or modified solution (Mintzberg et al. 1976). Furthermore, they argue that the custom-made solution is a complex and iterative procedure. Moreover, since the designing of custom-made solutions can be expensive and time consuming, firms tend to be unwilling to spend resources on several alternatives (Mintzberg et al. 1976). However, generating additional alternatives during the search routine comes at a rather low cost and can lead to several solutions, which is possible when the need for design is low in modified or ready-made solutions (Mintzberg et al. 1976).

The selection phase is the final phase of the decision process of Mintzberg et al. (1976) and it is typically a multistage, iterative process where the alternatives are examined in depth. The three routines in the selection phase, as described by Mintzberg et al. (1976), are screening, evaluation-choice, and authorization. The selection phase usually lies at the decision maker's table and is thereby deemed not relevant to go into depth since the final decision lies outside the scope of this thesis. However, it might be worth noting that screening is a rather superficial routine and to a great extent implicitly happens in the earlier mentioned search phase according to Mintzberg et al. (1976). More specifically, they view screening to be mostly a way of eliminating the non-feasible ideas. Thus, the screening routine is argued to be in the scope of this thesis although the rest of the selection phase is not. The evaluation-choice contains three modes where the decision is done through judgment by an individual, bargaining in a group with conflicting judgments by individuals or analysis which involves factual evaluation combined with managerial choice, which in turn involves judgments both groups and individuals Mintzberg et al. (1976). Overall, the section involves a large number of soft and non-quantitative factors which makes the phase hard to concretize (Mintzberg et al. 1976). Authorization is typically a binary situation with either acceptance or rejection (Mintzberg et al. 1976).

## 2.3. Interplay of Humans and AI in Decision Making

Ivanov (2023) presents three decision making approaches and ten characteristics. The author further analyzes how the human, as the principal, interacts with the artificial autonomous agents, shortened as AA. His three approaches consider the human in the loop, on the loop and out of the loop. Factors that Ivanov (2023) consider are the dependency on the AA, the human involvement in a principal-agent relationship and the responsibility for the decision outcome (Ivanov, 2023). In the loop refers to the AA recommending a decision where the human makes the confirmation, giving the AA a supporting role (Ivanov, 2023). He means that it creates a non-existent relationship between the AA and the human as there is low dependency on the AA, which in turn requires full involvement from the human in the decision making process as well as full responsibility of the human decision maker (Ivanov, 2023). When it comes to the human being on the loop, and considering the same factors, the author explains that the principal can override an unwanted decision. Therefore, Ivanov (2023) suggests that the dependency on the AA is medium and the human involvement is minimal to medium with an unchanged responsibility as it still belongs to the human. Ivanov’s (2023) third approach, when the human is out of the loop, represents a fully automated relationship where the principal has no control over the decision made by the AA, which results in a high level of dependency on the AA with no human involvement and the responsibility shared amongst human employees involved around the AA is considered to be unclear.

These three approaches described by Ivanov (2023) clusters ten levels of automation classified in a framework by Kaber and Endsley (1997) and Endsley and Kaber (1999), who in turn base their framework on a technical report in human performance and automated systems literature by Sheridan and Verplank from 1978. Sheridan and Verplank (1978) constructed ten levels of automation in man-computer decision making, in the context of undersea teleoperators. These are not to be confused with the ten characteristics. They stress in their report that the ten steps are merely examples and that there might be other variations to consider (Sheridan & Verplank, 1978). The findings of Sheridan and Verplank (1978) are thus incorporated into the framework by Endsley and Kaber (1999) , the technical report describes a scale of one to ten of human integration in the computer decision. Ivanov (2023) clustered the ten levels, which when viewed separately from the approaches shows vague distinctions between them. It is thereby inappropriate to discuss the particular levels within an approach for the purpose of this thesis and instead focus on three approaches whilst diving deeper into the characteristics that are relevant for this thesis.

### 2.3.1. Three Approaches to Decision Making

Ivanov (2023) excludes Level 1 since it involves completely manual control without any automation. The human in the loop approach encompasses the majority of the levels, 2-7, and covers a range of automation involving action support, batch processing, shared control, decision support, blended decision-making and rigid systems (Ivanov, 2023). In each of these levels, it is the human who ultimately selects the decision option, regardless of who generated it or who will implement the decision, whether it is the human or the automated system (Ivanov, 2023). The human "on the loop" approach pertains to levels 8 and 9, which involve automated decision-making and supervisory control, respectively (Ivanov, 2023). In these levels, he describes that the automated system selects and implements the decision option, while the human monitors the system. He also mentions that the human "out of the loop" approach overlaps with Level 10, full automation. Ivanov (2023) continues by emphasizing that the three approaches might be simultaneously present in a process.

### 2.3.2. Ten Characteristics Determining the Approach

Ivanov (2023) also provides ten characteristics that are more or less prominent depending on the taken approach. His characteristics regard, speed, frequency, level of algorithmization, negative consequences by human, negative consequences by AI, decision complexity, uncertainty for context, uncertainty for outcome, transparency required and ethical issues. Among the characteristics, only three can be considered as directly relevant with a high representation, since they encompass level 2-7 with the human in the loop and 8-9 with the human on the loop (Ivanov, 2023). The two extremes of no automation and full automation are left out consciously as it is not found reasonable to include them in the context of this study. For both of the approaches, level 2-9 decision complexity and outcome uncertainty is highly represented with ethical issues prevalent in level 2-7, or human in the loop (Ivanov, 2023).

According to Ivanov (2023), decision complexity involves the number of interconnected and affected decisions in the decision making process, number of involved stakeholders, and amount of data needed. He exemplifies that decisions with high complexity for instance are various strategic decisions for a firm as a new market entry which in turn entails a number of other decisions about marketing and pricing to mention a few. Furthermore, he argues that complex decisions will need to involve AI to deal with large amounts of data, still giving autonomy to the human to take the final decision. He applies this to the human in the loop approach and when given the autonomy to change the decision, he also applies it to the human on the loop approach. Outcome uncertainty involves the predictability of the outcome for a specific decision, where again a strategic decision as a new market entry is given as an example involving high outcome uncertainty (Ivanov, 2023). Ivanov (2023) further argues that both decision complexity and outcome uncertainty must thereby be addressed through a combination of AI and human intelligence.

Ivanov (2023) states that it is unquestionable that both humans and AI are imperfect, having biases that influence their decision making process. He further means that it is hard to integrate ethics in AI, resulting in that the AI itself is unethical or that practitioners use it unethically. Therefore, a human in the loop approach might be preferred in severe ethical issues (Ivanov, 2023). On the contrary, AI can also mitigate human biases (Ivanov, 2023). The conclusion from the ethical standpoint is that a human in the loop approach is the most suitable approach, however dependent on the technological ability (Ivanov, 2023), the AI's involvement might be more or less prominent.

However, there are some compelling points from the other seven characteristics. For instance it might be that context is important, since the context may vary depending on the industry and how well defined the market context is (Ivanov, 2023). It can be argued that the context is not important since the strategic decision making will be complex in either way, irrelevant of the context, as presented previously about decision complexity (Ivanov, 2023). When it comes to transparency, complex decisions, and in particular those that involve ethical issues or might entail potentially significant consequences, require high transparency (Ivanov, 2023). While simpler or repetitive tasks might hold little value of putting emphasis on transparency, as would be the case with the human of the loop approach (Ivanov, 2023). He further states that mistakes can differ in their implications, some entailing more negative consequences than others. When stakes are high, the best option is the human in the loop approach, with an inverse argument for insignificant consequences (Ivanov, 2023). He states that the argument is only valid as long as human decision making is superior to AA or if their decision making is indifferent in terms of quality, since humans are ultimately responsible for the decisions and can bear legal responsibility. The arguments for frequency are rather straightforward, where the general rule is the higher the frequency of the same decision, the higher the involvement with AAs (Ivanov, 2023). The argument rests on non-existing economic feasibility of using the AA in infrequent decision making, meaning that it is purely an economic question (Ivanov, 2023). The level of algorithmization is tightly related to frequency and should as a general rule increase with the frequency level (Ivanov, 2023). Ivanov's (2023) general rule for speed is the faster a decision needs to be taken, the less the involvement of the human is needed, inversely resulting in decisions that allow sufficient time for humans may not require automated assistance. Such rapid decisions which he considers to benefit from using AA are instances where decisions are made in less than a second. Therefore, using speed as a determinant in strategic decision processes is deemed to be less relevant for this study.

On a cautionary note, as Ivanov's article was published in 2023 but written in 2021, there is a risk that it is written with a perspective lacking the most recent and emerging LLMs and their capabilities. Much of what has been stated by Ivanov (2023) is still useful, however there are some points to reconsider as they might become less relevant in comparison to LLM development. For instance, the ethical aspects are included in the capabilities of the new LLMs, however still not perfect but mitigated to a certain extent. Ivanov (2023) pointed out that the current technological level of AI, when the report was written, had not reached a stage when AAs could evaluate themselves on ethical aspects in a holistic manner. Ivanov's (2023) research asks questions that this thesis will touch upon such as the ethical issues previously mentioned together with the question he poses for future research regarding the applicability of the approaches depending on the organizations they are applied in.

## 2.4. AI History and Overview

Shao, Zhao, Yuan, Ding and Wang (2022) provide a comprehensive chronological overview from the beginning of AI to the current state, by analyzing research and literature in the field.

They present a range of fields of AI, all branching out from the beginning of AI in 1950 with the Turing's test, and the company OpenAI can be found at the far end of one branch. Some categories of AI that relate to LLMs are transformers, deep learning, machine learning, neural networks and pattern recognition such as natural language processing and image recognition (Shao et al. 2022). It lies outside of the scope of this study to describe the technical aspects of these categories in detail. However, the functions of some of them and its relation to the LLMs will be briefly described to provide the reader some background to better understand the empirical data.

### 2.4.1. The Structure Under ChatGPT-3

When it was discovered that the human brain consists of a neural network, scientists started to investigate whether there could be an artificial brain (Shao et al. 2022). OpenAI's ChatGPT-3 is based on neural networks, which Wolfram (2023) explains can be thought of as the idealization of how a brain is assumed to work. Moreover, he expresses that the neural network under the ChatGPT-models can generate human language on a surprisingly good level. Traditionally, neural networks are connected through layers where all neurons on one layer are connected to both the previous and sequential layer, where an analogy could be of layers of images on top of each other where the pixels represent the neurons (Wolfram, 2023). All of the connected neurons have certain weights which are determined through the training data, where some of the connections amount to zero weight which simplifies the computation (Wolfram, 2023). The ChatGPT-3 neural network consists of 175 billion connections of artificial neurons, while the human brain could have 100 trillion connections of neurons as it has 100 billion neurons (Wolfram, 2023). Ultimately, the process under ChatGPT is rather simple because the artificial neural network is doing a simple computation on a collection of numerical inputs and combines them with certain weights to generate a new token, which requires 175 billion calculations (Wolfram, 2023). He further lifts that there is no theoretical way of explaining how many examples of text that is needed to train a human-like large language model, and states that ChatGPT was successfully trained on a few hundred billions of words in text.

As Wolfram (2023) describes it, after ChatGPT's training on the hundreds of billions of words in text, it can generate its own text from new inputs given as prompts. He furthermore explains that ChatGPT doesn't always give a correct output since it is designed to continue the text in a reasonable way based on its training data. He further stresses that the guidelines in the prompts have to be simple without going into deep computation with irreducible steps and if this is to be possible the neural nets, similar to humans, need to use external actual computational tools. It should be noted that the Wolfram (2023) article is referring to the GPT-3 version, and the GPT-4 version got access to such tools in the way Wolfram (2023) is describing, and Wolfram himself have been using GPT-4 with an API to his own computational knowledge engine, called Wolfram Alpha.

Wolfram (2023) makes a point that human language and the thinking process needed behind it seemingly have been seen as a peak of complexity, but continues by suggesting that language may be simpler than first thought. He means that in some way, even though the underlying computations are simple and straightforward, ChatGPT has successfully captured the essence in language and the thinking behind it by implicitly discovering the regularities behind how the language is put together. This is something that previously has been assumed to require a human brain, which leads Wolfram (2023) to believe that there is more simplicity and structure than previously thought in meaningful human language, as well as simple rules that describe how the language is constructed. However, the question of how efficiently a neural network can implement a model based on its ability to understand the human language remains to be answered according to Wolfram (2023), and this question lies close to the purpose of this thesis.

ChatGPT might still appear compelling when giving output that seems correct as it is based on what it sounded like in its training data. Wolfram (2023) lifts that without a connection with actual computational tools or external tools in general, ChatGPT is generating reasonable text without knowing if it is true or not, however with external tools it could fact check itself on its output. It can be noted that Ivanov (2023) makes a similar point regarding how technological ability can have an influence on what approach is most suitable ethically.

Wolfram (2023) means that the two greatest insights to bring forward from ChatGPT are the fact that a large number of simple computational elements have succeeded with remarkable and unexpected things and that we better understand the processes behind language and the thinking that generates it. Additionally, he highlights some important points in his thorough review on ChatGPT-3. A similar review of the architecture of GPT-4 is currently not possible to find or create, because of OpenAI's decision to not disclose such information (OpenAI, 2023). The next two sections will present what has been publicly available about GPT-4, mostly regarding its performance and implications.

## 2.5. Novel Large Language Models

### 2.5.1. GPT-4 Technical Report

In a technical report, OpenAI (2023) themselves describe the GPT-4 version to be different and improved, but still having some limitations as previously found in GPT-3 and 3.5. For instance, the context length amounts to twice as much, roughly 8 000 tokens. A limited version gives access to 32.000 tokens which is about 50 pages of text in both input and output form. Due to a competitive market and the safety concerning LLMs, OpenAI (2023) does not release detailed information about the architecture including but not limited to hardware, training compute, dataset construction, parameter count and dataset construction. They showcase impressive statistics on its performance in academic and professional exams, such as GPT-4 outperforming GPT-3.5 in most exams, such as the Uniformed Bar Exam where it scored in the top ten percent while the GPT-3.5 version scored in the bottom ten percent. Moreover, they show that the GPT-4 version also reached the 88th percentage compared to GPT-3.5 which only reached the 40th percentage in the LSATs. They provide more examples of major improvements, but also stagnated areas which show zero to minimal improvements, such as the Writing Graduate Record Examination. The GPT-4 version is also multimodal, meaning that instead of transferring text input to text output it can also convert image and audio to text (OpenAI, 2023). It outperforms other models in terms of interpreting infographics, as the technical report shows GPT-4's ability to explain photos of a text with graphs and other examples where it explains the underlying meaning, and even humor, of photos (OpenAI, 2023).

When OpenAI (2023) tested GPT-4's ability to follow user intent, results showed that in 70 percent of the prompts, humans preferred responses from GPT-4 over responses from GPT-3.5 humans. OpenAI (2023) also gave the GPT-4 an undisclosed task called Hindsight Neglect in which GPT-4 outperformed previous models significantly. Additionally, they present that earlier models were getting worse and worse at this task as the models improved in scale, but GPT-4 has reversed the trend and shows an accuracy close to 100. OpenAI (2023) admit that they did not predict this capability to improve and that there still are capabilities which are difficult to forecast. They further highlight that forecasting capabilities is an element they perceive to be important from a safety perspective for them as a company, but also for others in the same field.

OpenAI (2023) warns that the GPT-4 is not perfectly reliable as it is still hallucinating, but the newer model is hallucinating less than GPT-3 since it is better at factual accuracy. The company points out that it still has a limited context window and a capped knowledge after 2021 September when its pre-training data was cut off, and cannot be claimed to be learning from experience. Thus, they state that care should be taken when the generated output is used in high-stakes contexts that depend on reliability. Such care is exemplified by OpenAI (2023) to be reviews conducted by humans, adding context or completely avoiding using them when stakes are high. GPT-4 still misses subtle details, makes reasoning errors or accepts false statements from users (OpenAI, 2023). They show that it can fail to solve difficult problems like humans do, and the GPT-4 is not double-checking its work when it is likely to make a mistake. However, even though GPT-4 is not perfect, it is still better than GPT-3 at tests used as public benchmarks like TruthfulQA (OpenAI, 2023). Furthermore, they explain that the outputs from the GPT4 contain different biases and their efforts to manage and characterize these faults will take time. In all, they intend to make GPT-4 have sensible default behaviors that reflect the values of its users while also using input from the public to set boundaries for the systems.

### 2.5.2. Microsoft Testing ChatGPT-4

Bubeck et al. (2023) from Microsoft Research report in a paper their interactions with the GPT-4 for six months prior to its release, using the unrestrained model. Their study shows that the GPT-4 knows how and when to use different tools with little to no instructions. The authors give an example with a calculator and a search engine in which they ask who the current president of the United States is and the GPT-4 gives the correct answer by reading the snippets in the search engine, and in the snippets there were different names such as Trump, Harris, Biden. From this experiment, they suggest that it seems to know what tools to use and how to use them (Bubeck et al. 2023). This is one of the limitations that Wolfram pointed to regarding GPT-3 which now GPT-4 seems to have mitigated to some extent. Bubeck et al. (2023) testing of GPT-4 also demonstrates how it can make impressive estimations and guesses to Fermi questions, which is a mathematical problem requiring both quantitative thinking and general knowledge (Bubeck et al. 2023).

Bubeck et al. (2023) show that ChatGPT-4 proves to be a valuable assistant in some real world problems with an example of using it as a practical handyman. One of the testers (Bubeck et al. 2023) had a problem with their pipes in the bathroom and by going through a diagnostic together with ChatGPT-4, it found the problem. Bubeck et al. (2023) find that GPT-4 is good at incremental tasks but fails with discontinuous tasks, where incremental tasks can be exemplified as answers to factual questions or summarizing texts. The idea of a discontinuous task is a problem that requires an answer which cannot be generated in a gradual or continuous way, and it needs a novel way of looking at a problem (Bubeck et al. 2023). Examples of discontinuous tasks they give are solving a math problem with a new or creative method and coming up with a joke. Moreover, Bubeck et al. (2023) compare the discontinuous tasks to Kahneman's (2011) System 2-thinking. Kahneman (2011) presents System 1 and System 2, which represents two types of thinking. System 1 is the intuitive thinking which occurs automatically and without much or any effort and System 2-thinking requires attention, concentration and is used when solving complex problems (Kahneman, 2011). Consequently, Bubeck et al. (2023) perceive the ChatGPT-4 to master System 1-thinking, but fail to conduct System 2-thinking. They conclude that ChatGPT-4 is demonstrating early stages of AGI, despite being incomplete, since it is showing abilities which can be likened to human abilities.

OpenAI's two latest GPT-models have been presented in section 2.6.1 and 2.6.2, but it is worth noting that there are other companies that are developing LLMs, such as Google (2023) with their latest PaLM 2. PaLM 2 has been using other training data than GPT-4, where it differs the most on text from other languages, making it better at translation and linguistics (Google, 2023). This model has significantly less parameters, which are the connections in the neural network, than the largest PaLM model containing 540 billion parameters (Google, 2023). However, in some areas it is competitive with GPT-4, on presumably a comparable parameter count as GPT-3 (Google, 2023).

### 2.5.3. Open Letters and AI Legislation

As powerful AI technologies, such as the GPT-4, are showing improvements that are closing in on human intelligence, there are calls for slowing down the development of the technology while simultaneously catching up on how to manage it. In an open letter addressing AI developers all over the world, the Future of Life Institute (2023) vouches for a six-month pause in experiments on AI that are more powerful than the GPT-4. The open letter was published online on the 22nd of March in 2023 and signed by high-profile people such as Elon Musk and Steve Wozniak together with many other corporate leaders and academics. The letter's main point is to question what we as humans "should" do with the technology rather than what we "can", and specifically direct such questions to other actors than leaders within the tech-industry (Future of Life Institute, 2023). Additionally, they call for "robust AI governance systems" (2023, p. 2) to be put in place to regulate and oversee the development.

Currently, the first rules for AI are close to potentially being implemented in the European Union, as the so-called AI Act proposal is planned to be voted on by the European Parliament in the middle of June 2023 (European Parliament, 2023). The AI Act proposal has four objectives, with the first one addressing that AI systems should be safe and comply with already existing laws (European Commission, 2021). The second objective of the proposal is to enable investment and innovation in AI by establishing confidence in legislation, the third objective is to improve the enforcement of laws on AI systems and the fourth and last objective of the proposal is to assist the progress of developing a non-fragmented market for AI systems. These initiatives from many sides of society show that the control and management of AI development is something that is considered to be urgently needed.

## 2.6. Summary and Conceptual Model

The concepts of strategic decision making and AI found in the literature review were identified relevant in order to fulfill the purpose of the study. The most essential purpose of incorporating Barney (1991) and Eisenhardt and Martin (2000) is providing the connection from RBV to dynamic capabilities, which in turn involves strategic decision making, and should be seen as fundamental building blocks for the theory of which this thesis builds upon. Instead of going deep into the theory, a comprehensive description is found suitable for seeing the relevance for this thesis and theory that will be presented subsequently.

Mintzberg et al.'s (1976) decision making process constitutes the narrowed down lens for the discussion of the analyzed empirical results. The fundamental content from Mintzberg et al. (1976) is the decision making process which contains three phases which each have a set of central routines. Moreover, the concept of stimuli is important as it triggers the decision making process. Ivanov (2023) presents three approaches for AI involvement in human decision making which together with the connected characteristics for determining the approaches will be used. Critique, comparisons and thoughts on the ChatGPT-3 model by Wolfram (2023), in particular critiques on LLMs' computational ability, is provided to enable further discussion of LLMs utility in strategic decision making processes. Other technical reports (OpenAI, 2023; Bubeck et al. 2022) are included in the literature review to provide the reader with an understanding on what LLMs are, what their capabilities are as well as how large institutions such as OpenAI, Microsoft and Google view their LLMs and its current and future implications.

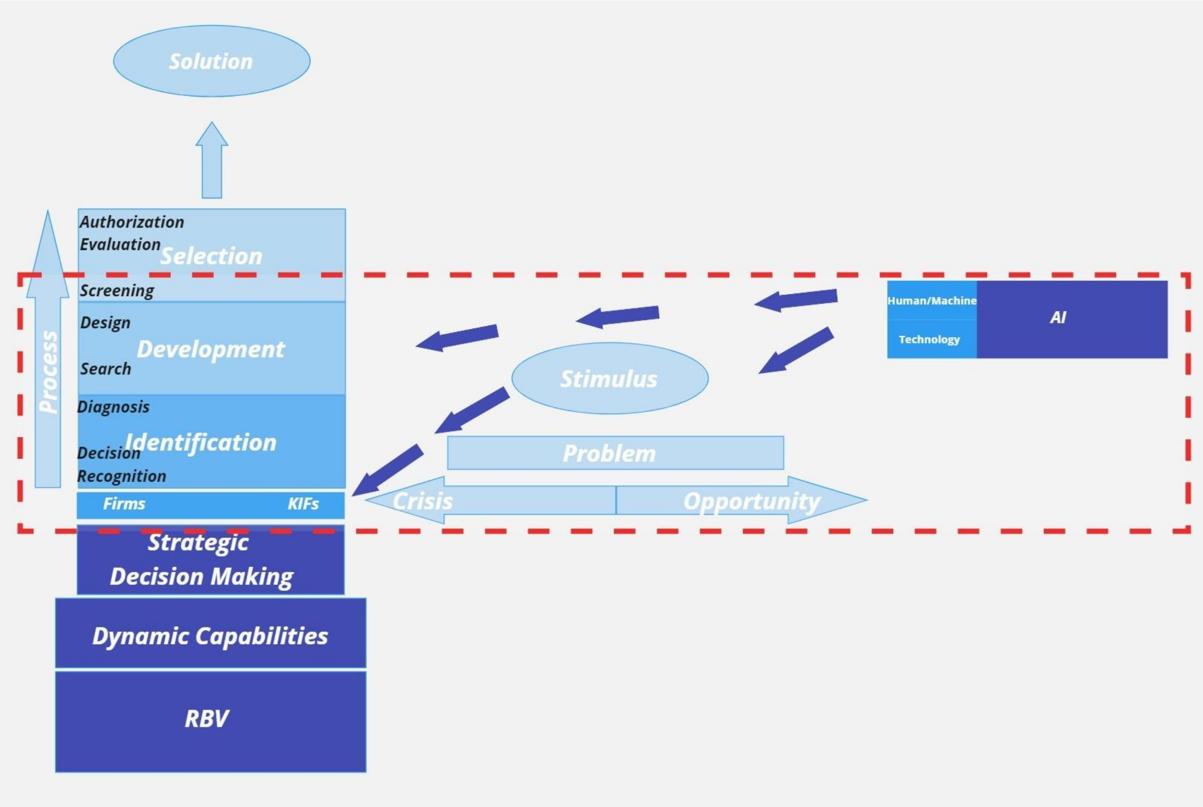



*Figure 1 - Conceptual Model*

The theoretical framework in the form of a conceptual model in Figure 1 is based on the key concepts which the literature review builds upon as well as on the key concepts used for discussing the empirical findings. The building blocks of the model are RBV, dynamic capabilities and strategic decision making and the previously argued connections between them are visualized. The red rectangle demonstrates the narrowed down lens of the conceptual model for which the analyzed empirical findings will be discussed. Followingly, firms and KIFs appear as the participants of the decision making process. Subsequently in the same direction, the decision making process of Mintzberg et al. (1976) is visualized with the phases and their respective routines. Note that the selection phase, excluding the screening routine, is outside the scope of the narrowed down lens for the empirical results. To the right of the process is the stimulus with the continuum which is affected by AI in terms of both the technology and the human and AI interplay. Lastly, the solution represents the output from the decision making process. Despite the final solution being outside the scope of the thesis, the development of the solutions are still mentioned in the development phase.

# 3. Methodology

## 3.1. Research Design

The purpose of this study is to contribute with an understanding of what role AI has, and might have, in strategic decision making processes, which led to the research questions:

> *How can Large Language Models be utilized in strategic decision making processes?*
> *What challenges of using Large Language Models are perceived by professionals in knowledge-intensive firms?*

For these research questions to be answered, a qualitative single case study was chosen for numerous reasons. A qualitative method is explained by Creswell and Creswell (2018) to be suitable when there are a multitude of unknown variables to study. This approach was deemed appropriate as the AI technology focused on in this study is showing rapid progress and becoming more accessible, thus being in a realm of uncertainty. Due to these characteristics of the study, we had reason to believe that the gathering of empirical data would stumble upon unknown parameters. Prior to collecting the empirical data, we elaborated on the fundamental theory for the study while keeping in mind that some instances of the theory could change depending on the character of the empirical material. An interplay between theory and empirical data occurred as the content of the data affected what theory that was considered needed in order to carry out a coherent analysis. This combination of a deductive and inductive approach can be classified as an abductive approach, which is explained to be a third way of reasoning within research (Bell, Bryman & Harley, 2022).

Additionally, a qualitative research design presents an opportunity to gather rich data, which stands in contrast to a quantitative study which instead favors generation of generalizable data (Creswell & Creswell, 2018). Again, as there is still much unknown about the use of LLMs in decision making in KIFs, getting deeper insights was found to be favorable over touching the surface on many points in order to come to a greater degree of understanding and thus fulfilling the purpose of the study. More specifically, a qualitative research design in the form of a case study was chosen as it is common when a researcher aims to acquire a comprehensive analysis of a case, such as an event or a process (Creswell & Creswell, 2018). Since the study involves KIFs and their use of LLMs, it was considered beneficial to not make a case with only one consultancy firm, but instead aim to gather empirical data from various consultancy firms or similar in order to gain an understanding of the context between the LLMs and KIFs. Despite gathering empirical data from different KIFs, the study can still be seen as a single-case study as it intends to create an understanding of a phenomena, namely the interaction between AI technology, KIFs and its clients in an early phase.

The main themes of the literature review of this study were also used as pillars when constructing the interview guide (see Appendix A), which was used when collecting the empirical material in the semi-structured interviews. This method of data collection is typical for research where the researcher is open to alter theory depending on the empirical data gathered, like the abductive approach (Bell, Bryman & Harley, 2022). The interview guide was slightly refined after the two initial interviews regarding the strategic decision section of the questions as we realized that the interviewees showed tendencies to refrain from speaking about their decision making as strategic. Consultants might not make strategic decisions per se, but they can be viewed to contribute to the foundation of strategic decisions of the clients, and are therefore in this study perceived to participate in the strategic decision process of their clients.

## 3.2. Method of Data Collection

### 3.2.1. Literature Collection

In conducting our literature review, we followed Creswell and Creswell's (2018) methodology of locating and summarizing. We used databases and keyword searches to gather a wide range of literature, then selected relevant articles for further analysis. We classified the literature based on relevance and created a visual map with categories related to decision making, automation and AI (see Appendix B). Our selection process involved snowball sampling (Bell, Bryman & Harley, 2022) as we used the found literature’s citations to guide what to review next. We considered both academic journals and research-backed reports as the goal was to build a strong theoretical foundation for our study by incorporating both broad and specific sources, primarily since the LLMs are developed rapidly.

### 3.2.2. Primary Data Collection

The semi-structured interviews were conducted with professionals in different KIFs in order to fulfill the purpose of the study of gaining an understanding of what role LLMs might have in strategic decision making processes. The interviewees were primarily selected based on them having either professional or personal experience of data science and AI, which became possible through a process of non-probabilistic purposive sample, in particular a snowball sampling (Bell, Bryman & Harley, 2022). In order to effectively get hold of potential candidates, our personal contacts were initially used and the study was promoted through a research proposal (see Appendix C). The purpose of the research proposal was to allow the potential candidates to see if they were applicable or to pass it on if they knew of someone who would be more suitable. This was also how the majority of our collection of the primary data occurred since an initial contact was made with possible candidates who were thought to be suitable, and sometimes the approached person accepted to participate but some suggested other, more appropriate candidates from their network instead. It was of importance to find appropriate interview candidates, which in this case is a rather limited group of people due to the nature of the technology and areas of its usage. The research proposal was thereby a helpful tool to improve the snowball sampling and mitigate any biases that could occur. For instance, the initially contacted people might have thought of the potential interview base to be too narrow or too broad if not for the research proposal. The platform LinkedIn also proved to be helpful to evaluate candidates' background and interest in AI. This evaluation occurred as an iterative process together with the snowball sampling, which also demonstrates a purposive sampling.

The interviews were based on an interview guide (see Appendix A) with open-ended questions to allow the interviewee to elaborate freely on the subject, and follow-up questions were asked in relation to some responses. In addition, all interviewees were ensured anonymity in order to allow them to feel more confident about sharing information with us and we have coded each participant (see Table 1). Our intention with this method was to gain as deep insights as possible from our interviewees, and Creswell & Creswell (2018) mean that this is favorable when the intent of gathering information lies between specifying the type of information that is to be collected and letting the information emerge from the participants. Moreover, the interviews ended with a question where the interviewee had the opportunity to talk about anything they thought we as researchers might have missed to ask. It was a conscious choice to include this question as we perceive the use of LLMs in consultancy firms to be in an initial phase with possibly many unknown dilemmas. Lastly, as the LLM is a novel area in knowledge intensive firms, we as researchers hold a limited knowledge, also not only regarding its application in the firms but also for the technology itself, which again highly strengthens the choice of open ended questions.

Some researchers, according to Bell, Bryman and Harley (2022), suggest that doing interviews face-to-face are superior to non-face-to-face interviews. However, Bell, Bryman and Harley (2022) argue that if those being interviewed are familiar with working online, this benefit is not as prevalent. We found offering to do interviews via online video-calls lowered the barrier of setting aside time for the interview. In addition, Bell, Bryman and Harley (2022) highlight that the Covid-19 pandemic encouraged more people to interact non-face-to-face. One downside of conducting the interviews online was the risk of internet disconnecting and consequently losing parts of the interviewees' responses, however this issue was not perceived to have affected the overall meaning of the empirical data necessarily. Since we used a snowball sample to gather the most suitable interviewees, there was a high possibility for candidates to be geographically far away, also considering that there are a limited number of people eligible for the study, as previously mentioned. The interviewees are presented in the following table in regards to their code names, positions, short description and what region they operate in.

| Code name | Position | Description | Region |
|---|---|---|---|
| **Delta** | Manager, Data Scientist and Responsible AI Lead | Leading a responsible AI launch at one of the major consulting firms. PhD in Computer Science. | Nordics |
| **Epsilon** | Senior Data Science Consultant | Generative AI expert. | Nordics |
| **Eta** | IT Solution Delivery Lead | Responsible for the automation by Robotic Process Automation and Natural Language Processing solutions. | Southern Europe |
| **Gamma** | Junior Management Consultant | Works with technology transformations such as technology enablement of strategic visions.<br>Engaged in internal discussions on AI. | Nordics |

| | | | |
|---|---|---|---|
| **Iota** | Big Data & Data Science Cloud Architect | More than 20 years of experience in a large international software firm. | North America |
| **Kappa** | Junior Management Consultant | Consultant within Financial Services, specifically Business Transformation projects. | Nordics |

*Table 1 - List of Interviewees*

## 3.3. Data Analysis

When attempting to make sense of qualitative data, Rennstam and Wästerfors (2018) put forward three problems which should be addressed by a researcher through three suggested activities. Firstly, they describe the problem of chaos as having significant amounts of empirical material which is not yet structured, and to solve this problem they suggest the activity of sorting by dividing or categorizing the material. Secondly, they explain that the problem of representation applies to the issue of not being able to include all the empirical material in the study, and thus they describe a need for the activity of sorting out what is considered less relevant in the particular study's context. Lastly and thirdly, they put forward the problem of authority which concerns what the study can contribute with in relation to previous research, and to tackle this issue they encourage using argumentation with a framework based on the material.

When analyzing the qualitative data of this study, we have used Rennstam and Wästerfors' (2018) concept of problems and activities as an overarching principle throughout the empirical processing and analysis. More specifically, we have followed Creswell and Creswell's (2018) five-step process in order to break down our data. In accordance with their first step, we organized and prepared the raw data for analysis by previously having recorded the interviews and importing the recordings into an automatic transcription tool. We then manually went through the produced transcript along with the audio-file in order to make corrections wherever needed to. This procedure made us get familiar with the data while ensuring our basis of analysis would be as close to the truth as possible, and led us into step number two of Creswell and Creswell's (2018) process in which the researcher reads and forms a general sense of the data collected. All interviewees' transcripts were gathered in one document in order to facilitate searching for specific words. At this point, we started to get an idea of what the empirical data could be used for in this study and highlighted sections we found particularly important. Coding the data is Creswell and Creswell's (2018) third step, and at this stage we put similar sections of the data into categories. Next, following their fourth step of generating themes, we created themes from the categories found in step three which serve as the main empirical findings of this study. When finding themes, Bell, Bryman and Harley (2022) suggest that focusing on repetition of topics across data sources can be useful. However, they stress that repetition in itself does not guarantee a theme to be relevant to the research questions. Additionally, Ryan and Bernard (2003 cited in Bell, Bryman & Harley 2022) suggest that researchers should observe if interviewees talk about a topic similarly or differently and see what data that might be missing. Finally, in Rennstam and Wästerfors' (2018) fifth step, themes are represented and conveyed by explaining the themes in greater detail. In the next chapter, the empirical findings will be presented and analyzed, and is followed by a discussion in chapter five.

As we utilized a snowball method of sampling interviewees, there was not ever a fixed list of participants determined before conducting the interviews. Instead, it was an iterative process of conducting and analyzing interviews throughout the study. Bell, Bryman and Harley (2022) describe that this interplay between collecting and analyzing data is typical for qualitative analytical processes. Bell, Bryman and Harley (2022) underline that thematic analysis is a rather flexible method of analysis and thus is the researcher's transparency of how the analysis is conducted even more important. Measures we have taken in order to strengthen the quality of the study will be further elaborated on in the following section.

## 3.4. Validity and Reliability

The quality of a study can be discussed in various parameters, and which parameters that are considered suitable varies from case to case and from writers to writers according to Bell, Bryman and Harley (2022). Creswell and Creswell (2018) present several validity strategies a researcher of a qualitative study can conduct in order to evaluate the precision of the findings of the study and provide the reader with a possibility to assess the accuracy of these findings. Adopting a selected number of their suggested strategies, we have strived to present the reader with such tools of assessment. One strategy of Creswell and Creswell (2018) that we have followed is providing a full and detailed description of the findings to guide the reader within the context of the study. In these descriptions, we have also presented information that is in contrast to our themes as Creswell and Creswell (2018) suggest that including several perspectives when discussing themes can make the discussion truer to real life. Another strategy they suggest is to be transparent as a researcher about any bias one might bring. As we both come from very similar academic backgrounds, there was a risk of having too similar approaches to processing and interpreting the empirical data, and thus risk being narrow-minded. Therefore, it was a priority to regularly question one another’s interpretations.

The reliability-parameter is further divided into external and internal reliability by LeCompte and Goetz (1982 cited in Bell, Bryman & Harley 2022) where external reliability refers to whether the study can be replicated. Case studies are difficult to repeat since their purpose is to create an understanding of the complexity of a particular case (Bell, Bryman & Harley 2022). However, by providing descriptions of the methodological processes of the study we have aimed to facilitate similar future cases studies. Internal reliability refers to whether researchers are in agreement of what they observe (LeCompte and Goetz, 1982 cited in Bell, Bryman & Harley 2022) and the analysis process described in section 3.3 was utilized to guide us through what we could observe from our transcripts.

# 4. Empirical Findings and Analysis

In this chapter, the empirical data gathered from the six conducted interviews will be presented and analyzed in order to answer the study's research questions:

> *How can Large Language Models be utilized in strategic decision making processes?*
> *What challenges of using Large Language Models are perceived by professionals in knowledge-intensive firms?*

To provide context, a shorter description of the different interviewees will be presented firstly, followed by the empirical findings and analysis structured by themes.

## 4.1. Case Description

The six interviewees all work in knowledge-intensive firms, more specifically professional service firms as they work in close relation to their market clients and customers as defined by Alvesson (2004). A majority of the interviewees currently work as consultants, while some, Iota and Eta, work in companies creating technical solutions for their customers. Most of the respondents are operating in countries in the Nordics, with the exceptions of Eta who is based in Southern Europe and Iota who is situated in North America. The level of seniority and experience varies across the respondents, for example Iota has been in his line of profession long enough to know "what business was like pre Internet" while the junior consultants Gamma and Kappa graduated recently. It is typical for the interviewees with seniority to both work internally and externally as Delta leads an internal AI project and Epsilon, as expert on AI, advises managers internally, while the junior consultants primarily work externally with clients. Even though the interviewees may not make the final strategic decisions, they are perceived to support the strategic decision making process as they advise their clients and customers in strategic questions. An overview of all the interviewees is illustrated in *Table 1* in section 3.2.2.

## 4.2. Interest and Potential Use

### 4.2.1. Increased Interest Among Clients

One common aspect of the interviewees' descriptions of the emergence of accessible LLMs is an increased interest in the technology among clients and customers. Iota described that there are more "executives and vice presidents" who are interested in utilizing AI in their businesses in a strategic manner, rather than a tactical manner, due the access to ChatGPT. Iota further stated that when it comes to demand, he is “seeing it from everyone” and mentioned industries such as retail, healthcare and manufacturing. Furthermore, he emphasized that corporate leaders are asking him specifically about how LLMs can be adopted into their business as they express an awareness of having to do it. He also described that there is a “fundamental change happening right now, that will play itself out over the next six months”. Moreover, Delta explained that it has sparked discussions and raised questions about the potential changes and impacts of AI on various industries and business processes. He further described that business leaders are particularly interested in understanding the value that AI can bring to their processes and thus ensuring they do not miss out on opportunities and remain competitive in the market. Iota and Delta’s statements about increased interest among customers is coherent with how Epsilon perceived the change. He has been working with “Traditional AI” until recently, and in the last six months he has been fully working with Generative AI, ChatGPT and similar technologies. This is another empirical example that demonstrates how the increased interest in AI, in the form of ChatGPT, due to its launch six months ago have contributed to his shift of focus within AI technology. He further mentioned that his experienced increase of attention can be related to his prominent position within the generative AI field. He also acknowledged that people within his firm are starting to look at generative AI as an “interesting business area”. This shift of focus is further supported by the following quote from Epsilon:

> "[...] more recently, a lot of companies have been asking me, we see ChatGPT, we see these generative AI technologies, what does it mean for us?” (Epsilon)

However, both Delta and Eta acknowledged that AI technologies have been developed for many decades, which may stand in contrast to the perception of it being a sudden increase in demand. Nonetheless, Delta further explained that as the accessibility of AI has increased with tools like ChatGPT allowing people to interact and experience a small part of AI firsthand, its popularity has increased: “It is just a website that you can look at, and start chatting with this tool.”. In contrast, both Gamma and Kappa described that they have not experienced any increase in demand of implementing AI in their work with clients. As they both are junior consultants, this could signal that such requests are primarily directed to those who are more senior and are experts in the AI field, such as Delta, Epsilon and Iota. It could also be an indication of that not all firms seeking consultancy are interested in implementing AI currently since Eta recognized that he has not experienced any change in demand from his customers.

### 4.2.2. Using AI Strategically

As the empirical data show that there is an increase of interest among clients in using LLMs strategically, a further analysis on what questions the clients are asking the interviewees follows. Epsilon described how the goal of a project is not always clear to the client when approaching him with it: “A lot of companies already know [the goal] whenever I'm entered into the process, but sometimes they don't.” Additionally, Epsilon described that clients do not only approach him when they are unfamiliar with what type of resource LLMs could be, as there are also those who already have started to use the technology to some extent but have gotten “cold feet”, questioning whether they are on the right track or not. He explained that this uncertainty could be attributed to the rapidly changing field of LLMs, and that he, in his role as an expert, can help to ascertain whether they are “going down the right path”.

These types of questions that clients bring to Epsilon can be seen to have strategic elements in them as they regard goal-setting, and Epsilon explains that these questions are becoming more common than before. This is further strengthened by how Epsilon specifically described typical inquiries to entail figuring out what impact it might have, what the opportunities, challenges and risks are together with staking out a path towards a potential future. After describing this typical request from a client, Epsilon himself identified that this process is similar to strategy. Epsilon also used words like “confusion” and “misunderstanding” to describe the state of mind of many clients. This is an additional indication of the need for guidance in an unexplored area, or perhaps an area where clients lack a significant portion of the knowledge needed to create a goal for it. This empirical example further strengthens the notion that the implementation of this type of AI technology in firm’s is still in an initial phase and there are a lot of unknowns to detangle.

Delta divided his projects into two categories, the first one entailed supporting businesses in their decision making processes with machine learning and the second encompassed helping on a strategic level by examining how data and AI analytics can be used in the client’s business. Thereby, both Epsilon and Delta described how AI can be used strategically for their clients. However, Eta acknowledged that it is not always straightforward to find how LLMs can be useful when scaling them. Furthermore, Eta explained that his strategic decision-making is driven by the goal of maximizing added value while minimizing costs. Epsilon believed that clients use LLMs primarily on an individual level, and that it is hard for companies to use them on an organizational level. Once again, he describes that “people are just getting out of this sort of starting hole at this point" and that the issue currently is in a "sketching phase". Clients have been asking Epsilon why not everybody is using ChatGPT currently, and he then typically answers that there has not been enough time to try it out yet.

Gamma mentioned that consultants have advised clients through different technology shifts and that they may have the upper hand in identifying use cases for generative AI for the clients and thereby may play a role in drafting better AI-solutions. Iota gave a similar statement where he explained:

> "[...] anybody who helps influence these decision makers and advises them and consults with them, has to help them understand how this technology can help their specific business needs." (Iota)

Iota continued by adding "the ones that adopt it the quickest are going to gain competitive advantages." Iota gave examples of firms using LLMs to support their software and products currently are those in the startup community and other software firms. He means that technologically-heavy companies are the ones that most aggressively can adopt LLMs. These descriptions by Epsilon, Delta, Gamma and Iota strongly indicate that expert consultants are approached with strategic questions regarding how to implement LLMs and firms who are able to incorporate such technologies might gain competitive advantages from it. Consequently, it could be suggested that consultants might have a significant influence over how this technology shift will play out.

### 4.2.3. Analyzing the Situation of Clients

Delta emphasized that AI is not meant to make decisions autonomously; rather, it is designed to support decision making and operational tasks in various domains. He also described that when humans make decisions, they rely on substantial information, identify patterns within extensive datasets, evaluate various options, consider potential outcomes, and then arrive at a decision. Delta also mentioned that in this decision making process, AI can be valuable by assisting in tasks such as processing large amounts of information. Delta, Epsilon, Gamma and Kappa stressed that LLMs can provide helpful answers to specific queries, serving as a tool for accessing information more efficiently. In addition, Iota viewed AI to be "beneficial for asking questions and getting information". Delta mentioned that without AI support, individuals would need to invest considerable time searching the internet to gather the same amount of information. Delta highlighted LLMs' promising potential and presented a concrete example of his firm collaborating with a startup founded by OpenAI, to customize language models, specifically ChatGPT-4, for supporting legal use cases thus already bringing transformative changes to the way legal consultants work.

Delta and Epsilon also mentioned that they are working with projects in a variety of industries. Delta expressed that it is hard to understand the business need and translate it into a technical solution, even if more obvious cases also exist. Delta explained how data is used to help the clients to make a better decision through creating support for decision making, not to replace or make the decision for them. Delta emphasized that data is used to make advanced analytics or machine learning models to forecast the firm's financial status, on the premise that they have a “sufficient amount of good quality data”. Iota mentioned that he uses data if he can, and similar to Delta it depends on the quality which in turn could be dependent on the maturity of the product meaning that if there is a lot of data on for instance the usage of a product, data will have a stronger say in the analysis. For new products this might not be the case. Iota mentions that in cases regarding new products, the clients are asking questions without underlying data, which could be interpreted to be the case when it comes to the usage of LLMs in products. Several interviewees, Delta and Epsilon for instance, have expressed that they need to handle a large amount of information or data. Or as Iota expressed, that the client might expect the firm to provide answers without data, which leads to an iterative process of seeking data. Eta and Delta mention perceived payoff and uncertainty, respectively, where Eta states perceived payoff as a single handedly determinant if the project is pursued or not. He once again emphasized the importance of return of investment when reviewing whether to apply AI as he described how “technology is at the end a cost”.

Delta described that his team works closely together with the business experts from the client side to make sure that the business insights are in interplay with the data, AI, and machine learning, when creating the models. The business experts help provide the context and understanding of the business processes as the numbers alone hold little meaning without that context. This can come from experts within their own firm, but also by involving clients with intimate knowledge of their own business. Epsilon, Eta and Gamma also bring in clients or business intelligence experts for similar purposes as presented by Delta. Using input from experts or clients can be helpful as a benchmark or objective, used to make reasonable assumptions according to Delta. The interpretation of the results from the model also requires human expertise or business insights, which is not quantifiable, but something that is rooted in the firm's knowledge about their own business. Iota also emphasizes the combination of data and how the quality in terms of robustness and accuracy determines how much it can be relied on.

Delta mentioned that the information derived from data and human expertise can challenge and complement each other. On one hand, humans may overlook hidden trends in large datasets, while on the other hand, assumptions, models, and data choices made by data scientists may be questioned by the business side. This interplay between data and human expertise serves to refine and reach a consensus on how to effectively combine both elements to make better decisions. Delta acknowledged that if there are contradictions between the data and human expertise, it suggests that something is amiss. Resolving such contradictions necessitates discussions between the technical and business teams to explore potential solutions, reflecting an iterative process. Delta also noted that encountering contradictions is not necessarily a negative outcome, but rather an indication of progress. Simply confirming preconceived notions held by the business side would render the use of data and AI redundant. This indicates that Delta utilizes a combination of AI and human intelligence, where AI is assisting the human in the decision.

## 4.3. Generating Solutions

### 4.3.1. Looking for Solutions

Eta emphasized that his involvement is more focused on decision tools for project execution rather than conducting business analysis of processes. He further explained this by describing that when he receives a business analysis from a colleague regarding a processing target, his role is to identify the next steps in developing the specific business needs. If he already has an existing application that can support a specific business process, his approach is "not to reinvent the wheel" but rather to utilize and modify what he perceives to be needed to accommodate the business process. Additionally, Eta focused on leveraging the existing environment to fulfill new business needs rather than creating something entirely from scratch. This can be related to his previously explained view on minimizing costs and maximizing added value to guide his strategic decision making.

Epsilon describes his decision making process as a typical innovation process which has a basic underlying structure even if it varies a lot depending on the case. In this process, he is typically responsible for killing off the many ideas that are initially generated. Epsilon explained that he uses his “gut feeling” in this process, but would rather describe it as something based on experience. Furthermore, he uses it in the first filtration in the feasibility phase where ideas are killed off, and he does not test every single case.

Eta exemplified a strategic decision process for a feasibility analysis that he conducted for the policy intake process by explaining that his task was aimed to identify opportunities for improving efficiency and accuracy in comparison to the manual process. Eta also makes a point about how the complex nature of his organization, and in turn understanding the existing processes and available resources, can pose a challenge in his decision making. Drawing from his previous experience as a Global Practice Leader for AI RPA (Robotic Process Automation), the focus was placed on leveraging a natural language processing platform. To determine the most suitable application, three prototypes were created and tested using actual documentation. Through a comparative evaluation, considering factors such as cost, ease of learning, maintenance, and quality of execution, a decision was made to proceed with the chosen solution. emphasizing considerations of cost, effort, time, and desired quality outcomes.

Eta implemented automation in processes that are repetitive as an initial use of the new technology. Additionally, he stated that “decision making for us in this moment is adjusting the classification of the documents in the policy intake [...]”, which shows that he is allowing AI to do some stages of the decision making process, but that in the end he controls and modifies the output. However, Eta did not specify what stages of the decision making process that the AI is doing.

### 4.3.2. Creating Solutions

Delta pointed out that there is a common misconception about AI, where people tend to associate it solely with tools like ChatGPT and LLMs. Delta further emphasized that AI encompasses a much broader scope, ranging from simple statistical analysis to more advanced machine learning models, including techniques such as classification, forecasting, clustering, and deep learning computer vision and various other complex technologies. Epsilon mentioned a similar scope of AI technologies when he described his previous focus to concern traditional AI and his new focus on generative AI. Epsilon further explained that when approached with a project or problem, he decides whether it is an AI problem or not. Delta suggested that it is important not to overlook these other aspects of AI when discussing and considering its potential applications and impact. Other tools could be more applicable or the best practice solution. Gamma gave a concrete example of when the business case spoke for a manual change of robot in RPA instead of using generative AI. However, Gamma's business case was before the widespread introduction of LLMs, in contrast to Eta's RPA project. Other than the given examples, the generating of solutions is described on a general level. Iota mentioned that every case is unique and its character depends on the client and both Epsilon and Kappa said that even if there is a basic underlying structure the process varies between use cases. It can be stated that neither of the respondents, Delta, Epsilon Eta, Kappa, Iota and Gamma, are forcing the use of LLMs if it seems to not be applicable. Additionally, in some cases, when the project's architecture is already established, Eta explained that he may collaborate closely with business analysts to delve deeper into the business processes and suggest adjustments or changes from an IT perspective. This suggests that the interplay between AI and human expertise is present when creating the solutions, just like when analyzing the client's situation.

## 4.4. Artificial Intelligence and Business Insights

Delta discussed the application of ChatGPT in financial planning and emphasized the need to understand the decision making process involved. He suggested that ChatGPT could potentially be used to support specific steps in the process for the purpose of providing better information, insights, and understanding. Furthermore, Delta highlighted that ChatGPT alone cannot replace the entire financial planning process, and it should be viewed as a tool to assist and enhance decision-making rather than making decisions on its own.

The absolute majority of Epsilon's use of LLMs is in the form of a partner that can structure thoughts and ideas. Epsilon could also use it as a tool for gaining general information about an industry or as a coding tool. Eta and Iota also expressed a similar use when it comes to coding. Gamma highlighted internal efficiency and a new way of doing things as an added value from LLMs. Iota and Kappa stated that when it comes to more advanced analytical tasks, the LLMs lose context and as Iota expressed: "fails miserably in understanding the context of an analytical type question". This indicates that the LLMs can be used to a certain extent, but is limited when a task becomes too complex.

Furthermore, Delta mentioned that data and AI is not anything magical, rather science and mathematics, still acknowledging that risk for unexpected events always exists, such as financial crisis or Covid which creates disruptions in the markets. Delta expressed that it is very difficult technically speaking to incorporate uncertainty in the market in these models.

Delta means that it has to be communicated to clients that the models are scientific forecasts based on statistics, not magic. On the note of magic, it is interesting to draw a comparability between Epsilon and Delta on the use of the word magic in the context of AI and LLMs. Where Delta communicated that the solutions are not magic in order to not oversell the technical solution and the risks that uncertainty implies, whilst Epsilon may tell his clients to pretend that AI is magic, in order to open up for creativity. Gamma also expressed that opening up for creativity is a direct use of the LLMs. This again demonstrates that the consultants may work with clients, starting from different stages in the decision making process. It demonstrates that the view on AI and LLMs may be subject to change in relation to how developed or critical the decision making process is at that stage. The view on AI and LLMs can be of a more leisurely nature in the early stages of a process, for instance in the decision recognition routine. The further the process goes, the more important it becomes to concretise expectations which Delta demonstrated by the use of the word magic.

Epsilon describes that he sees two potential different impacts LLMs might have in the future. The first one, “broad impact of ChatGPT and similar technologies” would entail everyone having constant access to a “conversational partner” to be used in many peoples’ daily work. He further describes that this type of use would not make any major changes in workflows or enable new products, but is a helpful tool which “increases productivity and quality of work”. All of the Respondents are already doing it or wish to use it in this way. Epsilon further estimates that this type of use is starting to take place in society already, referencing ChatGPT being the fastest growing app ever but still noted that many people still do not use it. The other impact would be if LLMs were used in new products and services and by doing so getting a new set of revenue streams or “fundamentally changing products”.

Epsilon highlighted a discussion concerning if there are going to be firm specific models or if there is going to be one big universal model for all companies. Additionally, Gamma showed skepticism for unique and specialized AI software and stressed that it needs to be user friendly, potentially integrated into already used applications in order to simplify the use for corporate clients.

## 4.5. AI Challenges

All of the interviewees expressed various challenges that are attributed to how LLMs can be used in strategic decision making processes, and the major categories of these challenges will be presented in this last part of the empirical findings. In particular, Delta highlighted the importance of considering the risks associated with AI adoption, such as data privacy, cybersecurity, ethics, and financial risks. Delta noted that these discussions encompass not only LLMs, but AI tools in general. The empirical findings are perceived to indicate that the development and wide uses of LLMs are in its initial stage, despite AI having been in existence for many decades due to the notable challenges.

### 4.5.1. Data Privacy

One issue that all respondents perceived as a challenge when investigating whether to implement LLMs is if they are safe to use in regards to data privacy. Epsilon brought up that an important consideration for him was how much of the client's info that could be shared with the LLM as he stated: “We are really, really sensitive about sharing client information”. Thus, he is limited when using an LLM if the data is closely related to client information, and he described it to be a “big limitation with the data sharing”. Likewise, Gamma emphasized that if he cannot use client data within ChatGPT, uploading documents, summarizing and comparing data the functionality of it disappears and he is therefore not using it in his work even if he sees great opportunities in using it:

“If I can't use client data within the generative AI or sort of ChatGPT-style, then I don't think it's really beneficial at all.” (Gamma)

Eta agreed that privacy is an associated risk with using ChatGPT, and mentioned that Italy temporarily banned it because of the GDPR. Eta, Gamma and Iota also expressed a need for centralized efforts on working on this problem as it is not only a technology problem, but also a regulatory issue, while Gamma mentioned a similar point but stated it to be a rather short term issue.

It should be noted that internal use across teams and business areas may be liberated from client related information thus enabling the use of LLMs. However, as Iota highlighted, the clients pose the question of "How to use it for decision making?" which is for "internal purposes to make decisions based on their proprietary and corporate confidential data". This suggests that clients feel that they may lose control of their data and Iota expressed that they want to bring this capability in-house instead. By extension, it could be assumed the issue of data privacy entails intellectual property issues. This is also mentioned by Epsilon as he sees that there is a division of clients that are willing to trust the cloud or not. Epsilon stated regarding the accessibility for the companies, if there is a need to leverage on LLMs instantly, it is currently only possible through a cloud service. Individuals are not under similar restrictions or policies as companies could be. Eta who operates in Italy did for instance get legally restricted by the Italian government, and by extension so was the firm. In all, this suggests that firms and individuals' contexts differ and will have different implications.

Additionally, one of the reasons why LLMs are not being used on a broader level is data security according to Epsilon. He calls it a “quality tool limitation” and refers to many LLMs being cloud-based, which is a security concern for large companies according to Epsilon. The solution is to use open source alternatives, and if they catch up it will not be as big of a problem as it is currently. This is another type of concern in comparison to what Eta and Gamma posed in the previous paragraph. Here, Epsilon aims at the concern of the actual use of it rather than its technical capability. Eta, Gamma and Delta have all expressed this concern when they bring up data privacy as a concern, however not stating the quality tool limitation as being cloud based as explicitly as Epsilon does.

Eta also explained that AI is an algorithm which will use the data that humans share input with, and that one should be careful about these datasets as the quality will have great importance for what its output will be. In relation to the issue of data privacy, he highlighted the need for high-quality data sets and caution against potential biases or inaccuracies that can affect the performance and outcomes of AI models. Gamma mentioned a similar concern where he sees a risk in not being able to detect or validate output.

### 4.5.2. Technological Capability

In addition to the concern of data privacy, the interviewees expressed concerns of the technological capabilities of the current LLMs. In particular, Iota and Kappa described the technological capabilities to be an issue with more advanced analytical tasks since the LLMs tend to lose context in such tasks. Delta expressed concerns of “uncertainty in the market, and how to incorporate that into your model” being very difficult technically speaking. Moreover, Delta did not see that AI will make decisions in anyone’s stead, but instead emphasized “that there must be human autonomy in this decision making process”.

Iota mentioned that the clients are questioning the technical ability in terms of how good LLMs are, what tasks they are good at, how they verify whether it is good and if that leads to any additional processes.

Iota explained that his firm is still in an evaluation mode, “still just learning about how to use it” and trying to “understand its capabilities and how it can be used effectively”. Additionally, he emphasized that a small change in the structure of a large, public firm can lead to big impacts, thus taking on a conservative approach. Iota also mentioned a unique technical challenge which none of the other interviewees mentioned. This technical challenge is adding a last step in automation by turning the output it generates into an action, as he exemplified that the LLMs would be “more useful, from a decision making turning into an action perspective” for tasks such as exchanging one hotel room for another without involving a human. Iota explained how it needs to be integrated in other systems if turning decisions into actions is to be possible. Likewise, integrating AI into already existing applications is something that Gamma stated to facilitate the adoption of AI.

Delta emphasizes the importance of ensuring the quality and reliability of the outputs generated by ChatGPT when used in strategic decision-making processes. Delta also mentioned the importance of considering potential use cases and challenges in different industries. He gave an example from the healthcare sector, where ChatGPT could be used as an alternative to phone helplines, providing quick and accessible answers to questions. However, he noted a need to address risks associated with ensuring the safety and reliability of advice provided by LLMs: “how do you guarantee the safety of the advice?”, which suggests that it could be too risky to use LLMs in contexts where the output cannot be incorrect.

However, Epsilon provided a slightly different view on the challenges of technical capabilities of LLMs as he considers knowledge to be the main obstacle in order to understand the technology and see its potential. In like manner, Eta argued that the level of skill influences LLMs capabilities to a large extent: “[...] if you haven't got skills, what can you do? Nothing.”. In all, the different challenges brought forward by the interviewees indicates that the empirical data support the existence of many potential improvements of the LLMs in need of attention.

### 4.5.3. Other Challenges

Delta acknowledged the need for non-technical aspects to catch up with technological advancements as he said: “[...] let's buy us some time for the non-tech part to catch up with the tech advancement”. This indicates that challenges associated with LLMs are closely connected to the assumption that it is in an initial phase. Moreover, Delta mentioned a need for better regulations, risk management standards, ethics models, and assurance and auditing systems to ensure responsible and beneficial AI use.

Delta suggested that implementing ChatGPT or any language model in critical areas requires careful consideration of the potential risks and the development of frameworks to guarantee the quality and safety of the AI-generated outputs. Delta explained that such critical areas could be in industries where safety and control are paramount, such as finance, banking, insurance, healthcare, and critical infrastructure sectors like telecom and energy. He expressed that it is crucial to have a holistic framework in place to ensure that AI is effectively managed to mitigate risks and that the outcomes generated by AI have a positive impact on society and the environment. Delta introduced the concept of responsible AI, which involves using AI in a responsible and ethical manner. He emphasizes that responsible AI encompasses multiple dimensions, including regulation, risk management, building trust, establishing standards and guidelines, and linking AI usage to corporate social responsibility.

Epsilon described that the main obstacles for using AI on a wider scale are knowledge, creativity, understanding the technology, identifying the problems, becoming aware of the potential and lastly making the solutions. Time is another obstacle that Epsilon identified for this technology to be used on a wider scale.

Eta recognized that there are many opportunities to apply AI technology in his line of work, in particular in decision making in classification of documents, risk management, analytics and data processing. However, he recognized that there still are many limits to it. The reasons for the implementation of LLMs not being extended into more areas and departments according to Eta are cost, time and effort. He also mentioned that “skills" and “capacity” are necessary, and that there are also cultural differences when it comes to the level of skills, in particular in Europe. He described a need for people to become educated about AI in order for it to become less of a hypothetical tool to many. In addition, Epsilon expressed that individual employees could do structured training on how to use these LLMs, but there is no particular interest currently. Gamma on the other hand mentioned prompt engineering as a skill that will be more and more helpful and sought for. Iota mentioned that the change is in all a big cultural shift for the clients since it regards a big shift in decision making processes. Iota explained that the client's adaptation will be slow as the technology matures.

Delta suggested that implementing ChatGPT or any language model in critical areas requires careful consideration of the potential risks and the development of frameworks to guarantee the quality and safety of the AI-generated outputs. Delta explained that such critical areas could be in industries where safety and control are paramount, such as finance, banking, insurance, healthcare, and critical infrastructure sectors like telecom and energy. He expressed that It is crucial to have a holistic framework in place to ensure that AI is effectively managed to mitigate risks and that the outcomes generated by AI have a positive impact on society and the environment. Delta introduced the concept of responsible AI, which involves using AI in a responsible and ethical manner. He emphasizes that responsible AI encompasses multiple dimensions, including regulation, risk management, building trust, establishing standards and guidelines, and linking AI usage to corporate social responsibility.

# 5. Discussion

The following chapter discusses the central analyzed empirical findings in relation to previous concepts outlined in the literature review.

## 5.1. Accessible LLMs as a Stimuli

It is of importance to highlight that the entire decision making process is iterative and recycling (Mintzberg et al. 1976), and likewise the routines (Mintzberg et al. 1976; Eisenhardt and Martin, 2000). Stimuli and the identification routines are in many ways interconnected where their relationship could be described in the following way: stimuli are all individual things that give clients impressions, the decision recognition marks the point for the accumulated stimuli when action is taken and the diagnosis routine aims to comprehend the stimuli through a cause and effect lens (Mintzberg et al. 1976).

### 5.1.1. Stimuli

The empirical findings clearly give an indication of increased interest in implementing LLMs since ChatGPT was made more accessible in 2022. More specifically, the analysis shows that such questions regard what use they can be for firms but also what they actually are. The empirical results show that many of the respondents are experiencing an increase in questions and requests to become more informed and create value. Connecting to the description of Mintzberg et al. (1976) concerning the decision process, this shows that corporate leaders are experiencing certain amount of stimuli from the emergence of accessible and powerful LLMs, which encourages them to act managerially to position the decision process on the continuum closer to opportunity rather than crisis by being passive. However, it should be noted that the junior consultants have not experienced the same increase in questions about LLMs, which could indicate that their clients still have not accumulated enough stimuli, which Mintzberg et al. (1976) describe is needed to take action, which happens when the managerial threshold is reached.

### 5.1.2. Decision Recognition

From the empirical results, it can be said that the clients of the senior data scientists interviewees, all except Gamma and Kappa, bring questions that intend to give answers on the implication of the LLMs for their firm. The questions vary in specificity depending on how easy the client's managerial threshold has been reached through accumulated stimuli, as described in Mintzberg et al. (1976) decision making process. Furthermore, they describe the matching phenomenon found in the recognition routine. In the context of this study, the phenomenon of AI together with a related crisis, problem or opportunity can determine if the clients seek professional help from consultants or not.

Determined by the accumulated stimuli and how high the managerial threshold for action has been, the posed question to the decision makers can be of an AI nature and a non-AI nature as described by Epsilon. The AI nature can thereby further  be divided into the traditional AI or LLMs, which makes LLM solutions or uses, one of three scenarios. No probabilities are discussed in this sense, it is simply used as a classification. In general, the empirical findings indicate that none of the clients of the interviewees seek consultancy in the crisis-problem spectrum of the continuum (Mintzberg et al. 2023) caused by AI, rather it seems to be more to the problem-opportunity spectrum. This interpretation rests on the time aspect of the phenomenon, where both the questions posed and the answers by the consultants revolve around a future scenario, not an immediate action. It is thereby not viewed as a crisis since clients seem to have time to read the situation. It can therefore be said that the question varies depending on the industry they operate in where different industries have reached different stimuli levels.

To illustrate with two examples on two extremes of matching quality, a high matching quality is the use of LLMs as a tool supporting legal decision making processes. Statements from Interviewee Delta, the tech report (OpenAI, 2023) and Microsoft's testing (Bubeck, 2023) demonstrates the legal area as to be one of the most prominent ones to have a direct use of LLMs. Factors such as lower uncertainty and high payoff could make the use cases reaching a high accumulative amplitude of stimuli. Judging from how far the decision making process has gone for the specific legal case which Delta mentioned, the questions and answers can be interpreted to have had a good match, thus leading to a fulfilled strategic decision making process. On the contrary, a low match quality example could be from areas where the use case is unclear, and the question asked by companies in such industries is simply what LLMs are and if the clients firm could have a use for it somehow for which there might not be a good answer, as Epsilon and Iota expressed.

Since the stimuli levels varies depending on the industries, but seemingly according to Iota, a broad range of industries are seeking help, and according to Epsilon and Iota some such inquiries rather basic, it can illustrate that the managerial threshold varies and that the more stimuli a client or industry has gathered, the faster and easier the firm can continue in the decision making process. It also shows that for those seeking consulting, there must not be a clear matching problem with opportunity to LLMs, simply the introduction of LLMs has led to it single handedly as a phenomenon, reaching a decision recognition point. This argument further strengthens that LLMs are today viewed as an opportunity, since managerial action comes from a single idea, or a future problem, where passivity and time can make an opportunity to later become a problem or a crisis in accordance with Mintzberg et al. (1976).

### 5.1.3. Diagnosis

In the case study by Mintzberg et al. (1976), a clear majority of the solutions were custom-made, which could explain the challenge Delta and Iota face of translating the business need to a solution. We see that consultants can be involved as a formal diagnosis from the companies, which is coherent with Mintzberg et al. (1976), when it comes to handling problems or opportunities which in this study revolves around AI. All interviewees except the junior business transformation consultants, Gamma and Kappa have expressed increased inquiries about LLMs. However, they still give input in various steps of the process, even though their clients' problems are not AI-related, contributing with a more nuanced KIF process perspective. All clients with AI problems or opportunities will presumably be handled by the data science departments or be in contact with people in these departments.

Depending on what the accumulated stimuli consists of, which marks the decision recognition point in the decision making process of Mintzberg et al. (1976), the next step in the process will differ. The strategic decision making involvement from the consultants will have different starting points which was particularly expressed by Epsilon. It also indicates that the transition from the identification phase to the development phase can be challenging according to what Delta expressed.

The decision makers, now seeking help from experts, could possess a large quantity of ambiguous data which contains problems and opportunities. In this routine of Mintzberg et al. (1976), the iterative character of the identification phase is prominent since viewing the stimuli from a cause and effect perspective requires consultants to analyze the stimuli. From this stage, the consultants need to make a judgment on factors from Mintzberg et al. (1976) such as perceived payoff, uncertainty and success probability, cumulated with frequency, pattern, clarity and consistency. Moreover, Eta and Delta mention perceived payoff and uncertainty and all other interviewees express factors such as clarity or uncertainty implicitly. When Iota is to do diagnosis without data, there is undoubtedly uncertainty and low clarity in the data, leading to an iterative process where stimulus needs to be analyzed further. Additionally, both decision complexity and outcome uncertainty is highly prominent, which Ivanov (2023) emphasized is typical for strategic decision making and where the first mentioned explicitly highlights the amount of data needed as a factor that raises decision complexity. Delta's statement about considering potential outcomes is also coherent with Ivanov (2023) regarding outcome uncertainty which in a more straightforward manner involves the predictability of a situation which Delta mentioned.

In this routine, most interviewees have in some way expressed the importance of data whilst emphasizing the combination of business intelligence and AI. Arguably, in order to make a good judgment for complex and uncertain decisions, the values for each stimulus factor should be accurate and of high quality, which is what "good quality data" expressed by Delta, is interpreted as. It should be noted that "good quality data" was emphasized as a premise to create a good plan for the future by Delta, and Iota expressed a similar point. This implies that depending on the stage of the decision process, where the clients involve the consultants, the starting point for the consultants will vary. Additionally, it suggests that interviewees go through an iterative process to get more accurate values for the factors affecting the stimuli, as this is expressed by Delta, Kappa, Epsilon, Iota.

The use of ChatGPT as support for the general purpose of supporting steps in the process by gaining better information, insights, and understanding can be interpreted as showing both decision complexity and outcome uncertainty from Ivanov (2023), since it involves handling large amounts of data with uncertain outcomes. In the decision making process by Mintzberg et al. (1976) these general benefits could be strengthened through ChatGPT in various routines. It can be interpreted as it would have some kind of overarching applicability throughout the process to some extent, perhaps more prominent in decision recognition, but LLMs could be used wherever these general benefits can be harnessed. It should be noted that Epsilon mentioned the two ways LLMs could be used in business in the future in the same way that Delta describes his current use of LLMs.

To summarize the discussion of the diagnosis routine, it is found typical for the consultants to first encounter the problems posed by the clients, both in accordance with Mintzberg et al.

(1976) and the empirical data. The consultants analyze the situation, which most likely is of a problem-opportunity nature when it comes to LLMs, with a combination of human business intelligence and data. Depending on the client's stimuli factors, industry, and matching phenomenon in general, in contrast to the threshold level, it sets the quality of the decision recognition point. This in turn determines whether the process continues to the development phase or if the routines need to be recycled.

## 5.2. Development

The development phase both express commonalities and uniqueness. The uniqueness is detected in the search routine whilst the commonalities lie in the design part. It is worth noting that the commonalities only exist when the design is viewed broadly where a deeper dive into how the design in the particular use cases for the firms would arguably have given diverse answers, since many of the interviewees highlight that the cases are unique. In relation to Mintzberg et al. (1976) it can be stated that this thesis most likely includes responses exemplifying both cases of ambiguity and of uncertainty.

### 5.2.1. Search

It should be noted that only two of the interviewees expressed the generating of solutions which can be interpreted classified as search. Eta lifted several elaborate examples whilst Delta shortly stated that there are obvious cases for the technical solution and pointed to the ones derived from traditional AI. In Eta's development, a memory search can clearly be distinguished in the RPA example, where the fully developed was used in this repetitive process. The theory from Mintzberg et al. (1976) does not exemplify or highlight how small of a modification that can be made to still have it viewed as a fully developed to ready-made solution, which entails that the distinction between search and design should be done carefully, It is consistent with literature from Mintzberg et al. (1976) that in case the design is ready-made, several alternatives can be generated in the design process and then all can be compared. If a small design is done, it has a low cost to bring forward several alternatives. It can be interpreted that in this scenario Eta used his previous experience in RPA. Presumably the resources for bringing forward several alternatives during design was lower than the potential resources being used for not having an optimal model. The resources that Eta mentions is also what Mintzberg lifts as the resources to consider.

Seen through Ivanov’s (2023) lens the decision making which Eta described as repetitive can be interpreted to have a high frequency, leading to a high level of algorithmization and suggestively could be close to an on the loop approach. The repetitive process could additionally be viewed as an incremental task and Kahneman's system 1 (Kahneman, 2011) where LLMs already are performing (Bubeck, 2023).

It should be noted that in addition to Eta, which in his search routine was able to develop several prototypes in the search part of the development phase to later select one, Epsilon’s typical innovation process involves killing of ideas when there is a lack of feasibility. It should however be noted that the screening Epsilon does differ from the screening that Mintzberg, et al. (1976) Describes in their theory. Epsilon’s screening is made for ideas neither mentioned to be ready made or custom made. Most likely they are in the design spectrum since the empirical material shows it as more common, as well as fully developed to ready-made alternatives probably have a higher feasibility rate since they have been used before and take the form of an already made solution rather than an idea. The most interesting point to derive from the comparison between Eta and Epsilon’s screening routines is that Eta’s is in accordance with the literature and Epsilon’s is not. For the feasibility analysis Epsilon means that his experience makes the screening process more efficient. System 1 (Kahneman, 2011) is used rather than testing all ideas before written off, using System 2. Epsilon’s screening routine implies that the screening also could happen in the design routine and necessarily should not be limited to the search routine. The finding is that experience based system 1 could be used for the screening purpose, also for an idea that categorizes as design. Screening should thereby not be limited to solutions but ideas for design as well.

It should be noted that the other interviewees' clients could be argued to be in a search phase from a more passive or trap behavior, where the behaviors from Mintzberg et al.(1976) is described in chapter 2. This goes for all clients that have succeeded on to the development phase but for which there is no ready-made solution and moving to design for some reason is deemed not motivated. These clients are thereby waiting for a ready-made solution, which should appear in the process. The early stage for LLMs together with the empirical material could indicate this where reaching out to consultants could be argued as activating a search generator.

### 5.2.2. Design

The separation of AI implies that specific tools or models are more or less appropriate depending on the case. As LLMs are one of these models it has certain areas it can be used in. Gamma's example of the RPA, Epsilon's questioning of the problem to be of an AI nature or not, and Delta who mention that other AI technologies should not be overlooked exemplifies that there is an emphasis on finding the most appropriate solution rather than forcing LLMs on use cases. It can further be compared to the matching phenomenon mentioned in the identification phase and the division of the solutions as non AI, traditional AI or LLMs where all interviewees in some way have expressed the LLMs to be of a novel technology which can indicate that translating the business need to a technical solution requires customization. Several interviewees, including Iota, have also expressed every case to be unique, some express there to be a basic structure but still highlighted the uniqueness of every case, as Epsilon and Kappa.

Comparing search to design in the development phase, the search examples clearly are of a lighter nature when it comes to resource consumption in accordance with Mintzberg, et al. (1976) As Mintzberg describes that the different development routines may happen simultaneously, it could be that the interviewees apart from Eta have used search in development and that Eta have used design as well. The interviewees could have focused on design or search depending on the one that seemed most convenient to give answers to, meaning that they are all trying both. Logically, all would choose to search if possible and the ones proceeding with design would still have a portion of search as well.

## 5.3. AI Challenges

Delta's statement about the risk and challenges for the adoption of LLMs in strategic decision making processes can be interpreted and related to stimulus and the threshold level (Mintzberg et al. 1976), it could be argued that the accumulated stimulus balances on a tipping point for the threshold level. In some areas, this threshold level seems to be lower, as for the legal services the stimulus exceeds the threshold level leading to an applicability and advancement in the decision making process. In less directly applicable areas, they are still in a decision recognition routine or have just reached the diagnosis routine. In general, Delta means that the technology side is evolving faster than the other parts of society which can be interpreted to create a stagnation or imbalance in stimulus, perhaps a matching problem.

### 5.3.1. Data Privacy

The issue of data privacy is expressed by all respondents as the most significant challenge, which implies that the issue is fundamental for all firms interested in utilizing the technology. More specifically, there are two major divisions of challenges within data privacy, the first one which all interviewees have expressed is the use of client information which cannot be uploaded to a cloud service. The second, which Epsilon and Iota has expressed but is deemed equally important as the first, is the clients themself uploading their data, which may lead to intellectual property issues. This implies that no one can upload company or client data to the LLMs as long as they are cloud platforms, because of trust and legal issues. It can be noted that the tasks, such as summarizing text, that Gamma expressed a desire to perform through ChatGPT are categorized as incremental tasks by (Bubeck et al. 2023), meaning that Gamma would have great use for the LLM.

Epsilon provides an answer about open source alternatives being the best option for the data privacy problem to reach a solution, and argues that currently only companies trusting the cloud service can benefit from the LLMs as ChatGPT.

Data privacy is not only a concern for firms wanting to use LLMs, but also those who are developing it (Future of Life Institute, 2023) and legislators (European Commission, 2021). This implies that several actors see a great interest in solving the data privacy issue and another answer in addition to the open source alternatives is legislation. From the EU, The general goal is to regulate and oversee the development to ensure safety and simulate the business environment around AI (European Commission, 2021).

### 5.3.2. Technological Capabilities and other Challenges

In all, the technical capability is seemingly the one that gets the least negative attention when it comes to the challenges. Only two of the respondents explicitly mentioned that LLMs failed at analytical tasks which depending on the question could be interpreted as discontinuous tasks, from the division made by Microsoft (Bubeck, 2023). Further, all praised its incremental tasks, even though a majority of the interviewees mentioned that LLMs are not perfect and still contain flaws. Additionally, Gamma expressed that other challenges, such as data privacy, could hinder the use of LLMs. Overall, the most important thing when it comes to the technical capability is the quality of the output.

OpenAI (2023) themselves and Wolfram (2023) touch upon the concern of how the quality of data will affect the output. Wolfram's (2023) critique is in regard to GPT-3, but it still seems relevant as OpenAI (2023) states that flaws from the GPT-3 models still exist even though they have been mitigated to some extent. Since they also recommend a human review, grounding with additional context or entirely avoiding high stake contexts, the posed concerns by Kappa and Gamma are sound and valid judged by the OpenAI's (2023) words about their LLMs capabilities.

Comparing the opinions of the interviewees and the capabilities both Microsoft and OpenAI (Bubeck; OpenAI, 2023) They are in general coherent, for instance when it comes to handling text in an incremental way by summarizing it, both the Microsoft report and several of the interviewees express coherence. The healthcare sector is also mentioned as a highly applicable area for LLMs (OpenAI, 2023). Again this showcases that depending on the industry and the problem LLMs could be more or less applicable. However, even if applicable, the risks remain which have to be taken into consideration. The sector thereby has characteristics from Ivanov (2023) that both speak in favor for AA and for human autonomy respectively. Some characteristics that point in favor of AA is the context which can be argued to be well defined with clear rules and formal procedures where some have a high frequency, potentially leading to a high level of algorithmization. On the other hand the context, even though well defined, might vary a lot and in the cases where all of the above mentioned is true, but with the addition of complexity, outcome uncertainty and negative consequences. This can all be applied to the example Delta gives about phone helplines where the questions and issues from the patient determine the appropriate approach. Suggestively, a natural path forward is the middle ground approach of human in the loop which also might be more or less towards a higher or lower level (Kaber and Endsley, 1997) depending on the case.

A wide array of challenges is emphasized by Delta as he described that there is a need for non-tech aspects to catch up with technological advancements in LLMs. The importance of regulations, risk management, and ethical frameworks is highlighted, particularly in critical industries. Challenges such as knowledge, time, effort, skills, and cultural differences are identified as obstacles to broader AI implementation.

# 6. Conclusion

As the typical firm arguably aims to make sound decisions as frequently as possible to gain competitive advantages, making sound strategic decisions can be imperative. The increased accessibility of Large Language Models has presented an opportunity for firms to potentially utilize the technology in various ways. This led to the purpose of this study, which was to contribute with an understanding of what role AI has, and might have, in strategic decision making processes. To fulfill this purpose, the following two research questions were posed:

> *How can Large Language Models be utilized in strategic decision making processes?*
> *What challenges of using Large Language Models are perceived by professionals in knowledge-intensive firms?*

To achieve this, a qualitative single case study was conducted by carrying out semi-structured interviews with professionals in knowledge-intensive firms. The theoretical framework was used to cover strategic decision making and AI. The empirical results showed that utilizing LLMs in strategic decision making processes is limited due to it being in an initial phase and encountered with several challenges which hinders a wide application of LLMs in firms. More specifically, the fundamental challenge of data privacy was found to be aggravating as the empirical material strongly suggests the use of client information cannot be uploaded to a cloud service. Furthermore there are challenges such as technological capability and ethical challenges.

The empirical material indicates that there is an increased demand in seeking consulting for strategic solutions supported by LLMs, either in internal processes or as a product. We conclude that this demand is due to the LLM ChatGPT becoming publicly accessible in 2022, thus being viewed as a stimuli as defined by Mintzberg et al. (1976) in an opportunity seeking behavior rather than for crisis solving. However, it is noted that even though a high number of firms are seeking solutions or opportunities, only a small number constitutes successful use cases, where legal and RPA are examples. We suggest, from the empirical findings, that this is due to a matching problem of problems, opportunities and crisis, which can be grounded on Mintzberg et al. (1976) theory. This means that clients do not experience enough stimuli to take action for a solution in the first place, therefore enforced an iterative process to find other stimuli or properly analyze stimuli. It could also mean that they do have motives, but get stuck in a search phase awaiting the matching solution. Epsilon's screening routine in his innovation approach implies that the screening also could happen for ideas in the design routine and necessarily should not be limited to the solutions in the search routine. The finding is that experience based System 1 from Kahneman (2011) could be used for the screening purpose, also for an idea that categorizes as design. This further indicates that a decision making process in accordance with Mintzberg et al. (1976), might resemble a traditional innovation approach.

Our empirical data demonstrates a higher utilization of AI/LLMs in the beginning of the process which partly is due to that the empirical material being geared towards the beginning of the process, since most LLM clients are in the beginning of the process, and partly is due to the information heavy character of the identification part in the beginning. It may also be due to the novelty of the technology. Since strategic decision making projects can span over several years it might be that it is simply too early time wise, despite challenges mentioned in the beginning. The general sound conclusion is that since the introduction of LLMs is in the beginning and there are several challenges, and LLMs in strategy encompass a highly complex and uncertain problem, everyone is still in an evaluation mode, waiting for a holistic picture to take form.

This thesis has shed initial light on a novel topic by contributing with an understanding of how advanced technology in the form of LLMs is affecting strategic decision making processes. LLMs and strategic decision making as an interdisciplinary area constitutes a highly unique research area and the research might thereby create an initial path for more research to follow the path of this study.

## 6.1. Limitations and Future Research

A methodological limitation of this study is the choice of a single case study conducted on a phenomena. A comparative study could potentially delve deeper into what differences and similarities there are between different types of organizations that are integrating AI in their decision making processes. Additionally, as the empirical findings build upon individuals' perceptions and descriptions, the study's generalizability is limited. Another limitation of the study is the experienced difficulty of finding appropriate interview participants as the topic of AI and LLMs limited the scope of possible candidates who were able to discuss the topics of the thesis. However, it should be noted that the participants who were found suitable held expertise and thus contributed empirically in a rich manner. An additional limit of the study is analyzing and discussing the later parts of the strategic decision making process. This is because the further into the decision making process the analysis and discussions go in this study, the more difficult it becomes to analyze the answers and apply the theory as the empirical data is richer for the earlier steps in the process due to how LLMs are being used.

As the LLMs recently became accessible to a wider audience of users, and the empirical results indicate that there are many challenges left to tackle before a wide implementation of LLMs to support decision making processes, there are still questions left unanswered. One question for future research could be what impacts different business models of LLM-developers could have on the identified challenges such as data privacy and ethics, as the empirical findings showed that open source alternatives could influence a firm's caution of data privacy issues. Additionally, future research could also study how top management of firms perceive how the novel LLMs can be used in strategic decision making processes as they are operating in the selection phase of Mintzberg et al. (1976), meaning that they make the final call. This could give new insights on what the interplay between human and AI is like in the top level of decision making in firms, rather than the strategic decision making process which was the scope of this thesis.

# References

Alvesson, M. (2004). Knowledge Work and Knowledge-Intensive Firms, [e-book], Oxford: OUP Oxford, Available through: LUSEM University Library website http://www.lusem.lu.se/library [Accessed: 6 May 2023]

Barney, J.B. (1991). Firm Resources and Sustained Competitive Advantage, *Journal of Management*, vol. 17, no. 1, pp. 99-120

Bell, E., Bryman, A. & Harley, B. (2022). Business research methods, Oxford: Oxford University Press

Berruti, F., Nel, P. & Whiteman, R. (2020). An executive primer on Artificial General Intelligence, McKinsey & Company, Available online: https://www.mckinsey.com/capabilities/operations/our-insights/an-executive-primer-on-artificial-general-intelligence [Accessed: 24 April 2023]

Bryman, A. & Bell, E. (2011). Business research methods, Oxford: Oxford University Press

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T. & Zhang, Y. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4, Available online: https://arxiv.org/pdf/2303.12712.pdf [Accessed: 22 April 2023]

Chandler, D. (2021). Sustainable value creation, [e-book], London: Routledge, Available through: LUSEM University Library website http://www.lusem.lu.se/library [Accessed: 10 May 2023]

Creswell, J.W. & Creswell, D.J. (2018). Research design. qualitative, quantitative, and mixed methods approaches, Los Angeles: SAGE

Eisenhardt, K.M. & Martin, J.A. (2000). Dynamic capabilities: what are they?, *Strategic Management Journal*, vol. 21, no. 10-11, pp.1105–1121

Eisenhardt, K.M. & Zbaracki, M.J. (1992). Strategic decision making, S*trategic Management Journal*, vol. 13, pp.17–37

Endsley, M.R. & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task, *Ergonomics*, vol. 42, no. 3, pp. 462-492

European Commission. (2021). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206 [Accessed: 20 May 2023]

European Parliament. (2023). AI Act: A step closer to the first rules on Artificial Intelligence, Available online: https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence [Accessed: 20 May 2023]

Future of Life Institute. (2023). Pause Giant AI Experiments: An Open Letter, 22 March, Available online: https://futureoflife.org/open-letter/pause-giant-ai-experiments/ [Accessed: 24 April 2023]

Google. (2023). PaLM 2 Technical Report, Available online: https://ai.google/static/documents/palm2techreport.pdf [Accessed: 18 May 2023]

Google Trends. (2023). Google Trend chatgpt, Available online: https://trends.google.com/trends/explore?date=2022-04-24%202023-04-24&q=chatgpt&hl=en-US [Accessed: 24 April 2023]

Hambrick, D.C. & Mason, P.A. (1984). Upper Echelons: The organization as a reflection of its top managers, *Academy of Management Review*, vol. 9, no. 2, pp. 193–206, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed: 17 April 2023]

Heracleous, L.T. (1994). Rational decision making: Myth or reality?, Management Development Review, vol. 7, no. 4, pp. 16–23, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed: 23 April 2023]

Ivanov, S.H. (2023). Automated decision-making, Foresight, vol. 25, no. 1, pp. 4–19, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed: 16 April 2023]

Johansson, O. (2023). Lunds universitet inför AI-robot i utbildningen: "Fantastisk hjälp för studenterna", SVT Nyheter, Available online: https://www.svt.se/nyheter/lokalt/skane/chattbot-integreras-i-utbildning-i-lund-helt-nodvandigt [Accessed: 17 April 2023]

Joshi, N. (2019). 7 Types of Artificial Intelligence, Forbes. Available online: https://www.forbes.com/sites/cognitiveworld/2019/06/19/7-types-of-artificial-intelligence/ [Accessed: 24 April 2023]

Kaber, D.B. & Endsley, M.R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety, *Process Safety Progress*, vol. 16, no. 3, pp. 126–131

Kahneman, D. (2011). Thinking, fast and slow, London: Allen Lane

Kumar, A. (2023). Large language models: Concepts & Examples, Data Analytics, Available online: https://vitalflux.com/large-language-models-concepts-examples/ [Accessed: 18 May 2023]

Mintzberg, H., Raisinghani, D. & Théorêt, A. (1976). The Structure of “Unstructured” Decision Processes, *Administrative Science Quarterly*, vol. 21, no. 2, pp. 246–275

OpenAI. (2022). Introducing ChatGPT, Available online: https://openai.com/blog/chatgpt#OpenAI [Accessed: 23 April 2023]

OpenAI. (2023). GPT-4 Technical Report, Available online: https://cdn.openai.com/papers/gpt-4.pdf [Accessed: 20 April 2023]

Rennstam, J. & Wästerfors, D. (2018) Analyze! Crafting your data in qualitative research, Translated by R. Ehnsiö, Lund: Studentlitteratur

Shao, Z., Zhao, R., Yuan, S., Ding, M. & Wang, Y. (2022). Tracing the evolution of AI in the past decade and forecasting the emerging trends, *Expert Systems with Applications*, [e-journal] vol. 209, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed: 15 April 2023]

Sheridan, T. B., & Verplank, W. I. (1978). Human and Computer Control of Undersea Teleoperators, MIT Man-Machine Laboratory, Cambridge, MA, technical report, Available online: https://apps.dtic.mil/sti/citations/ADA057655 [Accessed: 19 April 2023]

Taskan, B., Junça-Silva, A. & Caetano, A. (2022). Clarifying the conceptual map of VUCA: A systematic review, *International Journal of Organizational Analysis*, vol. 30, no. 7, pp. 196–217, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed: 18 April 2023]

The Alan Turing Institute. (2019). What Alan Turing means to us, Available online: https://www.turing.ac.uk/blog/what-alan-turing-means-us [Accessed: 24 April 2023]

Wolfram, S. (2023). What is chatgpt doing ... and why does it work?, Stephen Wolfram Writings, Available online: https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ [Accessed: 19 May 2023]

# Appendices

## Appendix A - Interview guide

**Intro**

This interview will be anonymous, your name and the company name will not be mentioned anywhere in this study.

1. Do we have your permission to record the interview?
2. Can you please describe your role in the company?

**Decision Making**

3. What type of strategic decisions processes are you involved with typically?
    - What characterizes the different decisions, are they different? Please exemplify.
4. What does the process for decision making look like in those different decisions?
    - Would you say that there are different phases in this process?
5. Do you experience that you have to rely on your gut feeling or intuition when having to make a decision, if so how often and in what situations?

**AI in Decision making - Large Language Models**

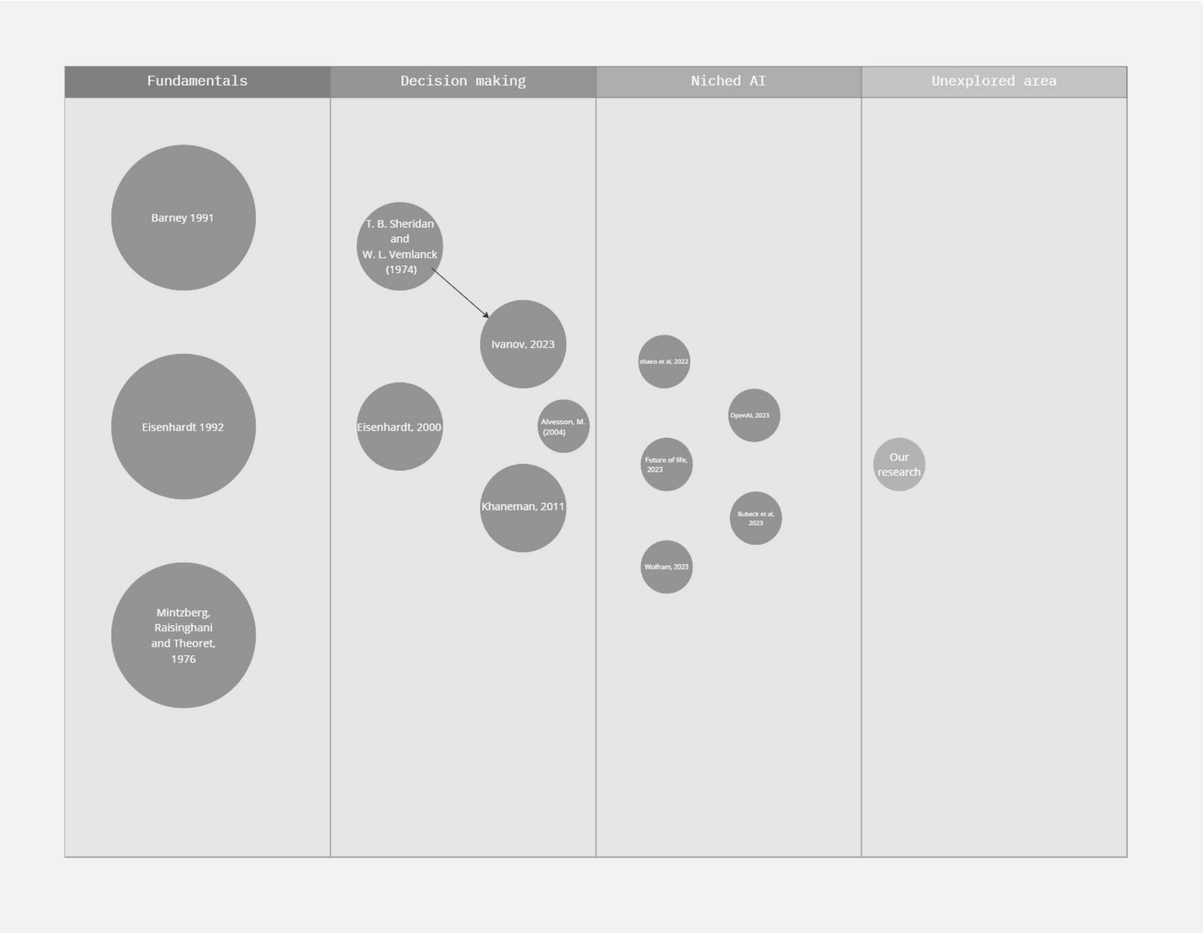
6. Do you work with AI to support your decision making in your job today?
    - If not, are there any particular reasons for why not?
    - Would you prefer to work with AI to a greater extent than today?
7. In what ways do you think AI can be a resource for your company?
8. Have you noticed any changes in demands from clients regarding the usage of AI?
    - Can you give an example of such questions they bring you?
9. Hypothetically speaking, how could AI be used in the different types of decisions you talked about earlier? Please Exemplify
10. What main flaws do you see in using AI in decision making?
11. What advantages do you see in using AI in decision making?

**Outro**

12. Is there anything you would like to add? Is there anything you think we have missed?
13. Do you know of anyone that would be interesting for us to interview?

Thank you for participating!

# Appendix B - Visual Map of Literature collection




Visualization of literature map, illustrating the articles from broad theoretical concepts to niche research on AI, from left to right.

# Appendix C - Research Proposal

## Proposal for participating in study - The role of AI in strategic decision making in consulting firms

There is a call for firms to make great use of the magnitudes of data available to them in order to make the right decisions, and sustain competitive advantages. Firms that are arguably reliant on making sound decisions are knowledge-intensive firms, such as consulting firms, and by extension their clients, as their service is to provide their clients with knowledge and solutions. AI is outstanding when it comes to structuring and codifying data, and is starting to show some tendencies to comprehend issues requiring human-intelligences. This leads to the question of how AI can be utilized by knowledge-intensive firms to provide decision making support for clients as well as knowledge-intensive firms' view on providing solutions that include AI and/or LLMs. As AI is constantly showing new capabilities, its future potential is to count on. Other areas in society, such as education to name one, are currently faced with the question of how to adapt to the new reality with AI.

*Purpose:*

The purpose of the study is to contribute with an understanding of what role AI has and might have in decision making. In order to contribute to this area of decision making literature, the following central research question will be answered:

Central Research Question:

- How is AI utilized by knowledge-intensive firms to provide decision making support for clients and how do knowledge-intensive firms view the providing of solutions that include AI and in particular LLMs?

Examples of Sub Questions:

- How do knowledge intensive firms view AI as a resource for the clients?
- In what areas of decision making is AI currently used?
- In what areas of decision making is AI planned to be used?
- How do AI and human intuition interplay in strategic decision making?

The study will use a qualitative case study approach and we will conduct semi-structured interviews. The interviews will focus on the usage/potential usage of AI in decision making and its impact on strategic decision making. The interviews will be recorded for transcribing purposes after permission is given, in order to simplify the analysis procedure. The company and the participants will be anonymous. The study is expected to provide insights into the current and planned usage of AI in knowledge-intensive firms externally to clients, the interplay between AI and human decision making, and the impact of AI on strategic decision making. The findings will contribute to a better understanding of the role of AI in decision making and provide implications for both consulting firms and their clients.