Dipartimento di IMPRESA E MANAGEMENT

Cattedra: INTERNATIONAL MANAGEMENT

ARTIFICIAL INTELLIGENCE AS AN INNOVATIVE DECISION-MAKING TOOL

RELATORE

Prof. LUIGI MARENGO

CANDIDATO: GIULIA BATTISTA

MATR. 710551

CORRELATORE

Prof. ANDREA FILIPPETTI

ANNO ACCADEMICO 2019/2020

Alla mia famiglia che non solo mi sopporta da 23 anni, ma che mi ha anche sopportato in lockdown mentre scrivevo la tesi (non credo serva aggiungere altro).

Table of Contents

Introduction

Effective managerial decision-making is at the core of every company's success. However, in the last decades, the economic environment is becoming decreasingly stable, thus making the decision-making process harder to perform while the taken decisions are bearing significantly more weight than they did in the past.

Managers and executives make important decisions daily and if the process is wrongly executed it can cause a considerable amount of economic loss for the companies and its shareholders. Since the process needs to be as seamless as possible but it becoming increasingly complicated this thesis proposes and analyzes Artificial Intelligence (AI) tools in order to simplify, speed-up the process and make it more precise with a particular focus on the decision-making process during the recruitment process.

The first three chapters of the thesis aim at providing a theoretical background on the decision-making process and on the obstacles that can arise because of human agents. As part of the human nature, individuals make decision based on their previous experience and can be subjected to bias, which prevents them from making optimal decisions. Since the thesis proposes AI as a tool for correcting human flaws, the first part of the thesis also provides background theory on what is really meant with the term AI and on the steps that organizations need to take to effectively integrate AI tools within their structure and the impact that the adoption of this technology will have on the labor market and on prediction tasks. Moreover, the impacts that AI can have on human behavior and on the economy are discussed. In addition, a theoretical background on why algorithms can be better predictors of decision-making is provided, along with the potential pitfalls that this technology can have.

The fourth and final chapter is the application of the theory illustrated in the previous chapters. In order to acknowledge the practical implications of the technology, a study is conducted to understand the current state of the AI technology and whether or not it is a valuable tool to adopt within the recruitment process. Six professionals of the recruitment sector have been interviewed, both in favor and opposing to AI, in order to gain a full spectrum of perspectives. The main aim of the study is to understand whether AI can be a valid tool for reducing human flaws and biases within the process. In order to assess this, three main themes, namely Limitations and opportunities to overcome them; Evolution of the recruiter's role; and Bias reduction, which are consistent with the theory presented in the previous chapters, have been analyzed. Moreover, the study was conducted with the following research objectives in mind:

1. Provide insight on how companies have overcome the concerns and successfully implemented AI (Theme 1)
2. Observe which concerns are raised by those companies that do not use AI within their recruitment software (Theme 1)
3. Investigate whether AI changes the recruiter's role once adopted (Theme 2)
4. Investigate if AI can be a valid tool for bias reduction (Theme 3)

Chapter I

Biases and Limited Rationality in Human Behavior

1.1. Behavioral Economics Background

Behavioral economics has been developed in response to the standard model of human behavior which views an individual as a Homo Economicus (Mullainathan & Thaler, 2000), suggesting that an individual is an unemotional and rational maximizer. Behavioral economics has challenged and later enriched this framework by analyzing and correcting the unrealistic human traits around which the standard model focuses. The traits are respectively: unbounded rationality, unbounded willpower and unbounded selfishness.

Herbert Simon (1955) was one of the earliest critiques of the standard model and introduced the term "*bounded rationality*" to more realistically describe human capabilities. By deepening his studies on the concept of bounded rationality, he observed that rationality is threatened by numerous factors such as psychological and physiological limitations that are taken as givens in optimization problems; external and internal constraints that define the problem of optimization; and ultimately by limitations of computational capacity. The importance of substituting the standard model's concept of unbounded rationality with that of bounded rationality is also emphasized by Conlisk (1996) who argues that not acknowledging bounded rationality in economic models would lead to ignore deliberation costs and thus give rise to a "*free lunch fallacy*" (Conlisk, 1996, p.686). In addition, since there are time constraints for decision-making, individuals tend to adopt rules of thumb to economize on both time and effort. Although people adopt these rules of thumbs, the standard model fails to recognize them, thus leading to relevant systematic errors (Tversky & Kahneman, 1974), again stressing the importance to account for bounded rationality instead of unbounded rationality.

The second unrealistic trait on which the standard model of human behavior relies is unbounded willpower. According to this trait the Homo Economicus should solve for the optimum and then choose it. However, in practice this does not happen as individuals fail to choose what would maximize their long-term profit because of their lack of self-control. To understand this concept, it is sufficient to simply think about smokers who acknowledge that it would be better for them not to smoke, but fail to quit even though they know it would be more beneficial for them in the long-term

(Mullainathan & Thaler, 2000). Once more the standard model proves to be weak when applied to real-life situations. Hence behavioral economic is fundamental in modifying the standard model by switching the concept of unbounded willpower with that of bounded willpower.

Lastly, according to the standard model, individuals are boundedly selfish and do not exhibit unbounded selfishness. According to economic theory, individuals act mainly because of self-interest as can be observed in an often-used example: the free rider problem. According to economic theory this problem is likely to arise because individuals are not expected to contribute to public goods unless there is a return for their own private welfare. However, people often take selfless actions and this is observed in games such as public goods games and prisoner dilemmas in which players cooperate, and in ultimatum games in which players turn down unfair offers.

Therefore, as illustrated above, behavioral economics is the result of the modification of the standard model of human behavior as the behavioral model takes in consideration the effects of psychological, cognitive and emotional factors in order to give a more comprehensive and complete overview of human behavior in economic environments.

1.2. Managerial Decision-Making Process

Decision making is an integral part of today's administration as decisions are a core function of management. Decisions are fundamental for a firm's growth as they are the foundation for organizational and managerial activities and, as a consequence, determine the organization's success or failure. Although successful decision-making is vital for an organization, there is no clear definition for decision. According to Simon (1960) decision-making is made up of three phases, respectively: finding the occasion; finding potential courses of action; and lastly selecting among the various courses of action. In contrast, Harrison's (1996) definition does not refer to phases and outlines decision as the moment in which expectations about a precise course of action lead the decision maker in selecting a specific path in order to achieve his objective. So, the question still holds, what is exactly meant with the term decision-making?

According to Harrison (1996) the decision-making process is a combination of the strategic gap and the managerial decision-making process. The concept of strategic gap arises since decisions create a relationship between an organization and the external environment in which it operates. The term strategic gap underlines the difference between the current strategic position and the organization desired position and is determined by comparing the firm's capabilities with the opportunities and threats found in the external environment. To measure the inherent capabilities, an organizational assessment is performed to evaluate whether an organization is able to capitalize on external opportunities, which implies significant internal strengths; to measure an organization's ability to protect itself from external threats, thus indicating an adequate knowledge of the internal weaknesses; and to evaluate if an organization succeeds in both areas as not many are able to. Once performed, the organizational assessment will reveal the effectiveness of the firm's management, technological level, policies and resources. The environmental opportunities and threats instead are measured through an environmental assessment which analyzes the environment's opportunities, threats, requirements and responsibilities. Once both assessments are concluded an organization will either have a positive (internal capabilities are greater than the environmental aggregates), negative (environmental aggregates are greater than the internal capabilities), or a zero strategic gap (the difference between capabilities and aggregates is caused by imperfect information).

The second component of decision-making is the decision-making process itself. Within this process managerial decision-making results from a set of decision-making functions that when grouped generate decision-making (Witte, 1972). The process is composed of three flows which all contribute to generate the final outcome of the process. The first flow is the primary flow which encompasses the main functions of the strategic decision-making process. Within this flow information gathered from the external environment is used to assess the strengths and weakness of the organization together with the threats and opportunities present in the environment. The second flow is the corollary flow which incorporates the additional functions of the process, such as the assessment of implemented decisions. Bypassing this flow would jeopardize the overall process as the combination of the corollary and the primary flow generates a guideline for successful strategic decisions. The final flow is the information flow. This flow consists of exploring all of the possibilities in order to find a set of alternatives and also deals with receiving the feedback on the information generated from the external environment in order to assess the success or failure of the implemented strategic decision.

Since the decision-making process results from the combination of the gap analysis and the process itself, it follows certain dynamics that need to be balanced

(Harrison, 1996). The decision-making process relies on three fundamental relationships: the external environment's influence on the decision-making process; the relationship between the gap analysis and managerial decision-making in order to ensure that the managerial objectives are coherent with the gap analysis; and finally, it is dependent on the continuous flow of information throughout the process, starting from the gap analysis until the evaluative flow of information generated by the external environment enabling to take corrective action.

1.2.1 Decision Making in Uncertain Economic Environment

Today managers worldwide have to deal with consistent amounts of uncertainty, which require fast and precise decisions to be taken by the management community. In uncertain environments the decision-making process becomes more complex as there is little time available to collect and evaluate information, which might be accurate, inaccurate, incomplete or simply not available, thus making decision-making an increasingly difficult task to perform. Furthermore, the decisions must be made taking in consideration the context in which the organization operates as it is important to account for environmental constraints and multiple players when evaluating the possible consequences that might be stem from the decision.

In order to succeed in a situation of environmental turbulence, a rational decision-making process is preferred as it would enable managers to act logically after having successfully processed the available information. Being able to act rationally is a fundamental quality for executives as this skill prevents them from distorting the reality under stressful conditions. If irrational, managers would otherwise rely on their believes when generating decisions. Although in theory rationality is easily accepted, in practice it is not that straightforward to always act in a rational way. In fact, managers, to face uncertainty, tend to adopt cognitive schemas (Harris, 1994). These schemas define the cognitive structure used to organize individual knowledge and are derived by personal experiences which are conceptualized as subjective theories that are then used to make decisions. Since they are developed form personal experiences and perceptions, the interpretation, prediction and understanding of events in an environment will be different between people, thus generating different decision-making schemas depending on the person.

Together with cognitive schemes, self-efficacy plays an important role in decision making. Self-efficacy indicates a measurement of the manager's degree of self-confidence that allows him to believe that he is able to handle a situation of uncertainty (Sayegh, Anthony & Perrewé, 2004). Self-efficacy is an important element in decision-making as it impacts the judgment process by which managers interpret outcomes and environmental variables (Bandura, 2012). Past research has shown that individuals who report a high sense of self-efficacy cope better in uncertain situations as they are more able to accept negative feedback (Nease, Mudgett & Quinones, 1999) and are increasingly prone to persist at performing their tasks (Lent, Brown & Larkin, 1987). Lastly, it is important to account for self-efficacy in decision-making as, according to Sayegh, Anthony and Perrewé (2004), it affects an individual's ability to regulate his emotions. Therefore, a strong sense of self-efficacy will result in a more effective decision-making process in uncertainty.

1.2.2. The Role of Experience and Stress in Decision-Making

As previously mentioned, even though rationality is often expected from managers, executives depend on judgmental responses which are generated from previous knowledge in decision-making situations. As experience is subjective, and judgment is based on it, it results that experts often arrive at the solution more rapidly and intuitively than novices, without being able to report the processes that leads them to that result (Simon, 1987). Moreover, the manager's ability to problem-solve springs from a retrieval process that employs a multitude of patterns which are stored in the executive's long-term memory (*ibid*). Since managers follow an intuitive, and not a standardized process of decision-making, the approach will be very different from manager to manager, thus resulting in numerous different outcomes. In addition, the process also varies under uncertainty as the consequences of a decision may be positive in certain environments, but may be detrimental in others. As the bad consequences of decision-making are passed on to the organization and to the employees, managers need to face stressful situations and need to deal with their mistakes in order to solve the problem. Research (Smith, Passos & Isaacs, 2010) has shown that managers cope with stress in either an adaptive or a maladaptive way. Adaptive strategies entail active coping, acceptance and looking for remedies, while maladaptive strategies include self-blame and self-destruction, thus being less desirable within an organization but unfortunately more frequent.

The common underlying of decision-making under stress is that stress is a powerful force that diverts behavior as it is frequently influenced by emotions. Stress generates nonproductive responses, especially under time pressure, as the need to diminish feelings of anxiety and embarrassment leads managers to take decisions that will benefit personal comfort at the expense of the organization's long-run profit. Lastly, the intuition of emotion-driven managers is unlike the intuition of the expert manager (Simon,1987). The expert's behavior is often generated from experience and learning and is adaptive to the environment, while the emotion-driven manger's behavior results from responding to primitive urges, thus making it inappropriate for the organization's survival and growth (*ibid*).

1.2.3. The Cost of Stress

According to Stanton (2006), stress levels can negatively impact the health and well-being of the individual and consequently his ability to make decision. Decision-making becomes impaired as stressed employees start to suffer from psychological symptoms such as low self-esteem, reduced motivation, job dissatisfaction and decreased organizational commitment (Smith, Passos & Isaacs, 2010). In addition, these symptoms influence people as they negatively affect the individual when trying to develop new skills, abilities and knowledge (Mikkelsen, Ogaard, Lindoe & Olsen, 2002), endangering the organization as a whole. The organization can also be damaged by stressed individuals as they tend to exhibit an antisocial behavior which can lead to aggressive conduct and might also cause the individual to perform criminal acts against the organization (Lambert, Lambert & Yamese, 2003), thus setting his decision-making process in a counterproductive manner.

Other organizational costs related to stress include poor employee performance and productivity, along with increased turnover and failure to develop organizational initiatives (Clarke & Cooper, 2004; Leka, Griffitths & Cox, 2003). Employee turnover is of great importance for organizations as it can impact restructuring and growth efforts, ultimately deteriorating the competitive advantage and consequently the organization's survival (LeRouge, Nelson & Blanton, 2006). If stress is not prevented, its costs can be detrimental for an organization as the costs of stress have been estimated to be at €20 billion in EU-15 (Brun & Milczarek, 2007).

1.3. Emergence of Biases Within Organizations

As previously discussed, human decision patterns can differ from those predicted by standard economic theory because of a multitude of factors including judgment and emotion. Moreover, different scholars (Thaler, 2000; Hogarth, 1987) have argued that individuals fail when it comes to judging probabilities as, in the words of Simon (1957, p.198), *"the capacity of the human mind* [...] *is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world"*. As individuals are not able to always act rationally in decision-making, the human decision-making process involves biases, which are systematic deviations from the standard assumptions of the rational paradigm in economics (Carter, Kaufmann & Michel, 2007).

Limitations in information gathering and processing do not allow individuals to analyze all the available solutions in an uncertain environment, thus forcing decision makers to use heuristics to simplify their decision-making process (Tversky & Kahneman, 1974). Although heuristics can be applied intentionally, they can also lead to baseless deviations from rationality (Tversky & Kahneman, 1986). These deviations thus lead the individual to adopt alternatives he considers satisficing rather than looking for, and later accepting, an optimal solution. At times, the deviations also lead to over-optimization or under-optimization of event probability evaluations, thus generating biases. The fields of economics, psychology and organizational decision-making have long studied biases and a review of the literature performed by Carter, Kaufmann and Michel (2007) revealed a total of 76 different decision biases or sources of decision biases divided in nine categories, namely: availability cognition; base rate; commitment; confirmatory; control illusion; output evaluation; persistence; presentation and reference point biases.

1.3.1 Major Managerial Biases and Organizational Consequences

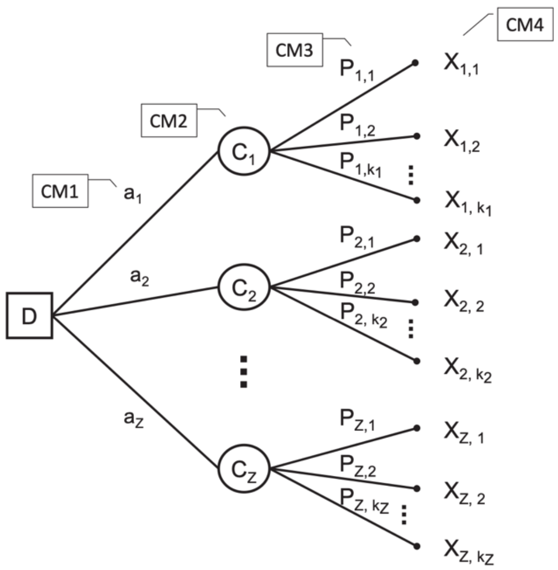
Among the many bias that can occur, overconfidence bias, anchoring bias, halo effect and confirmatory biases are the ones that hinder the decision-making process the most. The overconfidence bias can be found among the control illusion biases (Carter, Kaufmann & Michel, 2007), which occur when random events and non-representative samples are mistaken for fundamental elements of a project thus leading to unrealistic confidence in decision-making (Hogarth 1987). When individuals are subject to control illusion biases, they tend to generate an overly optimistic sense of control when evaluating multi-stage projects as they believe that all the past events are linked to the success of the project (Dawes & Hastie, 2001). The overconfidence bias is one of the major causes of the failure of joint projects developments, between buyers and suppliers, because it hinders the meeting of deadlines and leads to frequently exceed the planned budgets. Moreover, the overconfidence bias is also generated by the individual's poor perception of randomness (Ayton, Hunt & Wright, 1991). Because of this deficiency, individuals mistake non-representative samples and a large amount of data as essential characteristics of a project, therefore decision-makers develop an unjustified feeling of control which in turns leads to unrealistic confidence in judgment (Ayton & Fischer, 2004).

The second most relevant bias with respect to decision-making is the anchoring bias, which falls in the category of the reference point biases. These types of biases occur when the decision makers' judgments are biased in the direction of a reference point (Carter, Kaufmann & Michel, 2007). When it comes to judgement, humans tend to begin from an initial point and then adjust for opinions and evaluations (Tversky & Kahneman, 1974). The problem with this process of judgement is that the amount of adjustments from the initial position are insufficient (Slovic, Fischhoff, Lichtenstein, 1977), hence making the reference point the dominant factor in judgment, causing the presence of biases. Furthermore, research by Epley and Gilovich (2005), has shown that even when an anchor is determined randomly and the decision-makers are aware of its arbitrary nature, they still cannot escape the reference point bias. When it comes to practical managerial decision this bias can have a large impact. An example of the consequence of this bias is a buyer demanding only incremental improvements of price to a supplier because the current price is relying on the buyer's price anchor. In reality, the price set by the supplier might already be too high, but the buyer is not realizing it thus causing a waste of money for the organization.

The last major managerial decision-making biases, confirmatory bias and halo effect, both fall into the confirmatory bias category. When subject to confirmatory biases, individuals tend to search for evidence that supports their judgment while failing to acknowledge disconfirming information for desired outcomes (Carter, Kaufmann & Michel, 2007). The confirmatory bias is very relevant as it is completely opposite to one of the fundamental tenets of the scientific method which states that information against a thesis should be regarded as more valuable than information which supports that thesis. The inability to accept disconfirming information leads to unjustified confidence in behavior (Russo, Medvec & Meloy 1996). In addition, from the confirmatory bias stems the halo effect. The halo effect occurs when a decision-maker relies on his personal beliefs rather than on probabilities (Lynn & Williams, 1990). The individual's impression of a person or product will then influence the decision-maker's thoughts about the entity's overall character and properties (Lachman & Bass, 1985). For example, an individual will believe that a person will automatically perform well in all areas after having successfully done well in a certain area regardless of whether those areas are related or not. Because of this way of forming judgments, managers will regard sources of confirming information as more desirable and reliable than sources of disconfirming information (Gordon, Franklin & Beck, 2005).

1.3.2. The Impact of Biases on Decision-Making Under Uncertainty

Under uncertainty, individuals tend to use decision trees (Raiffa, 1968) as depicted in Figure 1.



Source: Montibeller & von Winterdeldt, 2015

The first step an individual needs to take to generate a decision tree is to identify the decision-making alternatives (Montibeller & von Winterdeldt, 2015) represented by CM1. The second element of the decision tree, CM2, is the identification of event nodes and their future outcomes. Next, is the assessment of the probability of the event nodes, CM3, and lastly is the estimation of consequences of alternative features, CM4.

When identifying alternatives, individuals start by identifying a set of decision alternatives A = {a1, a2, . . . , aZ), which start at decision note D. This step is crucial for the decision-making process as identifying good alternatives is of fundamental importance (Keeney & von Winterfeldt, 2007). Although evaluating alternatives is decisive, individuals often tend to consider only one alternative (Nutt, 1998), thus defining the problem as a binary choice of few options which are not the optimal ones. The inability of an individual to correctly evaluate alternatives might be generated by reference point biases, such as the anchoring bias. This inability suggests that an individual's decision-making process under uncertainty might not be accurate because of biases.

Biases are also present at the second step, the identification of event nodes and their relative outcomes. In the above tree the node C_r represents a chance node, with r = 1,2,…Z, which underlines the uncertainty associated with the set of mutually exclusive and collectively exhaustive events. Biases are relevant also at this stage as the set of outcomes is subject to the overconfidence and confirmatory biases with respect to the exhaustiveness and the range of the outcomes.

The overconfidence and anchoring biases are also relevant with respect to the third stage of the decision-making process: the assessment of probabilities of the event node. At this step conditional probabilities $P_{r,1}$ (with 1 = 1, 2, … , kr) are assessed for each of the rth events defined at the chance nodes. The different methods used in order to excerpt the probabilities may cause biases to arise, especially the anchoring and overconfidence biases (von Winterfeldt & Edwards, 1993). These biases need to be taken in consideration as they might influence and distort the probability estimates, thus generating a non-accurate decision-making process.

Even the last stage, the estimation of consequences of alternative attributes, is not exempt from the biases' influence. At this stage the consequences $X_{r,1}$ (with 1 = 1, 2,.., kr) of executing the lth event are assessed through data collection, modelling of systems and through expert judgment. Because judgment is rarely free of noise, this step will be subject to overconfidence and anchoring biases, once again proving that human decision-making in uncertainty is unreliable because of the impact of biases (Montibeller & von Winterdeldt, 2015) and that alternative methods of decision-making should be adopted.

1.4. There Is No Judgment Without Noise

Besides the problem of biases, which arises because of inconsistent judgment and decision-making by individuals, organizations also need to face the problem of noise. The main difference between biases and noise is clearly explained through an example provided by Kahneman, Rosenfield, Gandhi and Blaser (2016) in which they use a scale as an element for comparison. A scale is considered to be biased if the readings are either to high too low, meanwhile a scale is noisy if the weight depends on where you place your feet. In addition, noise can be measured without knowing what an accurate response to that noise would be.

Although an organization is always, or at least most of the times, aware of biases and noise in decision-making by its employees, collecting information is not always an easy task. The first complication when collecting information is that the results of a decision can be forecasted, but the final outcomes of that decision are not known until a future date. Moreover, even if organizations are aware of noise, they tend to overlook the fact that noise can interfere with the reliability of professional judgment. The organizations inability to relate noise to experience is the consequence of two phenomena. First, experienced professionals tend to exhibit high confidence in their judgment and second, they tend to view their colleague's intelligence with respect (Kahneman, Rosenfield, Gandhi and Blaser, 2016). The two phenomena together inevitably lead to an overestimation of agreement, where individuals expect other decision-makers' judgements to be closer to their own than they actually are, hence without imagining plausible alternatives to their judgment decisions. Noise can be significantly dialed down at high levels of skill and when actions are immediately followed by a clear feedback. Unfortunately for managers, few of them get immediate feedback. Since feedback is not readily-available, professionals tend to learn to make judgment by listening to their colleagues' explanation and critiques about a decision taken, which is less effective than learning through experience, thus failing to dial down noise. In addition, experience tends to increase confidence in decision-making, but confidence is not a guarantee for accuracy of decision. Hence, algorithms could be a valuable tool to avoid incurring in noise and bias disturbances.

Chapter II

Organizational Relationships With Artificial Intelligence And Their Impact On The Economy

2.1. What is Meant By Artificial Intelligence?

The first time the world heard about the term artificial intelligence (AI) it was back in 1956 when John McCarthy, an American computer and cognitive scientist, held a conference on the subject (Smith, McGuire, Huang & Yang, 2006). After the conference the topic was not left to drift and many scholars, including Alan Turing in his paper *Computing Machinery and Intelligence* (1950), further analyzed the topic and wrote papers discussing whether machines could simulate human beings and their ability to do intelligent things, such as playing chess. But what does artificial intelligence mean? There are multiple possible definitions of AI, but the majority of them agree on the concept of creating computer programs and machines that are capable of behaviors that would be regarded as intelligent if performed by humans (Kaplan, 2016). McCarthy, Minsky, Rochester and Shannon in 1955 described AI as the process of "*making a machine behave in ways that would be called intelligent if a human were so behaving*". However, this definition is deeply flawed as there is no agreement on what is considered to be intelligent and what is not. Moreover, some of the mostly used methods to assess intelligence, such as the speed at which students can add or subtract numbers of a certain list, would not be consistent when applied to machines as machines will beat any human at this task. Machines can also perform a wide range of tasks, such as security programs, that individuals cannot perform, so is it accurate to compare machines' intelligence with that of humans?

Despite the fact that the answer to this question is still open to debate, the essence of AI is clear and coherent between most scholars. AI is the ability to make suitable generalizations at the right time based on limited available data. The broader the domain in which the generalization can be applied, the quicker the conclusion will be drawn despite having minimal information. In addition, Kaplan (2016) also argues that machine intelligence can be the ability to perform sequential generalization by taking prior experience into account to perform future analysis.

2.2. The Influence of Artificial Intelligence on Human Behavior

AI is starting to become increasingly present in our daily lives and because of this, it is set to drastically transform economies and societies. Moreover, AI is set to affect behavioral economics and Camerer (2019) points out three major implications. First, AI can be used as a tool to identify variables that affect behaviors. Second, obstacles in implementing AI can help individuals understand common limitations of human cognition. And third, behavioral economics is an essential element to understand whether AI can overcome human limitations.

In strategic deviations humans tend to diverge from equilibrium because of their inability to assess the other individual's behavior. This will inevitably happen also when the other player is an AI agent. Hence, a question raises spontaneously. As the interaction between AI and humans will become more frequent, will humans ascribe the wrong amount of rationality to AI agents?

Research conducted by March (2019) examined the behavior of human players against computer players (CP) in a variety of games including auctions, mixed-strategy equilibrium games, public good games and bargaining games among others. March's research revealed some fundamental insights of human behavior when faced with machines. First, human behavior changes when the opponent is a CP. Second, individuals tend to behave more selfishly, which is contrary to what is expected in behavioral economics, and more rationally when interacting with computers as they believe that the counterpart cannot be harmed and because they believe the CP is better at predicting their moves than a human player. Third, individuals can learn to exploit CPs even if they do not follow a fixed strategy but are responsive to the individual's choice, and if the human player only has little prior information on the computer available. The last element that can be observed when human players interact with CPs is that there are ceilings to exploitations as sophisticated algorithms can successfully outperform human players in given environments.

2.3. Artificial Intelligence and Its Integration Within Organizations

Nowadays organizations are increasing the adoption of AI in a wide variety of tasks and the rapid adoption of AI from organizations can be attributed to four major causes (Von Krogh, 2018). First, in the past two decades there have been significant progresses in technology and science underlying AI methods. Moreover, many companies started to provide these AI technologies under open-source licenses, such as Amazon Alexa and Google's Tensorflow, thus increasing their adoption. Second, information technology is becoming a successful tool for efficiently capturing and storing task related data within organizations. The stored data is fundamental as it is the foundation of the algorithms of AI which are the premises for task automation. Third, even though AI requires extensive computation, the price of computer hardware is becoming increasingly affordable and the computational power's price is decreasing. Fourth, AI is becoming available to different types of organizations as the cloud-based services are improving and are increasing the available space.

As machines are becoming increasingly powerful and affordable, their business value is no longer limited by their computational capability, it becomes limited by the managers' ability to create and apply new processes, procedures and organizational structure to take advantage of the machines' capabilities.

2.3.1. Information Age Organizations

Information technology, which refers to computers and to the related technology and digital communication (Brynjolfsson & Hitt, 2000), has the power to reduce the coordination, information processing and communications costs. Hence, investments in information technology are of fundamental importance as they are linked to higher productivity and organizational transformation. However, in order to take advantage of these investments, organizations need to match their structures and capabilities to the technological capabilities, thus transitioning to information technology-intensive organizations (*ibid*).

In order to become effective information age organizations, companies need to restructure their businesses. Considering the fact that one of the organization's main activities is to engage in information processing, which entails transforming data into decisions, Mendelson and Pillai (1999) suggest that to succeed in dynamic and information-rich environments, firms need to engage in a blend of decentralization of decision-making, which allows the firm to respond quickly to new information; adoption of practices which aim at the promotion of absorption and diffusion of information, thus making accurate information available to the decision-makers; and finally development of extended inter-organizational networks, in order to prevent information overload through outsourcing. Moreover, Mendelson and Pillai (1999) defined concrete numerical measures for Informational Age organization efficiency and combined them into an index they refer to as organization's IQ, which they also regard as a good predictor of an organization's success.

Galbraith (1998) proved that decentralization enables faster and more effective decision-making in information-rich environments. In fact, the Information Age structure, by adopting decentralization, relieves information-processing bottlenecks as it moves decision-making to lower hierarchical levels and co-locates them with relevant information (Anand & Mendelson, 1997). Since Information Age organizations adopt a decentralized structure, they are able to take advantage of the tacit knowledge of line employees who are better informed on the necessary adjustments that need to be made at the shop floor level, thus allowint to adapt the production process to the present needs. According to the IQ measurement, organizations that give line employees the possibility to participate in decision-making scored a higher IQ than those that do not. Hence, decentralization should be applied also to the other levels of the organization. Moreover, it is important to assess whether the project manager, who has specific knowledge about the project with respect to the available resources, and not a manager at a higher level, who is not personally involved in the project, is the one taking the decisions. Thus, the second variable of organizational IQ is higher for organizations in which the project manager takes the final decisions regarding the project design and requirements. Another IQ variable is linked with decentralization and assesses who has the final decision in determining a project's termination. In order to score a high IQ level, the decision should be taken by the product development manager rather than by a high-level manager, who again has less information about the project.

Because decentralization is key for Information Age organizations, firms should provide incentives in order to ensure the optimization of the chosen organizational structure. As pointed out by Milgrom and Roberts (1992), decentralization can cause agency problems, hence it is important to develop an incentive structure to align employees' utility functions with those of the organization's objectives. The lack of alignment between employees and the organization will lead to costly conflicts, not only between the organization and the employees, but even among the employees themselves. Because of this, rewards should not be based on locally measured performance, but should be measured on the entire organization's performance. Group performance incentives increase the employees' willingness to share information, whereas individual performance incentives are likely to result in information withholding. Therefore, another important IQ variable is measured by observing whether incentives are based on the end results or not. In order for the incentive structure to work it is also important to assess whether lower level employees receive the same rewards as higher level managers since a reward structure providing the same variety of rewards would facilitate the decentralization of decision rights.

Organizations also need to always be aware of the state of technology, of its customers, products and of its competitors in order to effectively and quickly respond to changes in the external environment. Without monitoring the situation, organizations lack a clear market orientation, which is fundamental to ensure profitability (Jaworski and Kholi, 1993). Hence, two variables of organization IQ measure the extent to which discussions with customers and marketing personnel contribute to generate new ideas for product development, and secondly the extent to which customer preferences influence cost reduction objectives. In addition to external information, decision-makers should also have access to internal information, including tacit knowledge, in order to enable product development. According to Clark and Fujimoto (1991) there are four alternative methods to organize for product development, namely: functional structure; the lightweight project manager systems, in which a project manager coordinates the functions; the heavyweight project manager structure, where a project manager is responsible for a project organized by functions; and lastly the project team, in which the heavyweight project manager works with a team whose members are assigned to the project full-time and report to the project manager. Hence, another dimension of Information Age organization is to assess if they adopt product development by project teams, as it proves to be the most efficient method to enhance knowledge assimilation.

The evolution of organizations from mechanistic structures to Information Age structures is required in order to fulfill the higher information-processing required by the increasingly uncertain environment and because of the increased presence of machines within organizations. Moreover, this type of structure will favor the adoption of the increasingly present and relevant information technology.

2.3.2 Information Age Organizations and Organizational Design

To create a successful information age organization, Tushman and Nadler (1978) highlight the differences between mechanistic and organismic organizational structures as the different structures determine the business' capability to process information and cope with uncertainty. Organismic structures are characterized by decentralization and highly connected internal communications. These elements increase the scope for feedback, thus enhancing the problem-solving capacity of individuals. On the contrary, mechanistic organizations exhibit rigid communication channels due to their hierarchical structure, thus limiting the flow of information and making the organization more susceptible to information overload.

The difference among the two structures proves that organismic structures are more effective in dynamic environments where fast decision-making and ability to act on local knowledge is fundamental for the organization's survival (Mendelson & Pillai, 1999). Moreover, contingency theorists Lawrence and Lorsch (1967) argue that organizational design is responsive to environmental uncertainty. Hence, when tasks are subject to uncertainty, organizations respond by differentiating their structures. If the tasks present high variance among them, there will be a high degree of differentiation among the multiple parts of the organization, thus creating the need for integration so that information from different tasks can be recombined.

The relationship between task uncertainty and information process requirements, calls for a combination of strategies to reduce the volume of information processing while increasing the information-processing capacity (Galbraith, 1973). Galbraith (1973) suggests decentralization as an information-reduction strategy as decision-making is moved closer to the source of information, thus travelling through fewer levels. However, for a decentralized authority to work, it should be supported by strong vertical and lateral information ties through information technologies and team structures.

Alternatively, transaction costs economics provides a foundation to the theory of organizational design which focuses on institutional arrangements as tools to organize for economic transactions (Williamson, 1976). According to Williamson the boundaries of the firm are imposed by balancing the costs of carrying operations within the firm, with those of conducting the same operation through market transaction with third parties. When using third parties it is important to account for asset specificity, uncertainty and transaction frequency, as an increase in any of these characteristics augments the cost of using market transaction compared to the hierarchy mechanism. Because of these restrictions, Williamson (1991) recognizes that firms could adopt a hybrid mode that lies between the hierarchy and the market mechanism. In a hybrid mode, a firm is considered a "*quasifirm*" which operates through close, stable and long-term relationships between a general and a specialized contractor. Conducting transactions through sub-contractors will be more effective than vertical integration since the organization will not require a wide range of labor specialties. Moreover, a stable long-term relationship will also reduce transaction costs for both parties (Mendelson & Pillai, 1999).

Another organizational structure that can be adopted as a natural response to higher business dynamics is the network organization (Powell, 1987). According to Powell (1987), as the pace of technological changes increases, product life-cycles shorten and markets become more specialized, vertical integration may become a disadvantage. In response organizations disaggregate, hence limiting the size of the workgroups and set up ventures with suppliers and distributors. Since rapid technological change requires up-to-speed knowledge, networks which are bounded not by ownership but by information sharing start to form. A network organization is more agile than a hierarchical one and is quick and efficient with respect to information sharing. Hence, networks have a critical role in the development and sharing of knowledge, as the locus of innovations shifts from the individual firm to the inter-organizational network (Powell, Kenneth, Koput & Smith-Doerr, 1996).

The three organizational structures presented above are possible solutions to organizing a business in a dynamic and information rich environment. In fact, the Information Age Organization is based around three main characteristics: decentralization, resulting in co-location of decision rights; information practices which promote awareness for external information and information sharing; and network structures which view the firm as part of a larger inter-organizational network. The three structures are closely linked as they share the same rationale, that of supporting efficient and effective information processing in dynamic environments.

2.3.3. Data-Driven Decision Making

As the technological advancements progress, new opportunities to collect and leverage data have led to changes in managerial decision making as managers are starting to rely more on data and less on intuition. The change in the managerial decision-making pattern is well underlined by a quotation by Jim Barksdale, the former CEO of Netscape: "*If we have data, let's look at data. If all we have are opinions, let's go with mine*" (Barksdale, 2019). Even tough better data creates opportunities to make better decisions and the majority of organizations started to adopt data-driven decision-making (DDD), the rapid diffusion is uneven. The discrepancy arises as DDD is more present in organizations that show three key characteristics: correct size, high levels of complementary investments, and awareness of these practices and their implementation (Brynjolfsson & McElheran, 2016).

Complementarity between DDD and information technology (IT) related investments is a critical component for the successful implementation of DDD. This relationship especially holds for single-unit plants as younger populations might be more sensitive to IT advancements that make it more powerful and less expensive over time. As DDD requires up-to-date IT, it is quite intuitive that if organizations make substantial investments in IT, it will be easier from them to have greater rewards from DDD and vice versa. As greater levels of IT are adopted, a general movement towards standardized and structured management can be observed (Brynjolfsson & McElheran, 2016). Because standardized management is a product of the increased IT investments, this type of managerial structure is also positively correlated with the adoption of DDD. For firms with highly variable production processes it is fundamental to focus on greater instrumentation of process, mechanization and standardization (Bohn, 2005) as they need to invest in managerial decisions that allow them to develop a process to effectively select and collect data. To reach this type of process firms will have to engage in a procedure that allows them to acknowledge what they already know by discovering the knowledge scattered through the firm by consulting with employees. Hence, this process also becomes useful for capturing tacit knowledge from employees through less formalized channels. Although, standardization seems to be positively correlated with DDD, this could also be a size effect as large organizations are more likely to present both a structured approach to management and DDD (Brynjolfsson & McElheran, 2016).

Another important characteristic for the successful adoption and implementation of DDD are the organization's background characteristics. If the majority of the workers and managers within an organization are educated, the more likely for the organization to adopt high levels of DDD. The relationship between education and DDD is reliant on the existent complementarity between skilled labor and DDD (Brynjolfsson & McElheran, 2016). In addition, having a larger number of managers and consequently of layers of management may require objective and standardized measures to facilitate coordination among the hierarchical levels.

The age of an organization also plays an important role in the adoption of data-driven decision making (*ibid*). Older plants are less-likely to adopt DDD as in older organizations there might be common resistance to new technologies as employees might believe that their experience and tacit knowledge is substituted by objective data.

A variable that is negatively correlated with the adoption of DDD is the CEO's experience. Managers with a high level of experience tend to rely more on their knowledge and experience than on data-driven practices and formal data collection to infuse their decision with authority (Porter, 1996). In addition, even if these managers tend to relay on data, it might be possible that they attribute their subjective perception to it, thus downwardly biasing it.

The last variable that contributes to the successful adoption of DDD by an organization is the organization's awareness of this practice. Variation among methods adopted by different organizations has a significant impact on whether the firm will adopt DDD or not. Simply enough, those organizations that do not put effort in learning, therefore absorb less about new managerial practices, are less likely to adopt DDD.

As data-driven decision-making increases productivity by 3% (Brynjolfsson & McElheran, 2016), organizations should do their best to adopt this managerial practice. Furthermore, there is a significant discrepancy between adopters and non-adopters as more DDD is always associated with better performance. Moreover, early adopters can manage to remain ahead of competitors who do not realize the benefits of DDD in time, thus leading to increasing performance differences and to gain a competitive advantage, at least in the short term.

2.3.4. Artificial Intelligence Alignment

According to the two fundamental welfare theorems (Arrow, 1951; Debreru, 1959) perfectly complete and competitive markets will determine the final distribution of goods or social welfare weights chosen by society. However, in practice most markets fail to be perfectly competitive and complete, and because of this, cost caused by failure in alignment are introduced (Hadfield-Menell & Hadfield, 2019). Since welfare functions are based on subjective assessments of personal utility, any coherent social welfare function should be based on collective judgements on what values to pursue, hence there is inevitable misalignment between human values (Sen, 1985).

When focusing on the design of AI agents, misalignment is the cause of economic losses associated with delegation of decisional power. AI designers are challenged with the task of achieving the intended objectives while acknowledging the limitations that arise from transforming those goals into implementable algorithms to guide an artificial agent's behavior. Because of this it comes spontaneous to think of agent and principal misalignment only in terms of design, which sometimes is the case, especially when there is a misspecification from the designer's part. However, the most relevant cause of misalignment is incomplete contracting as complete contracting is routinely impossible and costly (Hadfield-Menell & Hadfield, 2019).

There are a variety of causes which lead to contract incompleteness in the human context, hence only the most commonly cited will be mentioned. The first cause is unintended incompleteness, also known as bounded rationality. This situation occurs when contract designers are not able to fully identify the circumstances that can potentially affect the value of the contract (Simon, 1955). A second possible cause of contract incompleteness is that contract designers often tend to economize on costly cognition and drafting, hence contract terms do not cover all the possible circumstances (Shavell, 2006). Similarly, contract designers tend to economize on enforcements costs by leaving out terms which are expensive to enforce as they require costly evidence and because their potential disputes can augment with the intricacy of the contract (Schwartz & Scott, 2003). The last major cause of contract incompleteness is non-contractibility. Some decision-makers choose to leave particular contingencies out of the contracts because they cannot be observed or because they are highly expensive to verify. The costly verification of those actions can be caused by hidden information or by the difficulty of describing the action in unambiguous terms (Maskin & Tirole, 1999).

The major causes of contract incompleteness in the human context can be translated to the AI context. A contract can be viewed as an implemented reward structure for an individual, hence the non-contractibility within AI context can be thought of as a learning problem which cannot be solved by using the known techniques (Hadfield-Menell & Hadfield, 2019). Within the AI context, rewards might fail to address all the possible circumstances as algorithm designers cannot think of all the possibilities. Moreover, costly enforcements in human context can be analog to the AI context as there is a problem of AI alignment which occurs because of the differences arising between the specified reward function and human actual values as a result of engineering limitations (Hadfield-Menell & Hadfield, 2019).

Although contractual incompleteness is undesirable, as it generates economic losses which arise from low costs of drafting and excessive simplicity, in some settings it is desirable as completeness might but feasible but not optimal (*ibid*). In some cases, information at the time of contracting might be incomplete, but the new information is anticipated in the future. An example of this context is planned negotiation. In this scenario, at the start of the relationship contract designers intentionally choose to draft an incomplete contract, instead of writing a contract based on incomplete information, which they expect to renegotiate in the future once more relevant information becomes available (Bolton & Faure-Grimaud, 2010). Another situation in which complete information contracts are not optimal is when optimal completion of contracts from a third party is feasible. Once again, instead of drafting a complete contract based on incomplete information at the start of the relationships, contract designers choose to write incomplete contracts. They later expect the contracts to be filled by a third-party adjudicator with better information about the future (Shavell, 2006). Hence, by exploring contexts in which optimal contracts do not result as optimal, a new challenge emerges for contract designers as they need to choose between developing a contract with a complete reward structure today or deferring the building of the rewards until more information has been acknowledged. From here also another concern is raised, that of increasing the risk of misalignment by immediately releasing an AI system while being aware of the fact that delay of release could allow for the gathering of better information for reward design.

Lastly, economists and legal scholars have proposed that incomplete contracts might be the result of strategic behavior (Hadfield-Menell & Hadfield, 2019). Contract designers might want to implement strategic protection of private information. A party that has private information about a missing contingency will not contract to cover the contingency as doing so would reveal their private information which then reduces the value of the contract (Ayres & Gertner, 1989). Similarly, both parties might choose to deter strategic investments in costly cognition. In this case the parties might choose not to cover all the possible contingencies as learning about them would be costly and biased. Hence, the parties make efforts to shield against strategic wealth transfers that will occur if the contingency arises (Tirole, 2009). The last strategic behavior that might lead to incomplete contacting is strategic ambiguity. The parties might choose not to include all the contractable and known contingencies in order to control strategic behavior in response to non-contractable contingencies. As AI agents will not directly bargain with human designers over the reward function, human strategic interactions cannot by directly translated to the AI context. However, the human designer's strategic consideration could still lead to design choices that deliberately deprive the AI agent of complete specifications of everything the designer cares about. Strategic incompleteness in reward design is a key element to observe as it is very relevant when designing increasingly advanced systems, more than those we have today (Armstrong, 2015). If a robot is able to predict that a human may rewrite the reward structure, the robot, implementing the initial reward, may start to behave strategically and withhold information to influence the rewriting in order to keep the initial rearward structure.

Reward misspecification therefore is fundamental and is not solely a result of poor engineering as it might become predictable, unavoidable and routine (Hadfield-Menell & Hadfield, 2019). However, aligning AI machines with humans will require building advanced technical tools that allow AI agents to do what human agents do naturally. This would imply incorporating into their assessments of reward the costs associated with taking actions that would be regarded as wrong by human communities.

2.4. The Impact of AI: Prediction Tasks and Labor Market

A great deal of the public attention paid to artificial intelligence concerns the impact that it might have on jobs. However, the increased use of machine learning and AI agents by organizations does not represent an overall increase in general artificial intelligence of the kind that is able to substitute all aspects of human cognition, but it represents an increase of one specific aspect of intelligence, that of prediction (Agrawal, Gans & Goldfarb, 2018). Prediction is defined by Agrawal, Goldfarb and Gans (2019) as using existing data to fill in the missing information. Prediction is very useful as it is a valuable input for decision making, but at the same time prediction has no value if it is not followed by a decision. Stated in simple terms, prediction is a complement of decision making as it denotes the confidence of a probability in an uncertain environment.

Although fundamental, prediction is not the only component of decision-making, as decision-making also involves multiple steps such as collection of data, and taking actions based on decisions and judgments to evaluate the payoffs associated with the possible outcomes. Even if prediction is not a unique component of the process, advances in prediction technology still affect human labor in a variety of ways such as substituting capital for labor in prediction tasks; by increasing the returns to capital versus labor through automatic decision; by enhancing labor when automated prediction tasks augment labor productivity, thereby increasing the relative returns to labor versus capital; and lastly by generating new decision tasks when automatic prediction decreases uncertainty thus enabling new decisions that before seemed unfeasible (Agrawal, Gans & Goldfarb, 2019).

The first concern that the public has with respect to AI is that it may substitute capital for labor in prediction tasks. This is an understandable concern since some tasks, such as demand forecasting, are increasingly being performed by AI agents instead of humans. Moreover, some tasks, which previously were not regarded as prediction tasks, are being transformed into prediction tasks in order to be performed by AI agents because of improvements in machine learning and because it reduces prediction costs. An example of a task which has been transformed into a prediction one is recruiting. Recruiting involves prediction as based on CVs, cover letters and interviews one needs to choose which candidate will perform better for a determined position. Consequently, promotion also involves prediction as one needs to predict which employee will perform better at a higher-level. Lastly retention also is a prediction tasks as organizations need to predict which employee might leave and which are the available incentives they can offer to encourage him to stay. Today a growing number of multinational companies is adopting AI for recruiting including Five Guys Burgers and Fries, AT&T, Hilton and many more (Jackson, 2019).

The second way in which AI affects labor in a task-based framework is by augmenting the relative returns to capital versus labor in complementary decision tasks. Intuitively, machines reaction times are faster than human reaction time. In addition, sometimes AI agents are able to make better predictions than humans as they can process a wider variety of data, as it happens with prediction tasks related to vehicles since machines have access to cameras. Hence, once the prediction task is automated, it increases the return of some complementary tasks.

Third, automatic predictions tasks might not impact the productivity of capital performing a complementary task, but it might increase the labor productivity. Automatic prediction tasks are increasingly adopted by organizations through their CRM software such as Salesforce and Dynamics 365. CRM software adopt automatic prediction performed by AI agents through a variety of different tasks. An effective example can be portrayed by taking in consideration a delivery company. Clearly enough a delivery company will need to deliver products to its customers and will do so through a wide array of delivery men, each of whom will have a different vehicle and will prefer to delivery in a certain area. It is sufficient for the human agent to insert the drivers' profiles within the software and from that moment on automatic prediction will substitute the human labor. It will be enough for the human agent to specify the delivery address and the type of product to deliver and the machine agent will take care of the rest by predicting which delivery man is more suitable, by understanding which one has the correct type of vehicle while also being able to predict which delivery man will be faster and facilitated to make the delivery based on their location and on the traffic situation. Hence, automated prediction can improve human choices as it will improve the speed with which the decision is taken. Furthermore, the human agent's work results will improve as automated prediction is able to take into account more variables than a human agent would be able to in a lower amount of time allowing him to focus on more human-related tasks.

The last type of direct impact that automated prediction tasks have on human labor is when automated prediction reduces uncertainty and enables new decisions that were not possible before. Whether the new tasks will be performed by capital or labor will depend on their costs (Agrawal, Gans & Goldfarb, 2019). Moreover, as some tasks become unfeasible because of high uncertainty, automated prediction technology will make them feasible as it reduces the level of uncertainty. The idea of technology supporting human labor relates to the reinstatement force described by Acemoglu and Restrepo (2019) where the return to technologies that use labor for new tasks increases as labor is freed by automation.

2.4.1. Indirect Effects of Automated Prediction

Besides the four above mentioned direct effects of automated prediction tasks on human labor, there are also indirect effects. Since some tasks will become more efficient with automated prediction, demand for upstream and downstream tasks might vary consequently. A useful example is provided by Brynjolfsson, Hui and Liu (2018), who describe how an automated trading platform enhances international trade, but at the same time affects translators and all the upstream and downstream labor involved in the trade.

The impact of the indirect effects will be subjective for each individual worker as the forces depend on the degree to which their core skill is based on prediction. If the worker's core skill is not prediction, he might find automated prediction to be a useful tool for his occupation. On the other hand, if the worker's core skill is prediction, such as for those who work in human resources, they might find that their occupation is being harmed by automated prediction.

2.4.2. Automatic Prediction and The Development of New Tasks

As AI agents improve prediction and decrease uncertainty, they allow for decisions to be taken where it was previously unfeasible because it was either impossible or too costly to make them. As emphasized by Simon (1972), individuals rely on rules when faced with bounded rationality. These rules can have various forms, however when uncertainty is reduced, the general rules may be replaced by probability-driven decisions.

Cockburn, Henderson and Stern (2019) addressed the improvements in artificial intelligence as being *"research tools that not only have the potential to change the method of innovations itself, but also have implications across a wide range of fields"*. In addition, Agrawal, McHale and Oettl (2019) underline how artificial intelligence is able to influence the knowledge production function and further analyze the implications of using AI in order to produce a map of the search of space ideas in order to reduce the costs of prediction. Simply stated, besides increasing the demand for the existing tasks, AI will most probably generate innovations that will lead to new industries and new types of jobs within those industries.

2.5. The Impact of Artificial Intelligence on Economic Growth

The last decades of economic progress have been driven by automation, and artificial intelligence might become the next tool for progress. Zeira (1998) provided a model of automation which considered the following production function:

$$Y = AX_1^{\alpha 1}\, X_2^{\alpha 2} \cdot \ldots \cdot X_n^{\alpha n} \text{ where } \sum_{i=i}^{n} \alpha i = 1$$

While Zeira interpreted X_i as intermediate goods, by following Acemoglu and Autor (2011) X_i can be referred to as tasks. Non automated tasks can be produced one-for-one by labor, but once a task is automated one unit of capital can be used instead

$$X_i = \begin{cases} L_i & if\ not\ automated \\ X_i & if\ automated \end{cases}$$

Hence, if labor (L) and aggregate capital (K) are assigned optimally to the tasks, the production function can be expressed as

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$$

where α, a constant, reflects the share and importance of the tasks that have been automated.

Next, Aghion, Jones and Jones (2017) embed in the setup a neoclassical growth model with constant investment rate. The portion of factor payments going to capital is expressed by α and the long-run growth rate of $y \equiv Y/L$ is

$$gy = \frac{g}{1-\alpha}$$

where g is the growth rate of A. It can be observed that an increase in automation will lead to an increase of the capital share α and will increase the long-run growth rate because of the multiplier effect associated with capital accumulation.

Zeira's framework (1998) predicts that growth rates and capital shares should increase as automation rises, however this has proven not to be accurate and Acemoglu and Restrpo (2018) provide a new approach to solve the issue. They underline that research can follow two approaches: either discover how to automate existing tasks or discovering new tasks to be used in production. In their framework α represents the fraction of tasks that have been automated, leading them to discover a possible resolution to Zeira's shortcomings by simply implying that humans are inventing new tasks at the same pace at which we are automating old ones.

Aghion, Jones and Jones (2017) also consider another aspect of automation as a tool for economic growth, that of Baumol's cost disease. Baumol (1967) observed that fast productivity growth in some sectors relative to others could generate a "*cost disease*" as the unit costs of labor in slow growth sectors increase while wages in high growth sectors increase. Hence, it is important to observe if automation can be a factor behind the creation of the cost disease.

GDP is a constant elasticity substitution (CES) combination of goods with an elasticity of substitution lower than one

$$Y_t = A_t \left(\int_0^1 X_{it}^{\rho} di\right)^{1/\rho} \quad \text{where } \rho < 0$$

At = A_0e^{gt} represents standard technological change, an exogenous variable. Because the elasticity of substitution is less than one, the tasks are gross complements, meaning that GDP is limited by the output of the weakest links (Aghion, Jones & Jones, 2017), causing the Baumol effect.

As it is in Zeira's framework, technical change is also caused by automation of productions, non-automated goods can be produced one-for-one by labor, while automated goods can be produces with one unit of capital

$$X_{it} = \begin{cases} L_{it} \; if \; not \; automated \\ X_{it} \quad\; if \; automated \end{cases}$$

The rest of the model is neoclassical (Aghion, Jones & Jones, 2017):

$$Y_t = C_t + I_t$$

$$K_t = I_t - \delta K_t$$

$$\int_0^1 K_{it} di = K_t$$

$$\int_0^1 L_{it} di = L$$

For simplicity, fixed endowment of labor is assumed.

Let β_t be the fraction of goods automated at time t and assume that capital and labor are allocated symmetrically to tasks. Hence, K_t/β_t are employed in each automated task and L/(1-β_t) units of labor are employed for each non-automated tasks. The production function will therefore be

$$Y_t = A_t \left[\beta_t \left(\frac{K_t}{\beta_t} \right)^{\rho} + (1 - \beta_t) \left(\frac{L}{1 - \beta_t} \right)^{\rho} \right]^{1/\rho}$$

which can be simplified to

$$Y_t = A_t \left(\beta_t^{1-\rho} K_t^{\rho} + (1 - \beta_t)^{1-\rho} L^{\rho} \right)^{1/\rho}$$

From this function it can be observed that the allocation of resources can be decentralized in a standard competitive equilibrium and that the share of automated goods in GDP is equal to the share of capital in factor payments. Also, the share of non-automated goods in GDP is equal to the labor share of factor payments. Hence, the ratio of automated to non-automated output is equal to

$$\frac{\alpha K_t}{\alpha L_t} = \left(\frac{\beta_t}{1 - \beta_t} \right)^{1-\rho} \left(\frac{K_t}{L_t} \right)^{\rho}$$

From the above equation it can be observed that there are two forces that move the share of automated economy. First, if the fraction of automated goods (β_t) increases, also the share of automated goods in GDP will increase together with capital share, holding K/L constant. Secondly, as K/L rises, the capital share and the value of the automated sector as share of GDP will decrease. If an elasticity of substitution less than one is present, the price effect will dominate. Simply, the price of automated goods will decline relative to that of non-automated goods.

As more sectors become automated, β_t increases, the share of automated goods and capital will increase. However, since automated goods are subject to faster growth, their price will decline as will their GDP.

2.5.1. Artificial Intelligence and Its Macroeconomic Effects

The advance of AI and its macroeconomic effects will depend on the behavior of the firm. By considering different firm's behavior it is possible to further analyze AI as first-order issues might emerge when market structures and sectoral differences are taken into account.

The market structure plays an important role in shaping the macroeconomic effects of AI. Previous research on competition and innovation-led growth highlight the existence of two contradictory effects. The first effect is that more intense market competition, or imitation threat, induces competing firms at the technological frontier to innovate in order to beat competition. However, the second effect implies that because of more intense competition, firms that fall behind the current technological frontier tend to be discouraged to innovate and to catch up with frontier firms. The dominant effect will depend upon the level of advancement and competition of the economy. With low initial levels of competition, the first effect is more likely to dominate, while in high levels of competition and less advanced economies the discouragement effect is more likely to happen (Aghion, Jones & Jones, 2017). At this point it comes naturally to ask whether AI can facilitate imitation of present technologies and products. By following the inverted-U logic (Aghion, Bloom, Blundell, Griffith & Howitt, 2005) it can be observed that with low initial levels of imitation, AI, through reverse engineering, can stimulate innovation, thus enabling the escape-competition effect. However, if the imitation threat is too high, innovation will be discouraged as innovators will be subject to excessive expropriation. Moreover, the introduction of AI might lower the time it will take sectors to become congested, thus leading to decreasing returns in innovation within existing sectors (Bloom, Garicano, Sadun & Van Reenen, 2014). On the other hand, it might also lead potential innovators to use more resources to create new lines in order to avoid imitation and competition.

The second item that should be taken in consideration to assess the impact of AI is sectoral reallocation stemming from knowledge diffusion. Baslandze (2016) argues that a knowledge diffusion effect has developed because of the IT revolution, which in turn lead to relevant sectoral reallocation from sectors that do not heavily rely on technological externalities to sectors that do. Baslandze's study proposes two counteracting effects of IT on innovation incentives which can also be relevant for AI expansion. On one hand firms can learn from each other and therefore benefit from knowledge diffusion more than other firms would. On the other hand, however, the increased access to external knowledge can increase the scope for business-stealing. The knowledge diffusion effect will dominate in sectors in which firms will benefit from external knowledge, while the latter effect will dominate in sectors in which firms do not strongly rely on external knowledge. Hence, A.I. adoption should lead to the expansion of sectors relying more on external knowledge at the expense of more self-contained ones.

2.5.2. AI Adopting Firms

When adopting ML, organizations start to questions how different aspects of the organization, such as vertical integration, size, price and returns to scale, will change. Whether and how the before mentioned factors will change will depend on multiple variables.

To understand how the firm size and boundaries will change, the relationship between fixed and vertical costs should be observed. If organizations are adopting costly and customized solutions to their issues, fixed costs are expected to be high and the organization's size must be large in order to amortize the costs. However, if a firm prefers to buy less expensive off-the-shelf services from third parties, fixed costs and minimum efficient scale would be quite small (Varian, 2018).

The availability of increasingly efficient ML also offers the opportunity to adjust prices according to customer characteristics, as it is done in auctions. However, it is fundamental to acknowledge that customers have themselves access to information, hence price differentiation should be done in a smart way. Airlines can adopt pricing strategies that link the price with the departure date, however reverse-engineering services have been developed and are able to advise customers on when it is more convenient to purchase (Etzioni, Tuchinda, Knoblock & Yates 2003), thus making airline pricing algorithms less efficient.

Another fundamental aspect that might change because of ML adoption are returns to scale, especially supply-side ones. When developing a software, firms incur in high developing costs and in a small variable cost of distribution. However, software development is not a one-time operation as almost every algorithm is updated over time, just as it happens with phone systems and is increasingly happening with smart televisions. This characteristic of software is blurring the distinction line between goods and services, since common goods such as televisions, are no longer static goods but are devices that allow the consumers to reach a wide array of linked services (Varian, 2018).

2.5.3. AI Providing Firms

AI providing firms are questioning their structure and future just as much as AI adopters are. One of the first concerns that arises within AI providing firms is how easily the customer will be able to switch between AI providers. Container technologies such as Google Cloud allow users to run the application independently of other processes, thus making it easier for customers to switch providers. Moreover, open-source platforms such as Dockers and Kubernetes are also readily available to customers, thus simplifying provider switching. Because of these increasingly available technologies, lock-in will not be a problem for small and medium size applications, however large and complex applications could encounter some issues as they require much customized work (Varian, 2018).

Another area in which AI providing firms are focusing is price. As it happens in other information-based industries, software is cheap to reproduce but expensive to produce. As hardware installations are easy to replicate at the level of motherboard and data centers, computer hardware tends to exhibit constant returns to scale (Varian, 2018). Moreover, ML is extremely competitive, thus providers will try and differentiate themselves by providing superior speed and accuracy. Firms that are able to offer the better services will also be able to charge premium prices to the amount that customers are willing to pay for those services. However, current speed and accuracy are very high, hence it is not clear how customers will react when higher levels of those dimensions are provided (*ibid*).

The last dimensions to which providing firms are increasingly turning to are the constantly expanding policy concerns raised by customers which will be further discussed in section 3.3.

Chapter III

Artificial Intelligence As A Tool For Improved Decision-Making

3.1. Managerial Deviations From Optimal Decisions

Managers make frequent deviations from perfectly rational behavior (Kahneman and Tversky, 1979) and this has a substantial impact on the value of firms and on the economy. But why do managers deviate from rationality? They deviate either because they are not aware that they are making suboptimal decisions, which is the most frequent cause, or because they do not care about making suboptimal decisions (Rode, 1997).

Most of the times, managers make decisions in the context of two theories: the descriptive prospect theory and the normative options theory. The latter theory is regarded as the extension of the net present value (NPV) analysis of a project. This theory presents some benefits as it incorporates the values of the flexibility that managers experience in decision environments. Meanwhile, the prospect theory augments the descriptive theory by adapting to it the traditional normative model of decisions by incorporating cognitive errors which are frequently made by decision makers.

Managers find themselves facing a multitude of problems, but an often recurring one is to decide whether or not to do something based on the expected results that will be generated from that action, such as in capital budgeting problems, strategic entry and resource allocation. Traditionally, it was believed that a valid solution to this problem would be to discount the present value of the future revenues. However, it was later realized that an important piece was missing in this analysis as it did not consider the value that would be generated by leaving one's opportunities open for the future (Kogut & Kulatilaka, 1994). As the traditional theory was lacking, the options approach became widely recognized as an improvement (Dixit & Pindyck, 1994). However, also this theory is not perfect as it suffers from normative insensitivity for the impact of constraints in the environment on managerial decision making (Rode, 1997). In addition, Trigeorgis (1993) argues about the actual impact of real options thinking. If a manager is not careful, he could view everything as an option, thus encouraging myopic behavior. Managers can also become overwhelmed by option thinking as it can become complex and as it requires rigorous applications, thus leading them to refuse to use option thinking in favor of simpler heuristic rules.

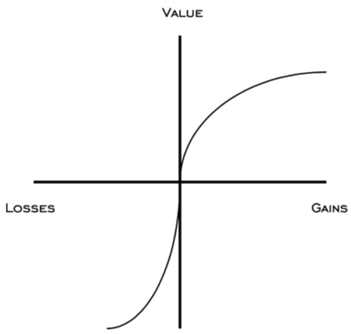
The second theory that is applied to the context of decision making is the prospect theory, a model of decision making under risk that accounts for cognitive errors that are found to systematically happen in decision contexts (Kahneman and Tversky, 1979). Classic utility theory defines utility functions only based on the final outcomes rather than taking into account gains or losses. However, Kahneman and Tversky (1979) note how gain and losses are fundamental in the decision-making process. They also discuss how the axioms of classic utility theory are violated in two areas, respectively editing and evaluation. Editing is the process by which values and probabilities of decision are transformed by the decision maker to make his decision simpler (Rode, 1997), such as coding. Simplifying the decision has a substantial impact on how the prospect's value is assessed. Humans tend to be risk averse for gains and risk prone for losses. Moreover, gain and losses are assessed from a reference point which could vary over time to reflect the changing asset position.

In order to show how prospect theory violates the axioms of the classic utility theory, Rode (1997) presents the following formula:

$$V(x, p, y, q) = \pi(p)v(x) + \pi(q)v(y)$$

The first scale, π, ascribes to each probability a decision weight $\pi(p)$, where π is not a measure of probability and $\pi(p)+\pi(1-p) <1$. The second scale ascribes to each outcome, x, a value of v(x) which represents the subjective value of the outcomes based on the present reference point. V represents the value of the choice between x, which is assigned a probability p, and y, with probability q. It can be assessed that utility theory defines the value of a choice based on the outcomes, while prospect theory defines the value of a choice based on the prospects. Simply stated, the prospect formulation, by relaxing the expectation principles of the utility theory and by allowing for editing procedures, permits the violation of the substitution axiom.

The areas in which prospect theory diverges from the expected utility theory are more clearly illustrated in figures 2 and 3.



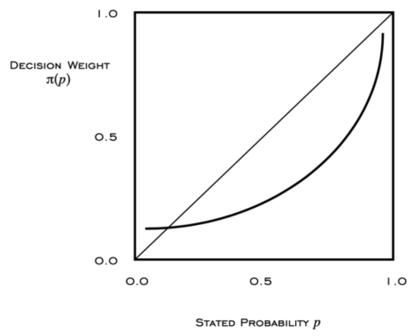


Fig. 2. Fig. 3

Source: Rode, 1997

In figure 2 the importance of losses and the incorporation of a reference point can be observed. These two factors may explain why individuals are eager to buy insurance contracts even though the price might not be accurate and it also explains why people may behave differently at different points in time. Instead, when looking at figure 3 it can be observed how probabilities can become a relevant problem for options. Decision-makers tend to over-weight small probabilities while under-weighting large ones. Hence, if the probabilities for accepting a project are near zero, a decision maker might treat that probability as being much greater than zero despite the huge losses it may cause. Moreover, since the option approach is dependent on probability assignments, in a discrete model, and on probability distribution assumptions, in a continuous model, systematic deviations will change the results of an option model, unless the decision maker is aware of these predispositions (Rode, 1997).

3.2. Algorithms as Better Predictors of Decision-Making

Recent advancements in artificial intelligence have generated excitements about the large potential of AI to transform medicine, transportation and the economy as a whole. In fact, research comparing the effectiveness of humans versus that of algorithms shows that the latter regularly outperforms humans (Meehl, 1954). Similarly, Grove, Zald, Lebow, Snitz and Nelson (2000) found that algorithms outperform human forecasters by 10% on average and that it is less common for human individuals to outperform machine actors.

Moreover, algorithms are better predictors of decision-making as biases in humans are very frequent and well documented. Research has highlighted how human judgments and decisions can be unconsciously influenced by the individual's characteristics, along with employers' biases which arise when giving different interviews to candidates with the same resumes but with names which are considered to reflect different ethnic groups (Rachlinksi, Jeffrey, Johnson, Wistrich & Guthrie, 2009). In addition, humans are also prone to misapplying information (Silberg & Manyika, 2019).

In many cases AI can reduce human subjective interpretation of data as machine learning algorithms are programmed to consider only the variables that improve their predictive accuracy based on the data used (Kleinberg, Ludwig, Mullainathan & Sunstein, 2019). Moreover, research shows that algorithms can improve decision-making by enabling it to become a fairer process (Miller, 2018). Another advantage of using AI agents for decision-making is that the AI decision process can be opened up, interrogated and examined to ensure that correct decisions are taken. The potential of machine learning for decision-making are endless and quoting Andrew McAfee "*if you want the bias out, get the algorithms in*" (Rosenbaum, 2018).

3.2.1. Fatigue Theory

Fatigue is a relevant influence on human life and is experienced by every individual during their lifetime. Fatigue can be felt by different individuals in various forms such as unfocused mental state, tension and low mood. Moreover, it disturbs the quality of life and in extreme cases can lead to incapacity of acting and lack of motivation. When looking at fatigue in terms of cognitive activities, it can have a deep impact on starting, completing and understanding tasks, such as decision-making, over a long period of time (Hockey & Hockey, 2013).

The level of fatigue is an important element to consider when observing cost-benefit analyses since, at a sufficient level of fatigue, preference reversals from high-rewarding options (HR), which can be obtained only through high costs, to low-reward, low-cost options (LR) can be observed (Iodice et al., 2017).

An experiment on mice performed by Iodice et al. (2017) reveals that when mice are subject to low levels of fatigue, they choose the HR options, but when their fatigue reaches 80%, they choose the HR option significantly less often than under lower fatigue settings. The possibility to manipulate the levels of fatigue highlights how humans and animals are subject to sharp preference reversals depending on their level of fatigue, thus underlining an index of behavioral flexibility.

Moreover, humans and animals are both able to evaluate the amount of effort required both rapidly and efficiently and prefer to maximize gains while minimizing energy expenditure (Shadmehr, Huang & Ahmed, 2016). This indicates that humans and animals consider the efforts to be made in a cost-benefit process which happens before the decision. In addition, the state of the current physiological state, or fatigue level, of an individual should be taken into account as it influences the decision-making process by promoting flexible forms of choice. Preference shifts from HR to LR are dependent on the flexible decision process which involves fast cost-benefits computations rather than slow re-adaptation to new contingencies through trial-and-error (Iodice et al., 2017).

The fact that humans exhibit preference reversals when subject to fatigue is a disadvantage when human efficiency is compared with that of AI agents. Algorithms and machines agents are clearly not subjected to fatigue, therefore will always choose the HR option if programmed that way by the designer, thus ensuring that the high reward option is always chosen with minimal effort on the human agent's part. Moreover, it is unlikely that human agents are only responsible for one decision making task a time, hence they will frequently experience high levels of fatigue which are then reversed on a high number of decision-making tasks. This can cause a great deal of economic loss to organizations since decision-makers will rarely set their tasks with the objective of reaching the high rewards. However, by using machine learning this would not be an issue as fatigue is only affecting the decision in a minimal portion. Moreover, human individuals, by outsourcing the decision-making task to machine agents, will also be able to better perform the tasks that need to be completed exclusively by human agents since they will be less fatigued and more motivated to achieve the high-level rewards.

3.2.2. Robust Patterns of Decision

Machine learning can facilitate reasoning by allowing for transparent and robust patterns of decision-making through prediction. Moreover, machine learning allows for algorithmic induction, which yields identical conclusion when it is applied by different individuals to the same data as it enables prediction by searching for complex, robust and replicable associations (Puranam et al., 2018).

Machine learning algorithms play a crucial role in facilitating inductive theorizing as they provide robust associative patterns in data, which are the foundation of prediction. The patterns can also be acknowledged as a robust stylized fact that needs to be explained by further theories and tested by additional data. Since the pattern itself is theory free but the pattern's explanation is subject to theory, separating the process of inductive theorizing in pattern detection, for which machine learning can be used, and pattern explanation, primarily human driven, presents some advantages.

Frist, pattern detection has high inter-subject reliability, hence the same algorithm used by different agents will lead to similar results (Puranam et al., 2018), allowing for replicability and for robust patterns of decision-making. Furthermore, separating detection from explanation reduces confirmatory biases, thus allowing to avoid false positives, and also enhances replicability of decision-making. The second advantage is that pattern detection through algorithms is not limited by comprehension constraints to which human individuals can be subject to (Puranam et al., 2018). Another advantage that comes from using machine learning algorithms is the possibility to tune the complexity of the patterns that the algorithm will detect depending on the objective. Third, using machine learning algorithms can shield from inductive results that are over-fitting to what is observed. Simply expanding the sample size can solve the problem of over-fitting, however it does not diminish the spuriousness of associations. Similarly, increasing the numbers of variables measured can help solve the problem of spuriousness but not that of overfitting. Machine learning algorithms instead can apply processes, such as regularization and cross-validation, that help decrease overfitting while also helping researchers observe reliable association patterns that replicate across a variety of data. This is an advantage if the aim is to build general theory from inductive efforts, however it is irrelevant if this is not the aim. Simply stated, algorithms that are designed with interpretability as their main aim can play an important role in generating stylized facts that are a fundamental input to inductive theory building.

Providing a replicable approach to detect robust and complex patterns of decision-making beyond those readily available to humans is one of the key functionalities that machine learning offers. In the fields of management and organizational research there are not many robust, or replicable, stylized facts that have been implemented. However, implementing more machine learning techniques can help generate more robust stylized patterns even within single studies. Moreover, by implementing robust patterns as starting point, internally consistent explanations can be developed, and later confirmed by being submitted to the critical hypothesis test.

3.2.3. Machine Learning and Automatic Generation of Solutions

Solutions can be hard to implement and to find for human agents, but in a wide variety of contexts machine learning can be a great tool to alleviate the pressure on human generation of solutions.

Plemenos, Miaoulis and Vassilas (2002) propose two techniques that allow the implementation of machine learning for declarative models by hierarchical decomposition. The first presented technique is based on neural networks, which allow for a reduction of the solution space thus enabling to generate only solutions compatible with the user's wishes. The second technique instead relies on the use of genetic algorithms which, starting from a set of scenes generated by the generation engine of the declarative model, allow to produce more solutions under the user's supervision.

Declarative scene modelling enables the user to create scenes by simply describing the desired properties without the need to specify how to construct them (Bonnefoi & Plemenos, 2002). Scenes are later created by the modeler using the properties given be the user. However, most of the time, scenes are described in an overly simplistic manner by the user and the description can become imprecise. The lack of precision can stem from two different reasons. The first reason is that the user does not know the exact properties of the scene to be designed resulting in imprecise descriptions. The second situation that leads to lack of precision is that the user believes he is giving an exact description when in reality his description allows for more than one solution (Plemenos, Miaoulis & Vassilas, 2002). Because of low precision the scenes proposed by the modeler as solution might not satisfy the user's desires.

Declarative modelling is composed of three phases: the description phase, in which the designer defines the scene; the scene generation phase, in which the modeler creates one or multiple scenes accommodating the description; and lastly the scene understanding phase in which the designer and the modeler analyze the scene and decide whether or not the solution is satisfactory or not.

Declarative modelling by hierarchical decomposition (DMHD) is a technique based on top-down designing of scenes (Plemenos & Tamine, 1997) and can often yield to two results. The first possible result is that if the scene is easily described and the properties are correctly implemented by the modeler, the design process is finished. However, if further description is possible, the scene is decomposed in sub-scenes and each one is described using DMHD. The advantages of DMHD is that it relies on top-down designing, hence descriptions are made locally for every scene without having to take into account other parts. It also allows for factorization of properties and enables generation in multiple levels of detail.

A declarative modeler can be used in two approaches: exploration and search mode (Plemenos, Miaoulis & Vassilas, 2002). When in exploration mode, the modeler starts from the user's descriptions and performs a full investigation of the solution space and then gives the user all the solutions found. This mode is most frequently used when the designer has insufficient knowledge of a domain and wants to learn more through exploration or when the designer is looking for new ideas. Because of the imprecision of the designer's description of the scene, the richness of the solution space is increased, allowing to obtain concrete answers from an unclear mental depiction. Moreover, since the exploration mode strongly relies on the use of imprecise properties, it is crucial to develop techniques that allow to reduce exploration costs by decreasing the number of tries during the solution search (Plemenos, Ruchaud & Tamine, 1998). Some of the solutions generated might not be of interest to the designer and since the modeler is not aware of the designer's preferences, machine learning can be used to teach the modeler what solutions should be regarded as interesting or not. In the solution search mode instead, the designer has a clear idea of what he is looking for, therefore the designer is aiming to obtain an immediate solution from using more precise properties. However, also in this case, the designer's description might be ambiguous, therefore it would be helpful to teach the modeler what solutions are not coherent with the designer's request and therefore should not be examined. This learning process would decrease the solution space as some scenes will not satisfy the initial idea and will be avoided. Machine learning in this case is a crucial tool as it simplifies the designer's role and contributes to increase efficiency.

The other technique used to implement machine learning in declarative models is by using genetic algorithms (Plemenos, Miaoulis & Vassilas, 2002). Genetic algorithms can be used as an aid to machine learning as they are meant to follow the natural evolution laws based on individuals. The implementation of genetic algorithm-based machine learning is composed of a chosen number of scenes generated by an engine with the purpose of generating subsequent scenes by verifying that the properties described by the user are constantly applied. Each initial scene is regarded as a chromosome encoded in a specific way. The main advantage of genetic algorithms is that they allow to generate solutions by using non-time consuming constraint satisfaction techniques. Moreover, the final solutions are obtained through an evolution process in which the search of new solutions is guided by the most efficient parts of the decision tree.

3.3. The Dark Side of Artificial Intelligence

Although AI presents many advantages, various concerns arise with its diffusion. To this day questions about safety, ethics and human-machine conflict play an important role in slowing the efficiency and the adoption of AI and machine learning from both individuals and organizations.

Research comparing the effectiveness of human and algorithm forecasts highlights how the latter outperforms the former. As research found that algorithms outperform human forecasters and that it is more common for algorithmic forecasters to outperform human than the opposite, it would be logical to prefer algorithmic forecasters to human ones. However, further research shows that the opposite is more likely to happen (Eastwood, Snook & Luther, 2011). Furthermore, humans put more weight on human input than on algorithmic input (Önkal et al., 2009), and harshly judge professionals who seek advice from algorithms rather than from humans (Shaffer et al., 2012). The dislike of individuals for algorithms is known as algorithm aversion (Dietvorst, Simmons & Massey, 2015). Some of the most common causes of algorithm aversion include the wish for perfect forecasts (Highhouse, 2008); the inability of algorithms to learn (Dawes, 1979); the notion that algorithms are incapable of considering specific targets (Grove & Meehl, 1996) and concerns about AI being able to make ethical decisions (Dawes, 1979). Furthermore, studies conducted by Dietvorst, Simmons and Massey (2015) emphasize additional causes of algorithm aversion. The studies highlight that observing an algorithm err induces the human agent to rely more heavily on the human forecaster rather than on the algorithmic one, even though the algorithmic model is more correct then the human agent even after the mistake. At the same time, witnessing a human err does not decrease the tendency to rely on the human agents. This trend suggests that people are more prone to abandoning algorithms that make mistakes than humans that err, even though human mistakes are often larger than algorithmic ones.

Confidence in the algorithmic model also plays an important role in AI adoption. The same study from Dietvorst, Simmons and Massey (2015) shows how participants are more inclined to learn from the model's mistakes than from human ones. Witnessing an algorithm make small mistakes decreases human confidence in the model, while witnessing individuals make large mistakes does not impact human confidence in other humans. Confidence reductions might ultimately lead to discontinuing the use of algorithms as they are not deemed as efficient. Furthermore, beliefs have an important role in the adoption or rejection of algorithms. People might choose human over algorithmic forecasts, even though they expect algorithmic forecast to outperform humans, on average, because they believe that human forecasts exhibit a higher chance of being perfect.

Algorithm aversion can also lead to negative consequences as it comes with high costs for the society. Various managerial decisions involving forecast generated by algorithms have proved, most of the times, to be better forecasts than those made by humans (Grove et al., 2000). The trend to discard algorithms after having witnessed a small mistake by their part is very problematic as it can become a barrier to adopting more efficient approaches to simplify decision-making tasks.

3.3.1. Artificial Intelligence and Bias

Section 1.2.2. highlighted how human judgments can be influenced by personal characteristics, experience, skills and by misapplication of given information. In many cases, AI can be a valuable tool to reduce subjective interpretations as the algorithms are designed to only consider those variables that improve predictive accuracy based on the used data (Kleinberg at al., 2019).

Although algorithms appear to be perfect at first glance, they can be exposed to algorithmic biases which mainly stem from the algorithm's underlying data (Silberg & Manyika, 2019). Some researches separate the model for algorithmic bias in two algorithms, respectively the trainer, which results to be biased depending on the underlying data and training process, and the screener, which is designed to make predictions based on the trainer algorithm (Kleinberg et al., 2019). The underlying data on which the algorithms are designed to work may contain human decisions or might reflect societal disparities, thus leading to the creation of bias when used. An example of such bias can be witnessed in hiring algorithms. Recently, a technological company discontinued its hiring algorithm as it was found that the system learned to favor words such as "*executed*" or "*captured*" which are most commonly found in men's applications, thus penalizing women candidates (Hao, 2019). This example raises another issue concerning machine learning algorithms, that of systems picking up illegal or socially unacceptable statistical correlation.

Another important source of bias is data generated by users which can create a feedback loop which later creates biases. A research conducted by Sweeney (2013) on online racial difference ad targeting revealed that searching for African-American identifying names resulted in ads displaying the word "*arrest*" more than would show when searching for white-identifying names. Sweeney also mentions that the ads might have been initially displayed equally, but that users might have clicked on different versions more frequently for multiple searches, thus leading the algorithm to display those ads more frequently. Since today an extensive number of algorithms reach billions of users worldwide daily, data generated by users is becoming an increasingly relevant source of bias.

3.3.2. Ethical Concerns

What does it mean for AI systems to make decisions and what are the societal consequences that stem from them? Can we deem AI systems responsible for their actions? How can we control these systems when the consequences are found in settings further away from their initial design? The way that society develops the answers to these questions will determine the level of trust individuals will have in AI and consequently AI's impact on society and its existence. As AI systems are constantly developing, it becomes increasingly important to consider the increasing need for AI to act responsibly (Dignum, 2018). The enlarged need for AI to incorporate ethics has led scholars to develop new frameworks such as the society-in-the-loop (SITL) framework by Rahwan (2017), which adapts the supervisory control concepts of the human-in-the-loop framework (HITL) to interactive machine learning.

In the HITL model a human agent is a fundamental tool in automated control processes as he is in charge of the supervisory tasks. A form of HITL in machine learning is interactive machine learning which is a valuable instrument to help machines learn faster by integrating interactive feedback from users (Amershi, Cakmak, Knox & Kulesza, 2014). Other fields in which HITL has been efficiently applied is human robot interactions (Cakmak, Chao & Thomaz, 2010), which includes adjusting the degree of autonomy given to robots (Crandall & Goodrich, 2001), teaching robots to adopt specific behaviors (Thomaz & Breazeal 2008), and helping the design of flexible human-robot teams (Johnson, Bradshaw, Feltovich, Jonker, Riemsdijk & Sierhuis, 2014). The presence of humans within the HITL framework covers two major functions: identifying misbehavior and taking corrective action, and secondly, providing for an accountable entity if the system were to cause any damage (Rahwan, 2017).

Although HITL is a useful framework it does not sufficiently consider the role of society. Moreover, what happens when AI systems perform various functions with broad societal implications? When this question rises it is helpful to shift from HITL to the Society In The Loop (SITL) (Rahwan, 2017). This framework aims at integrating the values of society as a whole in the algorithm. Shifting from HITL to SITL raises another issue, that of balancing competing interests of multiple shareholders, which is referred to as the problem of social contract (Skyrms, 2014).

When applying the STIL framework it is fundamental for the society to agree on two aspects. First, the society must solve tradeoffs between different values that can arise in certain contexts, for example the tradeoff between privacy and security (Kleinberg, Mullainathan & Raghavan, 2016). Secondly, society must agree on which shareholders will receive the benefits and the costs that they need to pay. Moreover, to apply the framework it is important to know the different behaviors that people expect from AI and to make policy-makers able to articulate these expectations to machines. Lastly, new tools to program, monitor and debug the algorithmic social contracts between machines and humans are required.

Even though all the steps to implement the SITL framework are known, the modern society is still not able to put it to action because of some barriers. One of these obstacles is represented by the cultural divide between engineering and humans. Legislators and ethicists are well able to reveal moral hazard situations and to identify ways in which constitutional rights may be violated (Castelfranchi, 2000). However, it is not as straightforward to code this knowledge through engineering, resulting in difficulties in infusing AI with this information. Another obstacle to the SITL framework's implementation is represented by the existence of negative externalities resulting from algorithms. Quantifying the externalities is not an easy task, especially when the externalities are the consequence of various events. Moreover, once the externalities have been identified the following step is to identify the tradeoffs. As for externalities, quantifying tradeoffs is a challenging task, especially in complex economic systems where tradeoffs are often unintended consequences of algorithm design (Rahwan, 2017). The task of identifying tradeoffs is also made harder as algorithms become able to learn from experience, thus leading to shifts in the tradeoffs being made beyond what is initially intended by the designers. Similarly, quantifying the behavior of systems in such a way that is able to be understood by ethicists is a hard task (*ibid*). The difficulty in quantifying AI behaviors makes it complicated to scrutinize the behavior of algorithms against set expectations.

Simply stated there is an increased need to encompass ethical considerations into AI programming, but at the same time society does not have the right tools to do so, thus harming the potential expansion and acceptance of AI technologies.

3.3.2. Safety Concerns

An accident in machine learning can be defined as a situation in which the human designer has an objective in mind but the results produced by the system result in harmful and unexpected outcomes. The errors arising in AI can be classified according to the stage in which the process went wrong. The first thing that can go wrong is the designer's faulty specification of the formal objective functions, leading to harmful outcomes. Negative side effects and rewards hacking are two causes of incomplete objective functions. The former arises when the designer specifies an objective function which focuses on developing the objective in a given environment, but ignores the other possible aspects of broader environments (Amodei, et al., 2016). Reward hacking instead is manifested when the designer admits an easy solution that maximizes the function but which compromises the initial intent (*ibid*). A second cause of faulty specification of the objective function occurs when the designer knows the correct objective function and knows how to evaluate it, but it is too expensive to do so, leading to harmful behavior caused by inaccurate extrapolations of limited samples. This phenomenon is referred to as scalable oversight (Amodei et al., 2016). Lastly, the designer might specify the correct objective function to reflect his objective, but negative consequences arise from poor decisions from incomplete data. In order to ensure that that reinforcement learning agents' exploratory actions do not lead to negative consequences that damage the long-term value of exploration a method known as safe exploration is implemented. A valid alternative to solve the problem is robustness to distributional shifts, which instead focuses on how to prevent machine learning systems from making bad decisions when facing inputs that are different from those practiced during training.

For AI agents operating in large environments presenting many variables, an objective function which focuses on one aspect only of the environment might lead to indifference with respect to all the other aspects (Abadi et al., 2016). Hence, an agent following this type of function might cause disruption to the broader environment, and give rise to negative side effects. A complete objective function should be formalized as "*perform task X subject to common-sense constraints in the environment*" (Amodei, et al., 2016, p.4) in order to avoid side effects as much as possible. Although side effects might be positive, the majority of times they are negative as they tend to disrupt the surrounding environment. In order to avoid them, multiple approaches can be adopted. The first approach consists in defining an impact regularizer. A regularizer is set to penalize changes to the environment, thus leading the artificial agent to perform the task with minimal side effects. The challenge however is to formalize the change. A basic approach would be that of penalizing the difference between the current state and the initial state. However, the problem is that the agent will resist all types of changes in the environment, including natural human evolution. A more sophisticated approach instead involves comparing the future state of the agent's policy to the future state under a potential policy (Amodei, et al., 2016). This approach attempts to rule out changes that naturally occur while leaving reliable only the changes caused by the agent. If these two approaches do not work out, a more flexible approach would be that of controlling the impact of the regulizer through extensive training, however this would require transferring learning. Because of this, it would be beneficial to separate the side effect component from the task component as most of the times the side effects are more similar than the main goal is. A different approach to avoid negative side effects is that of penalizing influence, meaning that the designer prefers the machine agent not to find itself in positions in that are prone to generating side effects (Amodei, et al., 2016).

The next issue with safety is reward hacking. Formal rewards reflect the designer's attempt to formalize their intent so that it can be performed by machines. However, finding an easy solution at the expense of the initial intent might be a complex problem. Reward hacking can be the consequence of multiple events, including feedback loops, abstract rewards and partially observed goals. In order to avoid this problem, there are multiple machine-based approaches. Some of them include adversarial reward function, adversarial blinding and counterexample resistance (Amodei, et al., 2016). The first method, adversarial reward functions, is adopted when ML systems have an oppositional relationship with the reward function. In normal settings the agent is powerful, while the reward function is static, hence it is incapable of responding to the system's attempts at hacking it. However, if the reward function were its own agent, it would be able to respond to the system making it much harder to fool. The reward agent might try and find scenarios that the ML agent perceives as high reward, but that are perceived as low reward by humans (Goodfellow et al., 2014), thus fooling the ML agent. To make this model work, however, it is fundamental for the reward function to be more powerful than the ML agent. The second possible method to avoid reward hacking is adversarial blinding, which involves the use of adversarial techniques to blind a model (Ajakan et al., 2014). By using this technique an agent can be made incapable of underestimating parts of the environment. More precisely, it can prevent an agent to understand the way in which rewards are generated, thus making it very difficult to hack. Another valid method to avoid reward hacking when abstract rewards are used, is counterexample resistance. With abstract rewards the human agent might be worried that the learned components of the system will be susceptible to adversarial counterexamples and to resist them adversarial training can be used (Goodfellow et al., 2014).

In order to train machine agents, a complete oversight of the situations and of the possible outcomes would be preferred. However, because of time and cost constraints designers have to rely on cheaper approximations (Amodei et al., 2016). The divergence between the complete function and the cheaper one might lead to unintended side effects and scalable oversight, but solutions to this problem can be found. One approach to solve this problem is semi-supervised reinforcement learning. In this framework a baseline performance for the agent can be designed by ignoring unlabeled episodes and by identifying proxies that predict rewards. Other approaches to solve scalable oversight feature supervised reward learning, a model in which rewards are predicted on a per-episode basis by taking into account lower confidence in estimated versus known rewards (Dewey, 2014); and unsupervised model learning, in which systems use the observed changes of the unlabeled episodes to improve the model's quality. Alternative approaches to scalable oversight include distant supervision and hierarchical reinforcement learning. The former method aims at providing hints to guide the agent to the correct evaluation instead of providing small fractions of the system's decisions. By providing supervision the agent will be less likely to commit undesirable actions. The second method, hierarchical reinforcement learning (Dayan & Hinton, 1993), requires top-level agents to make highly abstract actions which are later delegated by sub-agents favoring a synthetic reward signal that represents the correct completion of the action.

The last safety issue concerns exploration since all autonomous agents engage at least once in exploration. Exploration however can be dangerous as it involves taking actions that the agent does not understand well, which might cause multiple problems. These problems can be avoided by hard-coding the avoidance of calamitous behaviors, but this approach is only functional when a small number of things can go wrong as in more complex domains it becomes increasingly hard to identify all the possible catastrophic events. Research is increasingly focusing on this topic and it has identified different possible solutions. One branch of literature suggests a framework called risk-sensitive performance criteria, which suggests to change the optimization criteria from an expected total reward to other objectives that are more prone to prevent catastrophic events (García & Fernández, 2015). This approach includes optimizing worst-case performance and ensuring that the probability that bad performance will happen is small. Other research suggests to start exploration in simulated environments where catastrophes are less likely to happen. Later, exploration in the real world will be required, but exploring first in simulated environments will allow the designer and the agent to learn about the possible dangers, thus allowing for safe exploration in the real world. Another possibility would be to have human agents check for unsafe actions. However, this solution might not be entirely practical as a human agent would have to perform too many exploratory actions for human oversight (Amodei et al.,2016). Another concern for this method is time as a human agent will take more time than a machine agent would. Furthermore, the human agent would need to be able to distinguish between genuinely risky explorative actions and safe ones.

Simply stated, safety concerns are quite widespread in the modern society, but there are a variety of approaches to building robust and safe machine learning systems.

3.3.3. Cognitive Conflicts

Since machine agents' objectives are determined by the algorithm's designer, machine objectives can result to be perfectly aligned with those of the organization. Hence, there appears to be no machine agency problem. However, this is not entirely true as objectives are not always fully specifiable by designers. Because of this, when examining the AI context, cognitive conflicts may arise within the organizations since the cognitive frames of the machines are even less transparent than those of humans.

Marengo and Pasquali (2012) present a model that enables the principal to decide the most effective force between political and cognitive power. Their analysis revealed that when learning is not at stake, the choice between organizational structure and managerial structure is indifferent as they are substitutes. Diverging views between principal and agents can be bridged by careful organizational design, and managerial intervention can be used as secondary device. On the other hand, when learning is at stake, organizational structure and managerial intervention may complement each other, but have to be finetuned according to the environment's complexity. However, the attempt of principals at changing the agent's point of view is a hard task as the principal is not fully aware of what the agent thinks and of what is the right way to make him change his mind. The problem becomes increasingly difficult when the agent is not a human but a machine agent. Principals may not be aware of the algorithm that drives the agent, as to understand it they would need to have technical knowledge. In addition, since algorithms are programmed beforehand and machine cognitive frames are not clear, changing the machine's agent way of thinking would be time and money consuming and would imply a change in the organization's objective as the algorithms should be programmed to fulfill the organization's goals.

3.3.4. Robot Takeover

Another widespread concern is that of machines taking over humanity as it happens in many well-known movies such as *I, Robot* or *Terminator*. During the seminar *"Artificial Intelligence, Machine Leaning and it's place in our Society"* (Baguley & Patel, 2020) organized by VMware, a global leader in cloud and digital workspace technology, Joe Baguley, Vice President and CTO, was asked what he thought about this concern.

In order to illustrate his view on the problem, Baguley started by presenting Asimov's three rules of robotics, which, even if first developed in Asimov's short story collection *Runaround* (1950), play an important role in the field of robotics and AI. Asimov's rules are as follows

1. *A robot may not injure a human or, through inaction, allow a human being to come to harm.*

2. *A robot must obey the orders given it by humans except when such orders would conflict with the First Law.*

3. *A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.*

These rules cover all the basic things that machines should and should not do, however, they are valuable only if the humans that deploy and use the machine are doing so in an ethical way.

Putin famously said "*whoever becomes the leader in this sphere* [artificial intelligence] *will become the ruler of the world*" (Allen, 2017), a statement that might be considered as a sweet-and-sour statement if considered in the light of AI development and in all the ways that its increasing skills might be used against humans. Because of the negative side-effects that might be caused by the inappropriate usage of AI, Baguley firmly states that he is not scared of machines takeovers, and in fact, is excited that in many fields they already are. What he is concerned about is the ability of humans to inappropriately use machines against each other. Therefore, he argues that AI should be legislated and that AI should be controlled. However, creating legislation for something that is so ill-defined is not an easy task, thus shifting the takeover concern to an entirely different legislative dilemma (Emmen, 2015).

3.4. Machine Behavior

Three scales of inquiry for machine behavior exist, namely individual machines, collectives of machines and groups of machines which are part of a social environment along with human individuals in hybrid or heterogeneous systems (Shirado & Christakis, 2017). The individual scale of inquiry focuses on the study of the algorithm, the collective inquiry focuses on the interaction between hybrid humans and machines, and the collective inquiry studies the interactions between machines and humans.

3.4.1. Individual and Collective Machine Behavior

Individual machine behavior studies focus on specific algorithms and closely examine the properties of the machines that derive from their source code or design. In order to study individual machine behavior, two general approaches are often adopted. The first approach focuses on determining the behaviors of machine agents that are using a within-machine approach by comparing the machine's conduct in different settings (Rahwan et al., 2019). The second approach instead, examines how multiple machines act in the same context (*ibid*).

By adopting the first approach, a within-machine one, it is possible to investigate whether there are constants that characterize within-machine behavior of particular AI agents across multiple contexts. Moreover, it allows to examine how machine behavior changes over time and which environmental factors lead machines to act in certain ways. As some AI agents might exhibit only certain behaviors because their algorithm is generated and trained on particular data (Feldman, et al., 2015), the analysis in this case will revolve around understanding how the machine agent will behave if it is presented with an evaluation of data which is highly different than the one it is trained on. Other studies that have used this approach focus on studying individual robotic recovery behaviors (Bongard, Zykov & Lipson, 2006), on understanding the utility of using technics from psychology in understanding algorithmic behavior (Leibo et al., 2018), and on examining bot-specific characteristics such as the ones designed to influence human individuals (Subrahmanian et al., 2016).

The second approach instead focuses on examining the behaviors of different machines. A field in which this approach is widely used is that of marketing, as this approach allows to observe and understand the different advertising behaviors of AI agents (Carrascosa et al., 2015). In addition, this approach is increasingly being used to study the different behavior of autonomous vehicles in overtaking (Giusti et al., 2016).

Differently from individual machine behavior, collective machine behavior focuses on the interactive and systemwide behaviors of groups of machine agents. In some contexts, individual machine behavior may not make much sense, hence a collective study will allow to gain a more complete picture.

3.4.2. Human and Machine Behavior as Complementary Forces

The introduction of machines within society can alter human believes and behaviors. AI is increasingly used in the automation of industrial processes (Bainbridge, 1983), hence making humans start to question its reliability in developing a genuine picture. Individuals, for example, might ask themselves if news-filtering algorithms are altering public opinion by only showing part of the news. It is because of this that it is extremely important to investigate if small errors in algorithms or in data can produce society-wide side effects and how AI might change the quality of life (Lorenz, Weiss & Hirche, 2016) and influence human development (Westlund et al., 2017). In addition to these concerns many more exist, such as the way in which governments are using AI to alter democracy and transparency (Rahwan et al., 2019) or even the outcomes of elections (Lazer, et al., 2018), as some believe is the case with Trump's election. These concerns are rising because human are not capable of understanding how machines are influencing policy and welfare. Many of these questions are being researched and answered by scholar such as Brynjolfsson and Mitchell (2017) and continue to be of primary importance as it is fundamental to understand how human systems can be altered when machine agents are introduced in everyday life and tasks.

Machine behavior can affect the human one, however the opposite is also true. In fact, also humans can module and create the behavior of intelligent machines through the engineering of AI system and training on data that is generated daily by individuals (Rahwan et al., 2019). Human agents are always in charge of deciding which feedback and which algorithms to use (Thomaz & Breazeal, 2008), along with which data is more appropriate. Hence, human decisions can directly shape and change machine behavior. Moreover, a central topic for studies involves understanding whether training data is directly responsible for shaping particular behaviors in machine, or whether machine behavior is a mix between algorithms and data (Rahwan et al., 2019).

3.4.3. Human Machine Co-Behavior

The majority of AI systems work in environments in which human and machines co-exist in multifaceted hybrid systems (Gray & Wegner, 2012). Many studies in this field focus on examining typical behaviors in human-machine interactions such as cooperation, coordination and competition (Kramer, Guillory & Hancock, 2014), while another field of research focuses on the automation of human labor (Jackson, 2019). In this field there are two different views. The first one is that machines can enhance human efficiency, while the other is that machine will replace humans. This naturally leads to the question of whether it will be more plausible for machines to enhance or replace humans.

Because of increasing concerns, scholar started to examine human-machine interactions in laboratories and observed that simple bots can increase human coordination (Shirado & Christakis, 2017) and that machines can cooperate with humans as much as another human could (Crandall et al., 2018). However, there is still the need for studies to examine longer-run dynamics of hybrid systems with an increased focus on human-societal interactions and how they will change with the introduction of machines (Aharony et al., 2011).

Chapter IV

Case Study: Artificial Intelligence And The Recruitment Process

The recruitment industry is facing a significant issue as the traditional hiring process where resumes and interviews are used has found to be ineffective because of the growth of internet which could yield unreliable results (Sullivan, 2017). To overcome the obstacles related to the process, many companies have started to incorporate recruitment software within their process in order to scan more candidates faster and to acquire more talent. Some recruitment software also adopt AI within their technology, but being a novel implementation not much research has been performed to understand whether or not artificial intelligence (AI) is the best tool to power the software. Moreover, AI has been defined as "*game-changing for HR*" (May, 2016) which could mean it might have both positive and negative implications.

This study aims at understanding the state of AI that is available to companies and consequently the use that companies can make of AI within Human Resource Management Services, with a particular focus on recruitment. The study's objective is also to understand AI's practical implications on decision-making by providing an encompassing overview on the topic by delivering insights on both the benefits and the detrimental effects that AI can have on the recruitment process and on the applicants.

4.1. Artificial Intelligence and Recruiting

Artificial intelligence is becoming increasingly used within the recruitment process. Historically, recruiters have selected candidates based on a limited talent pool available to them and spent 13 hours per week (Entelo, 2018), on average, on sourcing candidates for a job position as there is just a limited number of resumes and applications that recruiters can look at. Since AI is able to rapidly scroll through millions of resumes, it can enable recruiters to reduce the time taken to distinguish high-potential candidates from low-potential ones. Moreover, AI can allow recruiters to source candidates from a wider talent pool, such as social media platforms and online career posts, as it enables to source information more rapidly and to examine all the available information faster.

Another useful application of AI within the recruitment process is the ability to efficiently identify past candidates, who make up 70% (LinkedIn, 2015) of the global workforce. Because of this, ignoring the potential of this group can affect recruiting effectiveness. In addition, 98% (Entelo, 2018) of talent teams claim that passive candidates are a precious source of talent and that they are a 120% (Jobvite, 2014) more likely to make a stronger impact within the company.

Former applicants and employees represent another undervalued category that could be highly beneficial and which could be worth to direct the recruitment efforts towards. Companies have a database of thousands of resumes that represent lucrative talent and AI can rapidly sift through the company's database and identify former applicants which are qualified for the current job offerings. Moreover, since former applicants and employees have already expressed interest in the company, the recruitment process would be sped up.

In addition, another potential application of AI within the recruitment process would be to reduce bias. The preferences of talent teams to prioritize applicants who share similar beliefs and backgrounds leads them to reduce the diversity within the considered talent pool. As companies with a diversified workforce tend to perform financially better than companies that do not, reducing bias would be a key benefit for companies. Along with reducing bias, AI is also expected to increase candidates' engagement through systems such as chatbots.

4.2. Literature Review

Artificial intelligence is a fast-moving technology because of its continuous evolution and, because of this, literature on the topic is limited. Moreover, the use of AI within the Human Resources (HR) process is rarely covered by academic studies and is more analyzed within professional reports.

In order to develop a qualitative analysis on the topic, thematic analysis was used. The themes were taken from arguments presented in literature and touch upon the influence that AI can have on recruitment and the impact that is has on candidates and employees.

The themes are as follows:

1. Limitations and opportunities to overcome them
2. Evolution of the recruiter's role
3. Bias reduction

4.2.2 Limitations and opportunities to overcome them

Artificial intelligence is a topic that is surrounded by a variety of limitations which range from technical to ethical ones. Advances in AI have the potential to automate human activities, however a survey conducted by Smith and Anderson in 2017 found that people more frequently reflect worries and concerns when asked about automated technologies rather than positive feelings. Raviprolu (2017) directly considers the role of AI within HR and acknowledges a number of barriers that need to be demolished in order to adopt it. The first barrier is the amount of accurate data that is required to successfully train a machine learning (ML) algorithm. Moreover, Yano (2017) emphasizes the importance of having accurate data by stating that AI is nothing more than a black box if data is not applied. Another concern that is raised for the application of AI is that since datasets are becoming increasingly protected, there is a higher risk that companies will train their data on cheaper, less-valid datasets (Campolo, et al., 2017). In addition, even supporters of AI within HR, such as Wisskirchen et al. (2017), argue that even the most advanced AI technologies can make mistakes. Another increasingly important challenge is GDPR as it can limit the access to data. There is an estimated 80% (Chaker, 2018) of recruiting companies that are not being compliant with the GDPR rules, which will have a big impact on the industry. In addition, Raviprolu (2017) suggests that AI has not reached strong communication abilities, hence according to Tandon et al. (2017), who agree with this view, AI agents cannot result in the full automation of the entire process. Another concern raised form the literature is that automated recruitment can have negative aspects on the process as a whole as it can impersonalize the process on both sides (Okolie et al., 2017). In addition, safety concerns are also frequent as AI, as of today, is only able to replicate the human decision-making process (Frey et al., 2013), and scholar such as Parnas (2017) argue that imitating humans might result in programs that are dangerous and not trustworthy.

Despite the fact that many concerns emerge, a study conducted by Korn Ferry (2018) reveals that even though professionals have less trust in AI than in human recruiters, nearly 72% argues that AI should be used within the recruitment process.

4.2.3. Evolution of Recruiter's Role

Not much literature directly addresses how the recruiter's role is affected by the adoption of AI as the literature focuses more on the workforce in general.

For many companies the first area in which AI is introduced is talent acquisition since companies are able to witness the results immediately. Hence, the recruiter's role is the first to experience change when AI is adopted. According to the Future of Jobs Report 2018 from the World Economic Forum (2018) by 2022, 75 million jobs will be displaced as AI takes over the routine aspects of work. However, always according to the same survey, these jobs will be substituted by different versions of the original jobs as 133 million new roles will be created and will require skills in emotional and technical intelligence. Thorpe (2018) also estimates that AI will create more jobs than it will eliminate. Because of the increasingly frequent adoption of AI, it is important to have an artificial intelligence ready workforce which can be obtained by prioritizing the upskilling of non-AI workers (Meister, 2019). With respect to the HR field, Meister (2019) also argues that it is fundamental to change performance management and to develop the skills needed in order to allow HR roles to understand how to use AI across the employee life cycle. Manyika et al. (2017) support this claim and also suggest that as many as 375 million workers worldwide might need to change job and learn new skills as 60% of occupations will have at least one-third of their work activities become automated.

4.2.4. Bias Reduction

Campolo et al. (2017) deeply analyze the social and the economic impacts of AI and observe that ML programs use existing datapoints in order to make decisions. If the used datapoint is biased, this could result in transferring cultural, gender and other types of discrimination contained in the dataset directly to the algorithm. In contrast, AI-supporters Wisskirchen et al. (2017) suggest that using AI in recruitment will remove all types of bias as it would allow to purely focus on data without emotion and sympathies interfering. According to Randstad (2018), approximately 70% of human capital and C-suite leaders suggest that technology is improving their hiring decisions. Hiring tools, such as video analysis, allow the recruiter to make faster decision and help to further diversify by eliminating unconscious human biases to which recruiters might be subject to (PwC, 2017).

However, algorithms are not completely bias-free and the literature also outlines how discrimination could arise when using an AI algorithm. Florentine (2016) explains that if a machine learning algorithm finds a statistical relationship between high performance in sales roles and quarterbacks in American Football, based on data derived from leadership skills, mental skills and decision-making, it could result in the algorithm eliminating all female applicants as females are not present within American Football teams. A similar situation to this one is observed in Amazon's recruitment algorithm which excluded females as the models were trained by vetting applications over a 10-year period (Dastin, 2018). Within the considered period most of the applications came from men, therefore the algorithm understood that male candidates were preferable to female ones.

4.3. Conceptual Framework

As portrayed within the literature review, three main themes have emerged when analyzing the application of AI within recruiting. The outlined topics are best characterized as: Limitations and opportunities to overcome them; Evolution of the recruiter's role; and Bias reduction.

In addition, since academic literature deeply focuses on the technicalities of AI, the research will have little considerations on how the technology works, but will focus on its application, and the impact that it has on the recruiting process. Moreover, the literature largely focuses on the employer's experience with AI over that of the candidates, therefore the research will also aim at providing insights on how the candidate can be impacted by the adoption of AI.

4.4. Research Aim and Objectives

The aim of this study is to perform a qualitative research to understand the current state of the AI technologies and to explore the extent to which, as of today, they are used within the recruitment process and the changes they can bring to the recruiter's role. Moreover, the study aims at understanding which factors are promoting or impeding the adoption of AI from recruiting companies.

The cognitive objectives of the study are as follows:

1. Provide insight on how companies have overcome the concerns and successfully implemented AI (Theme 1)
2. Observe which concerns are raised by those companies that do not use AI within their recruitment software (Theme 1)
3. Investigate whether AI changes the recruiter's role once adopted (Theme 2)
4. Investigate if AI can be a valid tool for bias reduction (Theme 3)

4.5. Methodology

In order to conduct the study, 6 recruitment software providers were interviewed, both users and non-users of AI technologies, in order to provide a full spectrum of opinions on the technology. Participants were subjected to semi-structured, shorter case study type interviews, in order for the interviewees to provide insightful commentary on the topic while allowing for deviations in those areas that were brought up but not yet researched on the topic of AI and recruitment. Interviews have been used as they are a useful tool to gather insights on the participants' perspective and each respondent gave permission to be featured in this research.

Semi-structured interviews can present negative sides. The first one can be the emergence of the interviewer's bias because of the lack of structure along with the possible problem of inaccurate articulation of the questions (Yin, 2018). Because of this, extreme care has been adopted in order not to introduce bias before a responded could answer (Bradburn et al, 2004). The questions posed to the participants were structured in such a way that the respondents could answer based on their expertise on the topic. In addition, to further counter bias, a final review of the questions and answers was performed and any answer that might have been a consequence of the question's phrasing was removed from the findings of the study.

4.6. Participants

The target group for the interviews were 6 companies providing recruitment software that have significantly contributed to change the recruitment processes. Because of their core business activity, the selected companies could provide knowledgeable insight and have considerable experience on the topic of AI adoption within the recruiting process and its potential benefits or pitfalls. The participants have been selected based on Capterra's (2020) list of recruitment software along with a through Google search to find the most relevant organic results when typing the phrase *"recruiting software Italia"*, which translates to recruiting software Italy, in order to be able to find the most relevant software for the Italian market. Moreover, the chosen companies and their software reflect the features that the current recruiting systems can offer to the companies that adopt them.

The interviews with the representatives of Altamira, Allibo, Cornerstone OnDemand and In-Recruiting were conducted in Italian and later translated to English.

The interviewed companies and their software are presented in the next sections in alphabetical order.

4.6.1. Allibo

Alliance Software Srl is a Milan based Italian company founded in 2010. Their core activity is that of developing vertical software for business organizations and the HR field. Alliance Software is currently the software provider for over 300 companies including Decathlon, Wind, Carrefour, and many more in a wide variety of sectors.

The customized software solution offered by Alliance Software is Allibo. Allibo, does not use AI within the platform and it is a GDPR compliant software designed with the mission to provide the best tools to help companies build the most efficient teams possible. Allibo offers different solutions that are designed to simplify the management of the HR process.

Allibo Recruit is a market leader in Italy. This software is an ATS (Applicant Tracking System) that allows to manage the recruitment process by simplifying the management of job-positioning on the company's websites, intranet or social media along with all the major job-sites both in Italy and aboard. Allibo Recruit also allows to collect and gather all the resumes received through web, mobile and e-mail within a single private resume database. According to Allibo, this is a valuable tool to speed up the screening process, to assess the candidates and to make video-interviews both live and recorded. Moreover, through query-reporting tools and data extraction, Allibo's clients are able to analyze the operative trend during different periods of time, thus helping them to track the performed actives and suggesting which phases of the recruitment process can be improved. In addition, this system will also increase the visibility of the posted job-offers, thus drawing more candidates to apply for the job offer. By using this system, the recruiter will then have more time to engage in the most relevant activities such as those requiring human empathy and human valuation.

In addition to the recruitment software, Allibo's platform also offers other tools such as Allibo Employees, which allows to manage already-hired employees; Allibo Train Up!, which is a training management system to manage personnel formation; Allibo Perf, that allows to evaluate employee's performance and Allibo Skill Iperwin, that generates an analysis of the skills that are present within the company. Although the beforementioned tools are very useful, they are beyond the scope of this research.

4.6.2. Altamira

Altamira S.r.l. was established in 1999 and is an HR software company. Altamira's HRM platform was designed to meet the HR department's needs, ranging from recruitment to employee management and from training to evaluation. Altamira has a wide range of customers coming from different sectors and some of them include Calzedonia, AS Roma and Lamborghini.

Altamira Recruiting is a GDPR compliant software, that does not make use of AI within the platform, and which aims to simplify the recruitment process while improving employer branding. At Altamira they believe the software needs to reflect the way the client's organization works, in fact their solution is fully customizable. Moreover, email and other recruiting tools make it hard to keep track and classify resumes, but Altamira Recruiting is able to perform full text searching thus allowing to create and choose the job positions required at that moment by the company. Moreover, the software allows to automatically match candidates to job requirements. Once the resumes are collected, Altamira's software ranks the candidates according to the specified criteria, thus allowing the recruiter to focus on those candidates that are potentially the best fit for the job-position. Moreover, in order to avoid spending time on analyzing less relevant resumes, Altamira's system offers a tool that automatically ads, or removes, resumes based on the requirements that candidates need to have to participate to the recruitment process. Profiles that meet the requirements will remain, while those that are not in line with the criteria are automatically removed. The system also helps to identify the best candidates by assigning to the resumes a score from 0 to 3 stars depending on their compliance with the criteria. Another advantage that can stem from using Altamira's software is superior employer branding as it makes it easier for candidates to find and gather information on the vacant position within the company. Through the use of Altamira Recruiting it is also possible to digitalize the onboarding process. The system automatically performs onboarding routine steps such as document exchange, reading of policies and assignment of preliminary tasks to the new hires.

Altamira's tools are not limited to Altamira Recruiting. The platform also offers Altamira Employees, which allows employers to organize already hired employees by organizing data and HR processes; Altamira Learning, which allows to simplify the employees' training process thus simplifying their work; Altamira Performance, which allows to evaluate employees' performance; Altamira Leave Management, to simplify the management of leaves; and Altamira Attendance to monitor attendance. The presented applications however go beyond the scope of the study.

4.6.3. Bryq

Bryq is a startup that helps companies to make better decisions during the hiring process through their scientific based hiring pre-assessment tool. This software measures cognitive abilities, leadership potential and the candidates' psychometric traits in order to help the recruiter and the candidates find the best fit for them. Bryq's goal is to facilitate the hiring process while finding the right fit. From the employers' point of view, Bryq's software can be helpful as it facilitates and speeds up the process of matching candidates to job positions while uncovering potential hidden talents that might get overlooked in hectic hiring processes. On the other hand, candidates can identify their ideal career path with ease, thus increasing their long-term job satisfaction. Some of Bryq's clients include Trafiko and Upstream. Moreover, their tool can be integrated with the most distributed ATS software such as Greenhouse and SmartRecruiters, and can also be adapted to satisfy customized integrated solutions.

"*Hire talents, not resumes*" is a key component of Bryq's perception of the hiring process and in order to be faithful to their aim, Bryq provides a bias-free tool. Being bias-free is a legal requirement and it also helps to create a diverse team capable of better contributing to create more profitable companies, and Bryq's software is able to objectively assess candidates, through the use of data, and match them to the company and job requirements in order to allow the recruiter to hire the best candidate without any bias. In order to do so, Bryq follows four simple steps: objective definition of job requirements, which can be customized to satisfy different companies' needs; automatic screening of candidates, by inviting candidates to take an assessment; shortlisting candidates based on fit, candidates are ranked based on the assessment results without ethnicity, gender, education age and other variables impacting the result; and lastly performing objective, which includes standardized behavioral interviews as Bryq provides an interviewing guide, based on the candidate's data, that allows to ask targeted behavioral questions. Moreover, what distinguishes Bryq from other pre-hire assessment tools is that they predominantly base their tool on scientifically derived research rather than on technological innovations. Their solution is based on proven industrial and organizational psychology frameworks and follows the Equal Employment Opportunity Commission's (EEOC) regulations and is able to measure 4 different cognitive skills (numerical, verbal, problem solving and attention to detail), along with 16 personality traits.

In addition to pre-employment assessment, Bryq also focuses on fostering internal mobility, thus allowing internal candidates to be assessed with the same unbiased treatment used for external candidates in order to pursue new careers within the company and take advantage of their strengths.

4.6.4 Cornerstone

Cornerstone OnDemand was founded in 1999 and it is a listed company with the aim of organizing, selecting and simplifying employee training. Cornerstone was born with a mission in mind, that of improving access to education worldwide thanks to online learning. Cornerstone OnDemand has a wide range of products and its Cornerstone Unified Talent Management application allows to satisfy the client's requirement when it comes to talent search, without using AI. Cornerstone OnDemand collaborates with companies such as Walgreens, Nestlé and Deutsche Post DHL. Moreover, thanks to the SaS (Software as a Service) solutions, users of the software are always able to use the most updated version of the software and are able to reduce IT and maintenance costs.

Cornerstone's Recruitment suite is designed in order provide the appropriate social recruitment tools to find the right talent. By adopting Cornerstone's software, employers will be able to attract suitable candidates by showcasing the employer brand and culture on an interactive media that allows to attract the best-fit candidates. Moreover, it will be easier to convert job seekers into applicants through intuitive applications processes that allow candidates to apply through social profiles and through branching questions in order to ensure the application process is relevant and seamless. In addition, the software enables to select best fit applicants by enlarging the talent pool as it extends the search to internal and external candidates, along with suggesting passive candidates who match the criteria. The recruitment software also offers features that go beyond applications as it helps with onboarding. The software is designed in order to provide personalized new hire onboarding portals that deliver interactive content to tailor resources to specific candidates while centralizing the onboarding activity. Moreover, the software makes data-driven talent decision based on recruitment metrics as it centralizes data and gives a comprehensive overview of the recruitment process thus enabling more informed decisions. In addition, by centralizing the candidate and hiring data, the process can be easily examined to reveal the areas that need improvements thus allowing the company to correct for those bottlenecks and make smarter decisions.

In addition to the Recruiting Suite, Cornerstone also offers a Learning Suite, that allows to personalize the employees' development; a Performance Suite, that enables companies to build high-achieving teams; an HR Suite, that assists in making data-driven decisions; and a Content Anytime Feature, which provides people with E-learning content. These features however lie outside the scope of this study.

4.6.5. In-Recruiting

Intervieweb S.r.l is a HR Tech company specialized in developing and selling software designed to simplify the HR and recruitment process. Intervieweb's core recruiting software is In-recruiting, one of the major ATS present on the market which is distributed through Cloud/SaaS and which allows to manage the recruiting process from start to finish.

The first version of the In-recruitment software was released in 2009 by a group of HR and IT consultants linked by a simple objective, that of simplifying the recruiter's daily job while offering him the best recruitment solution on the market. Moreover, according to the founders, recruiting requires performing many repetitive actions that do not provide high value if compared to the time it takes to perform them. Hence, performing these tasks without any use of technology is harming the recruiter's main job which is that of finding the best candidate in the shortest amount of time. Since 2009 In-recruiting continued to grow and today it is internationally available. Moreover, the tool is compatible with a variety of software, therefore making it easy for their clients to start using this service. The software is used by companies in different sectors ranging from insurance companies to production ones. Some of its clients are McDonalds, DHL and Lidl.

In-recruiting aims at finding faster and easier solutions for the recruiters, in fact In-recruitment developed a community of HR professionals that share feedback along with new difficulties they encounter so that the software can be adapted to solve the new needs. Thanks to the community's point of view, In-recruitment is able to upgrade its software 4 times a year on average in order to keep up with its customers.

Moreover, In-Recruitment makes use of AI through Inda (INtelligent Data Analysis), which is the AI technology that allows to automate the repetitive tasks that the recruiters need to perform and which enables to make the recruitment and acquisition processes faster. Through this technology, In-Recruiting is able to use technologies such as world embedding, which uses the semantic field of the words; computer vision, which is the ability to transform a resume's picture into a set of words; and effective computing sentiment analysis, which allows to understand the feelings expressed by a candidate both in written message and in pictures by analyzing the verbal and body language used in order to understand the candidate's personality through a screen. All of these tools are helpful to speed-up the process. Moreover, Inda also aims at developing a community by publishing blog articles about the new developments of AI in the recruitment field and also share their opinion on the future that AI can lead to.

4.6.6. Manatal

Manatal is an AI cloud-based recruitment software that aims to transform the recruitment process worldwide. Manatal's ATS, through its AI-powered platform, wants to make the entire process simpler while also keeping in mind time, automation, information and insight. Manatal is a recruitment software that enables the HR departments and recruitment agencies to source and hire in an effective way. The software was born by acknowledging the fact that the recruitment processes presents some issues, such as finding the right person at the right time, which is a challenge on a global scale. Moreover, Manatal addressed the difficulty of properly leveraging information and the sheer effort that is put in recruitment. Manatal's vision is that of creating an ATS that simplifies the recruitment process from souring the talent to onboarding and beyond.

The software offers a variety of tools to help within the HR field such as candidate sourcing, the applicant tracking systems itself and recommendations. In order to fulfill the first feature of the software, that of finding candidates, the system offers a variety of tools. By using Manatal, the company will be able to upload multiple resumes at the same time, and once uploaded Manatal will automatically create candidate's profiles based on the resume's information. Moreover, the software allows to create a job position and to simply share it across the largest job portals such as LinkedIn. Also in this case, the profiles of the candidates applying through the job portals will be automatically created, uploaded to Mantal and matched with the job. Moreover, every candidate applying for a job position will be scored based on his capabilities, work experience, skills and talents in order to speed up the recruitment process.

The second feature is that of the ATS itself, which is designed to simplify the process. The tools contained within this feature make the management of resumes and candidate's information easier as it helps to keep track of the candidate's profiles over time. Moreover, it also allows to monitor the state of new hires and placements through their work experience within the company by keeping track of starting dates, end of employment dates, candidates profiles and more.

Another relevant feature of Mantal's software are the recommendations. Finding the perfect candidate to match a job position might take a long time, but through its AI-powered recommendation engine, Manatal can speed up the process. This feature enables the system to suggest the most suitable jobs for the candidates and at the same time recommends the most suitable candidates for the jobs. In addition, the recommendation systems can compare candidates' skills, position, education and other job requirements and then rank the applicants based on a criterion selected by the company adopting the software. The higher the ranking, the more suitable an applicant is for the job position. Another tool is the Boolean search which helps to find candidates that possess specific skills. In this way, recruiters are able to identify candidates with unique experiences or advanced proficiency.

4.7. Findings and Results

Interviews were conducted with 6 experts via phone and meeting platforms. Stefano Michetti, CEO at Allibo, having directly experienced the benefits and disadvantages of adopting AI. Emanuele Rifaldi, Account Manager at Altamira, having contact with many professionals within the HR field and having participated at events with a focus on AI. Markellos Diorinos, CEO and Co-Founder at Bryq, having knowledge in providing an efficient recruiting software without the aid of AI. Francesca De Cristofaro, Senior Solution Consultant EMEA at Cornerstone, having relevant insights on the benefits that AI has on the recruiting process. Matteo Cocciardo, CEO and Co-Founder at In-Recruiting, having know-how with respect to the implementation of AI within the recruitment process. Tim Peltier, Head of Marketing and Communication at Manatal, having insights on AI adopting software and how it is perceived by their customers.

To analyze the transcripts of the interviews, coding was used. The three themes of Limitations and opportunities to overcome them (Theme 1); Evolution of the recruiter's role (Theme 2); and Bias Reduction (Theme 3) that were highlighted in the literature review have been used indirectly to code the interviews. The codes have been inspired from the original themes and therefore have been linked to them. The coding process has enabled a more effective analysis of the insights and knowledge provided by the interviewees. For these interviews the considered codes and the linked themes are as follows:

CODE	DISCUSSION	THEME
Capability	Which parts of the recruitment process are performed by AI?	Theme 1
Benefits	What are the advantages of using AI?	Theme 1, Theme 2
Concerns	Have your customers experienced concerns with the use of AI?	Theme 1
Limitations	What are your concerns with respect to the adoption of AI?	Theme 1
Bias	Can an algorithm reduce bias?	Theme 3
Supervision	Does an algorithm require human supervision?	Theme 2, Theme 3
Recruiters	How did the recruiter's role change with the adoption of AI?	Theme 2, Theme 3
Data	Is there an adequate amount of available and credible data?	Theme 1; Theme 3

Moreover, a summary table of the interviews is provided below

Questions	Software					
	IN-RECRUITING	BRYQ	ALLIBO	MANATAL	ALTAMIRA	CORNERSTONE
Which tasks are perfomed by AI?	CV acquisition phase and canidates' pre-screening phase	None	Had previous experience but as of today none	Candidate screening; suggestion of candidates; scoring; data mapping	No use of AI within the platform	Automatic identification of skills within resumes; CV passing; identification of skills based on interests and experiences
Any prejudice from clients?	Yes, because of AI's probabilistic nature and because of other company's ineffective use of AI	N/A	N/A	Not until now	N/A	Doubts on data privacy and security
Is an algorithm prejudice-free?	Depends on its training as it requires unbiased datasets	No since it is trained on biased data	Depends on its training	Depends on training: the larger the dataset on which it is trained, the less likely for a small minority to influence the model	Depends on who trains it and on what datasets it is trained	Yes ,provided it is sophisticated enough and is trained on non-biased data
Can an algorithm can reduce human bias?	Absolutely but the algorithm needs to be trained correctly	No because it is trained on biased datasets	Depends on how the models are trained	Definitely, if trained on an appropriate dataset	Not always, as it depends on datasets	Can help recruiters to have a more objective view on the process and helps them to make objective decisions, provided it is well trained
Is human supervision required?	As of today, yes	AI cannot replace humans	Yes, so it defeats its purpose	Yes	Yes	Yes, and is desirable
Main benefits of using AI	Economic; image; time, information availability; helps to overcome the recruiter's lack of multidomain knowledge	Can help in some recruiting phases where there is substantial data processing to be done provided it is trained on a bias-free dataset and is explainable	No benefits within the hiring process, but can be used to analyze the information on the uploaded Ids and is successfully used by job-posting websited (Ex. LinkedIn)	Accuracy in identifying skills; more efficiency in talent acquisition; professional can work faster	Can help during the pre-screening phase to reduce time taken; can help to understand candidate's language skills	Increases employee engagement; provides a wider view; automates time consuming activities; speeds the process; increases the talent pool
Why is AI is not a good tool for bias correction?	N/A	Trained on biased datasets; no explainability; lack of scientific easerch	Output validity cannot be assessed; can create "dangerous" outcomes; can be used inappropriately	N/A	Cannot be developed in-house (most of the time); high costs and time consuming to adopt	N/A
Are there any limitations or concerns with AI?	There should not be any software that allows AI to take decisions alone at the moment	Trained on biased datasets; lacks explainability; lacks scientific foundation	Leads to hiring "clones"(ethic concern); insufficient quantity of data to tailor algorithms; GDPR concerns; difficult to understand if AI is working properly	Needs to be carefully trained	The market is not thirsty for AI; needs training otherwise is useless; explainability concerns; ethical concerns	N/A
What is missing to ensure a successful implementation of AI?	N/A	Data and science	Clear idea on behalf of the employers of what candidate they are looking for	N/A	Financial means; effective training	N/A
How did the recruiter's role change after AI?	Recruiters can now focus on value adding activities while AI takes care of low value adding ones	By either using AI or science-backed systems it would improve	Did not change	Recruiter is now able to focus more on human oriented tasks	Did not change	Recruters are more efficient and faster

4.7.1. Findings on Concerns Bridging

From the interviews many concerns have emerged with respect to the possibility of using AI within the recruitment process. A concern expressed by Michetti is that the existence of pre-trained cognitive services and algorithms is leading the market to speculate on the concept of AI without being sure of what its actual capabilities really are. In addition, since people are not fully aware of what AI really is at the moment, they *"fears losing their job or Terminator-like scenarios, but this is far from reality"* as expressed by Peltier, thus contributing to the speculation around the topic of AI. Furthermore, Rifaldi pointed out that often companies are only using AI as a naming and are not correctly training their AI, therefore making it unable to perform efficiently, an idea to which also Michetti agrees. Hence, since *"any tool is as good as its application*" if the AI is not trained well-enough to perform, its application is useless as explained by Diorinos.

Being training a crucial part to ensure the model's efficiency, the data on which the algorithm is trained plays a fundamental role. As stated by Diorinos *"the problem is not the AI* […] *the problem is how you train it.* […]. *How do you train a supervised learning machine with data that is supposed to be bias free?"* Moreover, Rifaldi says there can be ethical problems which are created by the company that is training the AI since if the manager of that company presents some kind of prejudice, he will transfer it to the algorithm.

Another concern linked to training was presented by Michetti and is *"a technical problem, data quantity"*. As illustrated by Michetti, what works for a company might not work for another, an idea also supported by Diorinos, hence an AI system would need a large quantity of data in order to be tailored with the requirements of a specific company. In order to be able to do this, as expressed by Michetti, AI would also need to be compliant with GDPR since companies will need to ask consensus so that their data can be used by third parties to train an AI system. Hence, the problem that can arise from GDPR policies is the limited number of candidates that will agree to the terms and conditions of data processing once informed that their data will be transferred to third parties, since if they do not agree, their data cannot be used to train the algorithm. However, this issue is already being considered by companies such as Cornerstone as explained by De Cristofaro who illustrates that once gathered *"all the customers' data is anonymous and ends up in a data lake."* In addition, if the dataset is not wide enough, it can cause additional problems. Diorinos points out that algorithms might seem to be efficient within confined sets, however as soon as the model breaks out of the limited datasets, there can be issues as it might not be able to operate correctly. Moreover, Cocciardo explains that *"some algorithms, unfortunately, can have prejudice that comes from historical data or human opinions."* In fact, all of the interviewed AI-adopting companies, as of today, avoid using historical data as a source for training their algorithms.

Other concerns raised from the interviews are time and money. As Rifaldi states *"AI cannot be developed in house"* since this would require a large amount of time and money to be done properly. In fact, as stated by Rifaldi many companies are missing economic availability to be able to differentiate and develop AI in-house and, always according Rifaldi, many companies will never have this availability. Moreover, artificial intelligence, as expressed by Michetti, *"is a mechanism that tends to uniformity"*, therefore it has two problems. The first is to understand if there is a reason to fill a company with clones, while the second is to understand whether or not companies are able to fill themselves with clones. His answers to these questions were that there is no need to fill up companies with clones as the economy is not a Fordist market anymore and companies, even if they wanted to hire clones, are not able to do it since there is limited availability of data. In addition, in order to effectively use AI, companies need to be aware of what they are looking for because as Michetti states "*if you delegate the selection to the AI model by saying 'I am not sure of what I want, but you are going to understand it'*, […] *we will go towards hiring clones as the machine will choose what we need."*

In addition to training, another concern was raised, that of explainability. Diorinos states that what AI can do is replicate human intelligence and that because of this it is going to learn biased and is not going to be able to tell the human agent why it made those choices. Diorinos also provides an explanation of an experiment in which a machine was asked to recognized pictures of a horse (Lapuschkin et al., 2019). When exploring why the machine recognized it was a horse, the experimenters found out that it did not recognized the subject of the picture, but that it recognized the credit line as in its dataset the photographer was known to be a horse photographer. Hence, if the credit was of that photographer but the picture was not a horse, the machine would still say it represented a horse, thus undermining the capabilities of AI agents. Similarly, Michetti says that AI models will always provide an output but that the problem is the output's validity. Rifaldi also defines AI as a *"black box"* and agrees with Michetti and Diorinos on the fact that the reasons why a machine gives those results is not tracked and explained by the machine.

The interviewed AI adopting firms, however, are able to mitigate some of the expressed concerns, such as those related to training by using more objective from publicly available sources instead of using historical data containing prejudice to train the algorithms. Moreover, both Peltier and De Cristofaro when asked whether or not clients had any concerns said that as of today there were not any. De Cristofaro also expanded on the concept by saying that this could be because they might have prevented some of those concerns by sharing with their customers the basis on which the use of AI within their software is built, thus contributing to increase their credibility at the eyes of their customers.

At Manatal AI is involved in many of the offered tools including screening, recommending and scoring candidates for job positions along with data mapping. At In-Recruiting AI is involved in the CV acquisition phase and in the candidates' pre-screening phase, while at Cornerstone it is used within a variety of features including the identification of candidates' skills and in generating recommendations. From the phases in which AI is adopted, it can be observed that the full recruiting process is not entirely carried out by machines. Because of this, some of the above-mentioned concerns can be mitigated through human supervision since the human agent is always making the final decision. This is also highlighted by the fact that all of the interviewed adopters view AI as a tool for improving decisions and not as a component that can completely replace men. In addition, De Cristofaro and Cocciardo also highlight the positive impact that AI can have on the applicants. According to Cocciardo, AI in addition to saving time for candidates, also allows to overcome typing mistakes since the machine will recognize the searched word even if it is written incorrectly and will show the candidates as a result for the search, thus not penalizing them for a typing mistake. Moreover, according to De Cristofaro it can help with employee engagement as it is a valuable tool to define and suggest career paths. The fact that there are also benefits for employees in addition to those for the employers can also contribute to reducing the prejudice around this topic as candidates realize that the technology also simplifies their recruitment process and their future career paths.

Once and if the concerns are bridged, AI can have a variety of advantages such as helping to reduce the impact of the recruiter's lack of multi-domain knowledge as expressed by Cocciardo and can help to provide a broader and more objective view of the candidate pool as expressed by De Cristofaro. Always according to Cocciardo it can also have economic benefits as it reduces the cost of repetitive activities, along with image benefits with respect to the candidates who acknowledge that a company is always improving their way of working with advanced technologies. Moreover, as expressed by Peltier and Rifaldi it can help the recruiter to understand the candidates' skills. In addition, according to Michetti, a non-AI adopter, the use of AI can make sense *"when applied a on a single tool* [...] *as the recruitment process needs to use a variety of tools."*

4.7.2. Findings on the Evolution of the Recruiter's Role

When it comes to the impact of AI on the recruiter's role the opinions of the interviewed participants are divided in two opposite categories: those who believe that the recruiter's role will improve (Diorinos, De Cristofaro and Peltier), and those who believe it will not change (Michetti and Rifaldi). The remaining respondent, Cocciardo, argues that in order to understand whether or not the recruiter's role will change it is important to distinguish between two groups: companies that are not using any recruitment software and those that are already using some type of recruitment technology. According to him those companies that are not adopting any technology will experience a huge change, while for those already implementing a technology a change will occur, but not as drastic. However, he also argues that for any of the two types of companies the recruiter will benefit as he will be able to focus on value-adding activities, such as using human skills to choose the right candidate, while avoiding to perform low value-adding activities, thus making the recruiter faster and more precise. Similar to this statement, Diorinos argues that by using either AI or more science-backed systems, the recruiter's role is improved as he is provided with proper tools that allow him to do his job more efficiently. In addition, Diorinos also says that people make hiring decisions based and personal biases and on very short impressions and are left to discover all the problems going forward, therefore the tools provided by these systems can help to minimize the impact of subjectivity even if *"you can't take subjective out of the equation"*. Peltier is also of the opinion that the recruiter's role will change as recruiters will be increasingly able to focus on people-oriented tasks as AI can handle some of the tedious work. De Cristofaro also adds to Peltier's idea by saying that AI can improve the recruiter's role by automating activities, but also by allowing the recruiters to have a broader perspective than the one they might generate only based on their experience. Hence, all of these four respondents see AI, or in Diorinos case any recruiting software, as a tool to help the recruiter be more efficient by allowing him to focus on non-repetitive, people-oriented tasks.

On the other hand, Michetti, Rifaldi, and Cocciardo, in part, believe that the recruiter's role will not change if AI is adopted. Michetti explains that the reason for this is that as of today many companies are already implementing some type of recruitment tool that helps them perform the screening phase and are likely to continue adopting it even in the future. Hence, according to Michetti the main difference between using AI tools and non-AI adopting tools will simply be that the tools will work *"on different algorithms"* but the basic recruitment process will remain the same. Rifaldi, is also not of the opinion that the recruiter's role will change, however for a different reason. Being an Account Manager Rifaldi addresses the topic by considering the adopting companies' points of view. Rifaldi explains how some companies stated that it is fundamental for them to meet the candidates in person and to perform the recruiting process face-to-face as doing differently would go against their core believes and they would result as incoherent. Moreover, he also states that the same reasoning applies for all those companies that state that people are their biggest resources as virtual hiring would undermine their proposition. Hence, for many companies the recruitment process would still be preferably conducted the old-fashioned way, therefore not changing the recruiters' role.

In addition, Rifaldi also claims that "*no machine can read a person"* and that the capability to go beyond few objective terms can only be found in men, hence that AI cannot replace humans. This topic was agreed on by the other respondents and human contribution is also defined as *"desirable"* by De Cristofaro. Peltier states that AI works best as a *"tool to help recruiters, not as a stand-along thing that would do the work in place of a recruiter"* which is also supported by Cocciardo who states that "*technology is an aid for men"* which helps to make more rapid and accurate decisions. Cocciardo also states that *"as of today there should not be any automated choices where there is no human supervision".* Diorinos also argues the same by saying that the recruiter will always be a core component of the recruitment process as in all its related areas there is a strong human element. Moreover, this idea is further developed by Michetti that illustrates that the recruiters have a wide array of tools which might be technological, but also linked to their know-how or speaking ability and that AI represents one of these tools, therefore it always requires men to work. The necessity of human supervision also makes Michetti question the real point of using AI. Moreover, when asked about the potential benefits of AI he states that that from the recruitment point of view there are none, but that it is efficiently adopted in pure recruitment by all the platforms or websites that match job offers with candidates' profiles. On the other hand, adopters of AI such as Cocciardo say that AI can have benefits for the recruiter as it reduces the time taken to perform repetitive activities, thus speeding-up the process while also providing economic benefits. Moreover, Peltier made a similar consideration by saying that models that help recruiters identify the relevant candidates are of great use since analyzing all the applicants to find the few relevant ones will be a *"huge time investment for an HR team."* De Cristofaro also states that AI reduces the time taken for recruiters to go through the process.

The interesting insight that emerged from the interviews is that also non-adopters of AI state that there are benefits for the recruiter if their concerns are solved and overcome. Rifaldi says that AI could be a valid tool since it allows to identify a person's accent and language skill, a benefit that is also agreed on by Michetti who states that AI can be successfully implemented to evaluate fluency. In addition, also Diorinos says that there are phases of the process where AI can be implemented such as those phases in which it is possible to do *"a lot of processing that is meaningful"*.

4.7.3. Findings on Bias Reduction

From the interviews it emerges that whether or not an algorithm can correct for human bias strongly depends on how the algorithm is trained. When asked if AI can reduce human bias, 4 out of 6 participants (Cocciardo, Diorinos, Michetti and Peltier) pointed out Amazon's scandal with recruitment algorithms to stress the importance of effective training. Diorinos further expanded on the concept by explaining that if we take a sample of software engineers it is going to be prevalently males and that because of this the AI will pick up on the gender-related clues and will likely base its decisions by enforcing a prejudiced reasoning. Diorinos also says that the AI systems work exactly as they are supposed to, but that the problem is the data on which the algorithm is trained and that because of this the algorithm will be self-confirming as the dataset is biased from the beginning therefore no anomaly is perceived. Moreover, Michetti highlights how if the AI system is not working correctly it will guide the hiring process anyway, the problems is that it will guide it in an incorrect way. Moreover, when asked whether or not an algorithm can correct for bias the problem of data availability was brought up. As said by Michetti *"what is good for a company might not be good for another* […] *so the machine would need to be trained being aware of what works for a company and what works for another."* As presented by Michetti well-known companies which hire thousands of employees per year, still do not have enough data to correctly train the algorithm based on their specific requirement, hence it would be even harder for smaller companies to accurately train the algorithm. Rifaldi agrees with Michetti by saying that AI in order to correct for bias needs to be used *"well"* and that well depends on who trains the machine since the *"data on which the training is based needs to be genuine"*, hence he is not really of the idea that AI can correct for bias. He also adds that if the company that trains the AI has strong morals then the AI will help the company, however if the training company presents questionable morals and cultural barriers, the designed algorithm will reflect all of those questionable characteristics. Another interesting point raised by Michetti is that if companies have problems related with bias it is linked to the fact that they are not using any tool to help them reduce this problem. Because of this, even if efficient AI tools exist but companies refuse to use them, they will continue to exhibit the same flaws they are exhibiting today.

The Amazon case, which was mentioned by 2 out of 3 interviewed adopting firms, in addition to highlighting the issues related to data training, also served as a lesson for the AI adopting firms which are now required to be more cautious with their training. The difference between the adopting companies and the non-AI adopters is that Peltier, De Cristofaro and Cocciardo believe that AI can be a tool for bias reduction as they are training the algorithms on unbiased data and in fact are able to design bias-free algorithms. Peltier explains that at Manatal they use public data on the internet as the *"larger the data sample, the least chances you have of having a very small minority that influences the outcome."* Similarly, Cocciardo when asked if AI can be an efficient bias reduction tool said *"absolutely"*. When asked the same question De Cristofaro illustrated how humans are led to reason and form impressions based on experience, and that because of this there can be some prejudice which can be corrected by the algorithm which *"can help recruiters make informed and objective decisions"*.

Although positive of the fact that AI can correct for bias, Cocciardo also agrees with the other participants by stating that *"if you give it* [the algorithm] *data that is previously affected by human bias or opinions* […] *it is obvious that you will obtain* *an algorithm with that implicit prejudice”*. Because of this knowledge and after Amazon’s scandal, at In-Recruiting they decided to avoid using historical data to train algorithms and to exclude some information, such as name, surname, gender, age and nationality, from their models of candidate suggestion and scoring to ensure that they are completely unbiased. De Cristofaro also highlights the importance of training by saying that algorithms can correct for bias provided that the algorithm is sophisticated and trained on a valid dataset. Cocciardo however, also explains that in order to reduce bias alternative methods can be used as it would be sufficient to analyze the recruiters’ performance in order to ensure that there is no prejudice from the person in charge of recruitment.

4.8. Discussion

The efficacy of applying AI to the recruiting process is a very debated topic and in fact many conflicting, but also many agreed on, points emerged during the interviews as it can be observed in the table below.

		SOFTWARE					
		Allibo	Altamira	Bryq	Cornerstone	In-Recruiting	Manatal
CONCEPTS	AI is used within the recruitment process	✕	✕	✕	✓	✓	✓
	Technology is a tool for human agents	✓	✓	✓	✓	✓	✓
	Some datasets can be biased	/	✓	✓	/	✓	✓
	Amazon case example was made	✓	✕	✓	✕	✓	✓
	Recruitment always requires a human element	✓	✓	✓	✓	✓	✓
	Algorithms can perform well in some phases of the recruiting process	✓	✓	✓	✓	✓	✓
	Recruitment softwares are helpful for both employers and employees	/	/	✓	✓	✓	/
	AI can make the final decision	✕	/	/	✕	✕	✕
	Recruiter's role has changed	✕	✕	✓	✓	✕/✓	✓
	There is a sufficient amount of data available to train the algorithms	✕	✕	✕	✓	✓	✓
	AI can reduce human bias (if the algorithm is efficient)	✕/✓	✕	/	✓	✓	✓
	Effective training is crucial	✓	✓	✓	✓	✓	✓
	Different companies require different employees' charactersitics	✓	/	✓	/	/	/
	AI is sometimes adopted without sufficient scientific research and inefficiently by companies	✓	✓	✓	/	/	/
	Explainability concerns	✓	✓	✓	/	/	/
	Important to make hiring decisions considering factors beyond experience	✓	✓	✓	✓	✓	✓
	AI is as good as its application	/	✓	✓	/	/	/
	Ethical concerns with AI adoption	✓	✓	/	/	/	/

4.8.1. Limitations and Opportunities to Overcome Them

The extent to which recruitment experts believe AI can be applied to the recruitment process varies widely. The arguments presented by the interviewees were very consistent with their companies’ way of carrying out the recruitment process. Cocciardo, De Cristofaro and Peltier support the use of AI in some phases of the recruitment process. Whereas, Michetti, Rifaldi and Diorinos are generally opposed to the use of AI within the recruitment process but not to AI in a general sense, in fact they state that the technology itself has potential. The main reasons why they are against the use of AI within recruiting are the various concerns that have been brought up during the interviews. The interviews revealed that the concerns presented within the studied literature match with the issues that are presented in more practical terms.

The major concern, which is also shared by the interviewed AI adopters and which is also present within the literature, is that of training. The main issue related to training, as presented by Raviprolu, Michetti and Diorinos is that machines need an extensive amount of data in order to effectively train the machines. This would mean high costs for the AI adopters as bias-free data is becoming increasingly costly and hard to find since policies to protect data are constantly developing as also illustrated by Campolo. GDPR policies are also becoming increasingly important when it comes to data sourcing as the terms and conditions would need to be clearly stated to the candidates, thus making data collection harder as applicants might not want their data to be sent to third parties. Contributing to the rising concerns linked to AI is that fact that even AI supporters such as Wisskirchen and Cocciardo are aware that some algorithms, even the most advanced ones, can make mistakes because of incorrect training, thus making the reliability of algorithms wobble. In addition to the datasets used to train the algorithms, it is also important to acknowledge the ethical concern related to who is training the machine. Therefore, it is fundamental to assess if the scope for which people are training the machine is a beneficial or a detrimental one. Because of this, in-house creation and training of the algorithm would be preferred, although costly and time consuming, as it would allow AI adopters to clearly explain to their customers the data on which the algorithms are trained and what they are trained for. This would be initially costly for companies, however in the long-run it will result in more credibility of their algorithm and also in more reliability to the eyes of the customer and the applicants, thus favoring the adoption of the technology from third parties.

Another very relevant concern that is raised by Okolie et al. is that of impersonalization of the process both from the recruiter and on the candidate's part. This issue is very relevant in practical terms as presented by Michetti. According to him, if a recruiter is not sure of what he is looking for, and according to Michetti 95% this is the case for the majority of Italian companies, the AI agent will be trained to decide who he thinks is the best candidate and will therefore select the same type of candidate over and over again, thus leading to hiring clones. In the long-term this is detrimental for a company as diversity has proved to be beneficial for thriving, therefore solving this issue should be of primary importance for AI adopters. However, De Cristofaro states that AI can provide recruiters with a broader point of view, thus suggesting that when it comes to hiring, recruiters will have a wider range of elements on which to make decisions, thus avoiding to hire clones. Moreover, this also links to a concern presented by Parnas in the literature with which Diorinos agrees to, and it is that of AI being a supervised machine learning, meaning that AI can only replicate human behavior and that therefore it could result in untrustworthy programs. Partially because of this, but mostly because of the current state of technology, Tandon and all of the interviewees believe that AI cannot result in the full automation of the recruitment process, but that it can be successfully applied only to some phases. In fact, even the three AI-adopting companies that were interviewed use AI only for some phases of the recruitment process. In fact, Cocciardo, Peltier and De Cristofaro state that their technology is a tool for humans and not a replacement. Cocciardo also states that there should not be any automated decision and because of this the recruitment process is always supervised by a human agent who is then in charge to take the decision.

Another important concern that is highly important in practical terms has been presented by Rifaldi and Michetti who state that there are companies that are only flagging their products as using AI while in the end the tool is not really using any AI agent, or even worse, is using an AI agent in an inefficient way. Both situations are contributing to raise concerns with respect to the effectiveness of AI as customers are led to either believe that AI does not add much to the systems or that AI is ineffective as it is wrongly applied by its providers. Since a tool is as good as its application, if AI is used improperly it will not lead to the creation of benefits for the recruiter and in the long-run it will also jeopardize the company's recruiting process as biases will be included in recruitment instead of eliminated. Another reason why some of the interviewed recruiting expert (Diorinos, Rifaldi, Michetti) tend to stand clear of AI is because it has zero, or close to zero, explainability. This concern is very relevant in practical terms since if it is not clear why the algorithm chose that candidate, then the process might not be perceived as reliable since candidates might believe that fit is purely based on coincidences, as it was illustrated in the horse picture example.

However, this concern is partially tackled by AI adopters, such as Cocciardo, who claim that if a machine makes an error it can be easily tracked down and modified which is rarely the case with human behavior which is harder to change and correct.

Although all of these valid concerns have been raised by both the literature and the interviewee, independently of whether or not they adopt AI, companies such as In-Recruiting, Manatal and Cornerstone have been able to incorporate AI within the recruiting process, thus proving that it is possible to overcome some of the expressed concerns. These companies mainly use AI tools for acquiring resumes, to pre-screen the candidates and to provide a scoring of candidates and De Cristofaro also states that *"AI is fundamental to ensure that the recruiter is even more efficient".* Moreover, all the interviewed AI adopters underline that their companies avoid using historical data as they are biased and tend to use more objective data sources which are free of human prejudice, thus solving the concerns related to training as they are able to provide bias-free algorithms. By training algorithms on datasets that are free of prejudice, AI can have an efficient role when it comes to the above-mentioned phases of recruiting as it helps to speed-up the process while improving the candidate's perception of a company. Nevertheless, concerns might still be raised on the culture of the company that trains the AI and which could later affect the algorithm. However, that is not a direct error of the system as this error could occur independently of whether or not AI is used as the prejudice can also stem from the recruiter who will make the final decision. It can be therefore assessed that some concerns can be accounted for and solved, however as presented by Michetti, Rifaldi and Diorinos the recruitment process should not be solely based on resumes as the candidates' soft skills also play an important role and AI agents are not always able to fully grasp them. Concerns and limitations of the technology, such as the afore mentioned one, are recognized by the interviewed AI adopters who in fact see AI as a tool to help human agents and not as a stand-alone tool, a concept on which all the interviewee agreed on.

All of the concerns raised on the adoption of AI are very tangible and if not overcome can have a detrimental impact on both the software providing company and on the adopting one. As not all issues can be solved with the current state of technology, AI should not be in charge of carrying out the entire process alone. In fact, it is mainly used to take care of time consuming and repetitive activities, such as resumes screening, while leaving the human agent to take care of more value-adding activities, such as making the final choice. An efficient application of AI would help both the candidates and the employees as it would reduce the time taken to complete the recruitment process. However, the technology is not advanced enough to carry out the entire recruitment process alone and still has some limitations. These limitations however can be perceived as a double-edged sword. From one point of view the fact that AI is not automatically carrying out the entire recruitment process reduces the concerns that machines can take over humans, therefore reassuring individuals on their fear of losing their job. On the other hand, the current technological state of AI limits its potential since if an algorithm requires human supervision its benefits can be only partially experienced. Furthermore, the fact that humans are also employed within the process can also help to mitigate the explainability concern since it is the human who gives the machines the requirements to look for in candidates and in the end, it is the human recruiter who makes the final decision. However, can we really know why a human recruiter decided to choose one candidate over the other? Also, can we be sure that the reasons why he chose those candidates are not subjective? Hence, whether or not the recruitment process adopts AI agents or not, it will always be subject to some type of explainability concern.

4.8.2. Recruiter's Role

From the interviews, conflicting opinions emerged from the ones expressed in the literature as all of the participants agreed on the fact that the human recruiter plays, and will always play, at least with the current state of technology, a fundamental role within the recruiting process. Hence, the recruiter's role will not disappear in the near-future. This is mainly linked to the fact that machines are not advanced enough to acquire human skills such as empathy, therefore requiring a human agent to always supervise and intervene during the process.

Although the interviewees share the view of the human agent being a crucial part of the process, therefore agree on the fact that the role will not be eliminated, they have diverging opinions on the level of change that the recruiters will experience. According to Meister and Manyika et al., recruiters will have to change their role as they will need to develop skills in order to understand how to use AI across the employees' lifecycle. However, as it emerges from the interviews, no mention to the recruiter's skills is made. This is mainly because the recruiter's role will basically remain the same as he will have to perform the same tasks as before, such as performing hiring decisions. The only difference is in the fact that the machine will take care of some of the tasks for him. Therefore, the recruiter will need to learn how to implement AI, however the skills he will need to learn will be inferior to the ones he can let go of. The recruiter will actually require less skills, as he will no longer use skills such as resumes screening daily, and will be able to focus more on developing his existing skills related to value-adding activities. In addition, Diorinos argues that the recruiter's role is improved as AI allows to give to the recruiters tools to perform their job better. However, he is referring to tools and not skills, therefore it is granted that the employees will need to learn how to use those tools, but they will not need to learn new skills since the tools will take care of that job for the recruiter. Cocciardo instead has a mixed opinion on the future of the recruiter's role. He distinguishes among companies that are already adopting some recruiting software and companies that are not. He goes on to say that for companies that are not using any software, the adoption of AI will surely change the recruiter's role as it will completely revolutionize the recruiter's tasks by automating repetitive ones and allowing him to focus on more value-adding activities. On the other hand, companies already adopting a software will not experience that big of a change with respect to the recruiters' role as AI will add a technological level to the process and will reduce the time taken, but it will not affect the recruiter's tasks, thus completely opposing the theoretical foundations. De Cristofaro's point of view on the recruiter's role is also opposing the literature as she claims that it is *"fundamental to allow the recruiter to be more efficient without impeding him to intervene"* therefore the recruiter will have to hold on to his previously learned skills as they are still of great use within the process.

Also opposing the theoretical assumptions are Michetti and Rifaldi's views as they both believe that the recruiter's role will not change. Michetti argues that companies that are looking for AI systems already have some tools to help them during the screening phase, therefore the recruiter's role will not change as the difference relies on the type of algorithm used by the model. Rifaldi is also of the idea that the recruiter's role will not change based on the fact that some companies are simply not ready to completely automate the recruiting process as it goes against their principles and could damage their public image. Hence, for some companies it is important to pursue the typical recruiting process, therefore the recruiter's role will not change as the process will mostly remain the same.

The theory also suggests that most of the jobs will disappear or be substituted with the adoption of AI, however the recruiter does not seem to fall within this category of jobs. From the findings, the importance of having a human agent involved in the recruitment process is extensively highlighted as machines are not able to use human skills to assess the candidates. Hence, the crucial role of supervisions suggests that the recruiter's role is not going to simply disappear in the future. However, the recruiter's role has changed over the years and will continue to change as the technology advances. Although there have been many changes concerning the recruiter's role, not all of them are a direct consequence of the adoption of AI as some changes are the result of the creation of software that help the recruiter's during the recruiting process independently of whether or not they are adopting AI, therefore changing the tasks that the recruiters need to perform. Many companies are already adopting a software to speed up the recruitment process, therefore if these companies adopt AI models, they might have an increase in the model's performance but the recruiter's role will not be subject to any change as the tasks he performs will remain the same. In addition, AI tools can improve the recruiters' role by providing them with tools to fill their voids when it comes to skills, and by helping them by suggesting which candidates are more suitable for a position or which should be assessed for a different one with respect to what they applied.

The key point however, stays the fact that AI, as much as any other current recruitment software, is not a stand-alone tool. Hence, it will always require human supervision in order to be successfully implemented. What this means for the future of the recruiter's job is that the recruiter will always remain a relevant figure within the process, but his tasks will change as he will be able to focus on more people-oriented tasks while the machines take care of tedious and repetitive ones. Moreover, because of this, the recruiter will also be able to improve his human-related skills, such as his job interview technique, as he will have more time to focus on more people-oriented tasks. Hence, in light of these findings the practical impact of AI on jobs, such as the recruiter's role, should be re-assessed taking in consideration the current state of technology and its limitations.

4.8.3. Bias Reduction

A common theme that emerged both from the literature and from the interviews is that the key component to ensure a bias free algorithm is a bias free dataset on which to train the algorithm. Campolo et al. and Cocciardo point out that if algorithms are trained on biased datasets, the resulting AI agent will implicitly be subject to that bias.

This claim that is supported by all of the interviewed participants, therefore if an algorithm is wrongly trained, it will undo all the benefits that AI models could bring to a company. Moreover, this claim is strengthened by examples given by both Diorinos and Florentine that explain how a biased dataset can lead to gender discrimination as the algorithm will not directly recognize the genders but will pick-up on gender related clues and apply them during the selection process, thus discriminating certain candidates.

In addition, both the literature and the interviews are divided among supporters of AI as a bias reduction tool and those who believe it is not a valuable tool. Cocciardo, De Cristofaro and Peltier believe that AI can be a valuable tool to reduce bias, provided that it is trained effectively, therefore agreeing with the theory presented by Wisskirchen et al. If well trained, AI can analyze a larger quantity of data faster and more accurately and at the same time could help in identifying hidden skills within resumes and implicit biases within the human recruiters. However, as suggested by Cocciardo, these applications could also be performed by non-AI tools as any technology could help to analyze the prior decisions taken by the recruiter and analyze whether or not the recruiter is impartial, however AI might do it faster. Michetti agrees with Coccairdo on the fact that technologies that are independent of AI can already do this and also adds that if companies have issues arising from bias, it is because they are not using objective selection tools and that these types of issues are linked to deeper problems that are independent of AI, such as problems related to a company's internal culture. It is for this reason that both Michetti and Rifaldi are not entirely sure that AI can be a valid tool for reducing bias, mostly depending on the fact that this assumption would require efficient datasets on which to train the algorithm which are rarely available.

From the research it is clearly underlined how AI can be a valid tool for bias reduction if and only if it is trained on appropriate unbiased data sets, since if the dataset is biased also the algorithm will be biased. In order to solve this problem, AI adopting firms are eliminating biased datasets from their training models and are training their algorithms on bias-free data in order to ensure that the designed algorithms are completely unbiased. This allows to take the subjective aspects out as much as possible from the recruitment process, however it is impossible, contrarily to what Wisskirchen et al. claim, to completely remove emotions and sympathies as human agents are still involved in the final phases of the process. During the initial screening phases the process is completely free of bias as the algorithm is focusing solely on data, assuming the algorithm is well trained. However, the final decision is always taken by a human agent as machines are not able to make these decisions, therefore subjective opinions will always affect the process Hence, AI is a valid tool to mitigate the presence of bias, provided that it is correctly trained, but it is not capable of completely eliminating bias as humans, which will always have some type of bias even if they are not aware of it, will be involved in the process.

Moreover, it is important to understand whether or not biases are always negative for companies. From one point of view the presence of bias can be a problem as presented by PwC. According to PwC (2017) eliminating biases, even unconscious ones, could help make faster decisions and help diversify the employee pool, thus providing beneficial advantages to the company. On the other side, it can be argued that biases help to identify a specific type of employee, thus ensuring that the employees will all have the same, or very similar, values therefore easily fit within the company's culture at the expense of diversification. However, the presence of bias can lead to hiring clones, therefore creating obstacles to the long-term development of a company. In addition to clones, biases can damage a company's image since a portion of the candidates will always be excluded because of discriminatory issues instead of insufficient skills. Moreover, discriminating candidates is rightfully illegal therefore companies could face legal problems Therefore, the presence of implicit bias can create very small benefits to the company's structure that are largely outweighed by the advantages that bias-free assessments can bring to a company, hence leading companies to prefer bias-free, both for legal and practical reasons.

4.8.4. Case Study Applications

Although the case study focuses on the recruitment process, the findings can be generalized and can also be applied to a variety of other fields. Even though only 6 companies participated in the study, they all agree on one idea: human supervision is fundamental. This concept is very relevant as the modern society is always trying to automize every task in a variety of job positions as it is perceived as a tool to reduce the amount of work that needs to be carried out by human individuals and since it can reduce costs and time constraints. With the current state of AI technology however, this is not possible. Self-driving cars are becoming increasingly popular and debated among society as they would perform all the driving steps on their own. However, if we are not able to completely trust a machine to make hiring decisions, can we trust a machine to self-drive a car where a mistake can cause much worse consequences than a failed hiring process? Hence, it is important to thoroughly recognize the limitations that this technology has and adopt it with those obstacles in mind.

Moreover, the case study calls for a re-evaluation of the existing theory regarding the effects that AI can have on human job roles. The literature claims that many jobs will be eliminated and will be substituted by a newer version of the original jobs. However, with the current state of technology human supervision is always required in order to ensure that the processes are performed efficiently. Hence, many tasks and job roles performed by individuals will not be eliminated. Jobs such as the recruiter and the driver, or any type of job that requires important decision-making, cannot be completely performed by an AI agent only. Because of this the future theory on the topic could focus on dividing job tasks in different categories: those that can be completely performed by AI, such as cleaning processes in which a robot could simply automate the tasks, and those in which the humans will always have a crucial role, which are those in which critical decisions are made and those in which wrong decisions can have harmful impacts on society.

4.9. Study Summary

The current recruitment process has been the same for decades (Singh, 2017) and it has been found to have some pitfalls as technologies are developing. A solution to minimize these pitfalls can be the adoption of AI, however it needs to be carefully introduced in order to maximize the benefits while mitigating the pitfalls.

Three themes have been analyzed in order to gain an insight on what impacts AI can have within the process. The analyzed themes Limitations and opportunities to overcome them, the Evolution of the recruiter's role and Bias reduction, have led to interesting considerations with respect to the topic. From the literature and the conducted interviews, it emerged that there are multiple concerns linked with AI as it can be costly to develop, it has no explainability and as it needs to be trained efficiently in order for it to work. Although some of the concerns cannot be bridged, as demonstrated by the fact that AI is not able to carry out the entire recruitment process by itself, others such as the concern of machines taking over human agents can be minimized as a human recruiter will always be responsible for making the final decision. In addition, the crucial role played by the human recruiter within the process ensures that the recruiter's role will still be present in the future and that it will only slightly change since AI models can take care of tedious tasks for them but the recruiter will always have to focus on human related tasks. On the other hand, the required presence of a human agent can be detrimental when looking at AI as a bias correction tool. The initial phases of the recruitment process such as screening and resumes acquisition, will be completely bias-free, if the algorithm is correctly trained, but since the final decisions are taken by human agents their implicit biases could affect the process thus ensuring a bias-free process only at the initial stages.

Furthermore, another important factor to guarantee a bias-free assessment is efficient training. If an algorithm is trained on improper datasets or by companies that have implicit prejudices, the algorithm will turn out to be completely biased and therefore damage the integrity of the process and also the company's image. Therefore, as of today, AI can be a valuable tool to reduce bias in the initial steps of the recruitment process, as it can simplify the recruiter's decision-making, but the technology is not advanced enough to ensure an entirely automated recruitment process. Therefore, the adoption of AI within the decision-making process cannot guarantee a bias-free and efficient process from start to finish.

Moreover, the conducted study provides insights on the current level of the technology and on how it is perceived by experts in the recruitment field, but could be generalized to many more areas as the basic benefits and limitations of the technology are the same independently of the field. However, the study has some limitations such as the fact that it is based solely on literature and interviews and not on practical observations. Hence, for further studies it would be helpful to directly observe the effects that this technology has on companies through first-person observations.

Conclusion

This thesis aims at analyzing whether or not the currently available AI tools can be valid instruments to improve the decision-making process. After examining literature concerning the main reasons why human agents divert from optimal decision-making and on the topic of AI, a study was conducted by interviewing 6 experts of the recruitment field. The recruitment filed was chosen since the recruitment process requires different levels of decision-making on both the recruiters and the applicants' part, hence it is of great use to understand the practical implications of the technology. By interviewing experts of the field, some supporting AI and other against it, a full spectrum of perspectives was gathered. The interviewees were asked questions with the aim of obtaining answer to questions related to the following research objectives: understanding whether or not the concerns linked to AI adoption can be bridged; gaining insights on how the recruiter role will evolve with the adoption of AI tools; and assessing if AI can be a valid tool to reduce bias within the decision-making process. Within the case study, three themes have been analyzed in order to understand the impacts AI can have within the process. The analyzed themes: Limitations and opportunities to overcome them, Evolution of the recruiter's role and Bias reduction, have led to interesting considerations with respect to the topic.

The literature and the conducted interviews highlighted a variety of concerns related to AI adoption such as its high development costs, its lack of explainability and the importance of algorithm training, which if done wrong will jeopardize the efficacy of the tool. Although some of the presented concerns cannot be bridged, as highlighted by the fact that AI is used as an aid for human agents and not as an independent component, AI can still be valuable within the recruitment process. Some of the presented concerns can be minimized by the presence of human agents who are supervising the process and since the human recruiter will always be responsible for making the final decision. In addition, the crucial role played by the human recruiter within the process will ensure the durability of the recruiter's role which will not be eliminated, but will be improved, by the presence of AI. Since AI models will take care of tedious tasks, the recruiters will have more time to carry out human-related ones, thus being more able to focus on a lower number of repetitive tasks while speeding-up the process.

On the other hand, the required presence of a human agent can diminish the potential of AI as a bias correction tool. Even though the initial phases will be completely bias-free as AI will solely focus on data, provided that the algorithm is correctly trained, the final decisions will always be made by a human agent. As highlighted both in the literature and through the interviews, human agents will always engage in subjective decision-making and are often subject to implicit biases which will affect the process. Thus, even with AI, a completely bias-free process is not entirely possible as it will be free of bias only in the initial stages. Moreover, another important factor to guarantee a bias-free assessment is efficient training. If the used datasets are compromised, also the algorithm will be, thus making the entire process inefficient.

Since AI is not technologically advanced enough to completely automize the entire recruitment process, it can be a valid tool to speed-up the decision-making process and to reduce bias only in the initial phases of the recruitment process. In the starting stages, in which AI is efficiently used, the algorithm will base its decision on objective data and consequently will suggest candidates based on unbiased terms. However, the need to delegate the final decisions to a human agent does not allow for an entirely bias-free assessment as the individual will make the final decisions which will be affected by subjective components, thus re-inserting bias within the process.

The study also gives a brief insight on the impact that the technology can have on the candidates. Since the topic is not frequently addresses by the literature, further studies could address the effects that the adoption of AI has on candidates and how they perceive the recruitment process when being aware that they are being assessed by a machine.

References

Abadi, Martin, Andy Chu, Ian Goodfellow, Brendan H. McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. "Deep Learning with Differential Privacy." In *ACM SIGSAC Conference on Computer and Communications Security*. Vienna, Austria, 2016.

Acemoglu, Daron, and David Autor. "Skills, Tasks and Technologies: Implications for Employment and Earnings." *Handbook of Labor Economics*4 (June 2011): 1043–1171.

Acemoglu, Daron, and Pascual Restrepo. "Automation and New Tasks: The Implications of the Task Content of Technology for Labor Demand." *Journal of Economic Perspectives*, 2019.

Acemoglu, Daron, and Pascual Restrepo. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review*108, no. 6 (January 2018): 1488–1542.

Aghion, Philippe, Benjamin Jones, and Charles Jones. "Artificial Intelligence and Economic Growth." *NBER Working Paper Series*, October 2017.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. "Competition and Innova- Tion: An Inverted-U Relationship." *The Quarterly Journal of Economics, MIT Press*120, no. 2 (May 2005): 701–28.

Agrawal, Ajay, John McHale, and Alexander Oettl. *The Economics of Artificial Intelligence: an Agenda*. Chicago: The University of Chicago Press, 2019.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction." *NBER Working Paper Series*, 2019.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. *Prediction Machines: the Simple Economics of Artificial Intelligence*. Boston, MA: Harvard Business Review Press, 2018.

Aharony, Nadav, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. "Social FMRI: Investigating and Shaping Social Mechanisms in the Real World." *Pervasive and Mobile Computing*7, no. 6 (2011): 643–59.

Allen, Gregory C. "Putin and Musk Are Right: Whoever Masters AI Will Run the World." CNN. Cable News Network, September 5, 2017. https://edition.cnn.com/2017/09/05/opinions/russia-weaponize-ai-opinion-allen/index.html.

Amartya, Sen. "Damage Measures for Breach of Contract." *The Bell Journal of Economics*11, no. 2 (1980): 446–90.

Amershi, Saleema, Maya Cakmak, William Bradley Knox, and Todd Kulesza. "Power to the People: The Role of Humans in Interactive Machine Learning." *AI Magazine*35, no. 4 (2014): 105–20.

Amodei, Dario, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. "Concrete Problems in AI Safety." In: arXiv:1606.06565, 2016.

Anand, Krishnan S., and Haim Mendelson. "Information and Organization for Horizontal Multimarket Coordination." *Management Science*43, no. 12 (1997): 1609–27.

Armstrong, Stuart. Motivated value selection for artificial agents. In *AAAI Workshop: AI and Ethics*, 2015.

Arrow, Kenneth J. An extension of the basic theorems of classical welfare economics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1951.

Asimov, Isaac. *I, Robot*. Garden City, NY: Doubleday, 1950.

Ayres, Ian, and Robert Gertner. "Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules." *The Yale Law Journal*99 (1989): 87–130.

Ayton, Peter, and Ilan Fischer. "The Hot Hand Fallacy and the Gambler's Fallacy: Two Faces of Subjective Randomness?" *Memory & Cognition*32, no. 8 (2004): 1369–78.

Ayton, Peter, Anne J. Hunt, and George Wright. "Randomness and Reality." *Journal of Behavioral Decision Making*4, no. 3 (1991): 222–26.

Baguley, Joe, and Anvi Patel. "Artificial Intelligence, Machine Leaning and it's place in our Society*."* Webinar, VMware, London,UK, April 14, 2020.

Bainbridge, Lisanne. "Ironies of Automation." *Automatica*19, no. 6 (1983): 775–79.

Bandura, Albert. *Self-Efficacy: the Exercise of Control*. New York: W.H. Freeman, 2012.

Barksdale, Jim. "A Quote by Jim Barksdale." Goodreads. Goodreads, November 2, 2019. https://www.goodreads.com/quotes/655987-if-we-have-data-let-s-look-at-data-if-all.

Barnheim, Douglas B., and Michael D. Whinston. "Incomplete Contracts and Strategic Ambiguity." *The American Economic Review*88, no. 4 (1998): 902–32.

Baslandze, Salome. "The Role of the IT Revolution in Knowledge Diffusion, Innovation and Reallocation." EIEF manuscript, 2016.

Baumol, William J. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis." *The American Economic Review*57, no. 3 (June 1967): 415–26.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. “The Distinct Effects of Information Technology and Communication Technology on Firm Organization.” *Management Science*60, no. 12 (2014): 2859–85. https://doi.org/10.1287/mnsc.2014.2013.

Bohn, Roger E. “From Art to Science in Manufacturing: The Evolution of Technological Knowledge.” *Foundations and Trends in Technology, Information and Operations Management* 1, no.2 (2005):1-82.

Bolton, Patrick, and Antoine Faure-Grimaud. “Satisficing Contracts.” *The Review Of Economic Studies*77 (2010): 937–71.

Bongard, J., V. Zykov, and H. Lipson. “Resilient Machines Through Continuous Self-Modeling.” *Science*314, no. 5802 (2006): 1118–21.

Bonnefoi, Pierre-François, Dimitri Plemenos. “Constraint satisfaction techniques for declarative scene modelling by hierarchical decomposition.” *International Conference, Limoges (France*), 2000.

Bradburn, Norman M., Seymour Sudman, and Brian Wansink. *Asking Questions: the Definitive Guide to Questionnaire Design*. San Francisco: Jossey-Bass, 2004.

Brun, E., & Milczarek, M. 2007. *Expert forecast on emerging psychosocial risks related to occupational safety and health*. European Agency for Safety and Health at Work.

Brynjolfsson, Erik, and Kristina McElheran. “Data in Action: Data-Driven Decision Making in U.S. Manufacturing.” *SSRN Electronic Journal*, 2016.

Brynjolfsson, Erik, and Lorin M Hitt. “Beyond Computation: Information Technology, Organizational Transformation and Business Performance.” *Journal of Economic Perspectives*14, no. 4 (2000): 23–48.

Brynjolfsson, Erik, and Tom Mitchell. “What Can Machine Learning Do? Workforce Implications.” *Science*358, no. 6370 (2017): 1530–34.

Brynjolfsson, Erik, Xiang Hui, and Meng Liu. “Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform.” *NBER Working Paper*, 2018.

Cakmak, Maya, Crystal Chao, and Andrea L Thomaz. “Designing Interactions for Robot Active Learners.” *IEEE Transactions on Autonomous Mental Development*2, no. 2 (2010): 108–18.

Camerer, C.F. “Artificial Intelligence and Behavioral Economics.” *The Economics of Artificial Intelligence*, 2019, 587–610.

Campolo, Alex, Madelyn Sanfilippo, Meredith Whittaker, and Kate Crawford. “AI Now 2017 Report,” 2017.

Capterra. “Software per Selezione Del Personale.” Capterra, 2020. https://www.capterra.it/directory/30602/recruiting/software.

Carrascosa, Juan Miguel, Jakub Mikians, Ruben Cuevas, Vijay Erramilli, and Nikolaos Laoutaris. “I Always Feel like Somebodys Watching Me.” *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies - CoNEXT 15*, 2015.

Carter, Craig R., Lutz R. Kaufmann, and Alex R. Michel. “Behavioral Supply Management: a Taxonomy of Judgment and Decision-Making Biases.” *International Journal of Physical Distribution & Logistics Management*37, no. 8 (July 11, 2007): 631–69.

Castelfranchi, Cristiano. “Artificial Liars: Why Computers Will (Necessarily) Deceive Us and Each Other.” *Ethics and Information Technology Volume*2 (2000): 113–19.

Chaker, N. “State of Recruitment Marketing 2018.” Beamery Talent Operating System, 2018. https://beamery.com/academy/recruitment-marketing-2018.

Clark, K.B, and T Fujimoto. “Roduct Development Performance: Strategy, Organization, and Management in the World Auto Industry.” *Harvard Business School Press*, 1991.

Clarke, Sharon G., and Cary L. Cooper. *Managing the Risk of Workplace Stress: Health and Safety Hazards*. London: Routledge/Taylor & Francis Group, 2004.

Cockburn, Iain, Rebecca Henderson, and Scott Stern. *The Impact of Artificial Intelligence on Innovation*. Chicago, MA: University of Chicago Press, 2019.

Conlisk, John. “Why Bounded Rationality? .” *Journal of Economic Literature*34, no. 2 (June 1996).

Crandall, J.W.and M.A Goodrich. “Experimenting adjustable autonomy”. *Systems, Man, and Cybernetics, 2001 IEEE International Conference,* volume 3, pages 1624– 1629, 2001.

Crandall, Jacob, W., Mayada Oudah, Tennom, Fatimah Ishowo-Oloko, Sherief Abdallah, Jean-François Bonnefon, Manuel Cebrian, Azim Shariff, Michael A. Goodrich, and Iyad Rahwan. “Cooperating with Machines.” *Nature Communications*9, no. 1 (2018).

Dastin, Jeffrey. “Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women.” Reuters. Thomson Reuters, October 10, 2018. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

Dawes, Robyn M. “The Robust Beauty of Improper Linear Models in Decision Making.” *American Psychologist*34, no. 7 (1979): 571–82.

Dawes, Robyn M., and Reid Hastie. *Rational Choice in an Uncertain World*. London: Sage, 2001.

Dayan, Peter, and Geoffrey E. Hinton. "Feudal Reinforcement Learning." *Advances in Neural Information Processing Systems*, 1993, 271–71.

Debreu, G. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. New York: John Wiley and Sons, Inc., 1959.

Dewey, Daniel. "Font Size: Reinforcement Learning and the Reward Engineering Principle." *AAAI Spring Symposium Series*, 2014.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err." *Journal of Experimental Psychology: General*144, no. 1 (2015): 114–26.

Dignum, Virginia. "Ethics in Artificial Intelligence: Introduction to the Special Issue." *Ethics and Information Technology*20, no. 1 (February 2018).

Dixit, Avinash K., and Robert S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

Eastwood, Joseph, Brent Snook, and Kirk Luther. "What People Want From Their Professionals: Attitudes Toward Decision-Making Strategies." *Journal of Behavioral Decision Making*25, no. 5 (2011): 458–68.

Emmen, Ad. "VMware's Joe Baguley Talks about Cloud Legislation, OpenStack, Customer's Choice, and Mobile Device Management at CloudScape VII." PrimeurWeekly and PrimeurMonthly HPC, supercomputer and Grid Computing news, April 27, 2015. http://primeurmagazine.com/weekly/AE-PR-05-15-98.html.
Entelo. *2018 Recruiting Trends Report.* (2018). https://cdn2.hubspot.net/hubfs/202646/Entelo%27s%202018%20Recruiting%20Trends%20Report.pdf?t=1530708036795

Epley, Nicholas, and Thomas Gilovich. "When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-Generated and Externally Provided Anchors." *Journal of Behavioral Decision Making*18, no. 3 (2005): 199–212.

Etzioni, Oren, Rattapoom Tuchinda, Craig A. Knoblock, and Alexander Yates. "To Buy or Not to Buy." *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 03*, August 2003, 119–28.

Feldman, M., S. A. Friedler, J.Moeller, C. Scheidegger & Venkatasubramanian, S. Certifying and removing disparate impact. In *Proc. 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 259–268, 2015.

Florentine, Sharon. “How Artificial Intelligence Can Eliminate Bias in Hiring.” CIO. CIO, December 22, 2016. https://www.cio.com/article/3152798/how-artificial-intelligence-can-eliminate-bias-in-hiring.html.

Frey, Carl Benedikt, and Michael Benedikt Osborne. “The Future of Employment.” *Oxford Martin Programmeon Technology and Employment*, September 2017.

Galbraith, Jay R. *Competing with Flexible Lateral Organizations*. Reading, MA: Addison-Wesley, 1998.

Galbraith, Jay R. *Designing Complex Organizations*. Reading, MA: Addison-Wesley, 1973.

García, Javier, and Fernando Fernandez. “A Comprehensive Survey on Safe Reinforcement Learning.” *Journal of Machine Learning Research*16 (2015): 1437–80.

Giusti, Alessandro, Jerome Guzzi, Dan C. Ciresan, Fang-Lin He, Juan P. Rodriguez, Flavio Fontana, Matthias Faessler, et al. “A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots.” *IEEE Robotics and Automation Letters*1, no. 2 (2016): 661–67.

Goodfellow, Ian, Jean Pouget-Abadie, Mirza Mehdi, Bing Zu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. “Generative Adversarial Nets.” *Advances in Neural Information Processing Systems*, 2014, 2672–80.

Gordon, Ruthanna, Nancy Franklin, and Jennifer Beck. “Wishful Thinking and Source Monitoring.” *Memory & Cognition*33, no. 3 (2005): 418–29.

Gray, Kurt, and Daniel M. Wegner. “Feeling Robots and Human Zombies: Mind Perception and the Uncanny Valley.” *Cognition*125, no. 1 (2012): 125–30.

Grove, William M., and Paul E. Meehl. “Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy.” *Psychology, Public Policy, and Law*2, no. 2 (1996): 293–323.

Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson. “Clinical versus Mechanical Prediction: A Meta-Analysis.” *Psychological Assessment*12, no. 1 (2000): 19–30.

Hadfield-Menell, Dylan, and Gillian K. Hadfield. “Incomplete Contracting and AI Alignment.” *SSRN Electronic Journal*, 2019.

Hana Ajakan et al. “Domain-adversarial neural networks.” In: arXiv preprint arXiv:1412.4446, 2014

Hao, Karen. “This Is How AI Bias Really Happens and Why It’s so Hard to Fix.” *MIT Technology Review*, February 2019.

Harris, Stanley G. “Organizational Culture and Individual Sensemaking: A Schema-Based Perspective.” *Organization Science*5, no. 3 (August 1994): 309–21

Harrison, E. Frank. “A Process Perspective on Strategic Decision Making.” *Management Decision*34, no. 1 (1996): 46–53.

Herbert, Simon. “Theories of Bounded Rationality.” In *Decision and Organisation*. North Holland, Amsterdam: McGuire and Radner Eds, 1972.

Highhouse, Scott. “Stubborn Reliance on Intuition and Subjectivity in Employee Selection.” *Industrial and Organizational Psychology*1, no. 3 (2008): 333–42.

Hockey, Robert, and Bob Hockey. *The Psychology of Fatigue: Work, Effort and Control*. Cambridge: Cambridge University Press, 2014.

Hogarth, Robin M. *Judgement and Choice: the Psychology of Decision*. Chichester: Wiley, 1987.

Iodice, Pierpaolo, Claudio Ferrante, Luigi Brunetti, Simona Cabib, Feliciano Protasi, Mark E. Walton, and Giovanni Pezzulo. “Fatigue Modulates Dopamine Availability and Promotes Flexible Choice Reversals during Decision Making.” *Scientific Reports*7 (April 3, 2017).

Jackson, Amy Elisa. “Popular Companies Using AI to Interview & Hire You: Glassdoor.” Glassdoor , January 1, 2019. https://www.glassdoor.com/blog/popular-companies-using-ai-to-interview-hire-you/.

Jackson, Matthew O. *The Human Network: How Your Social Position Determines Your Power, Beliefs, and Behaviors*. New York: Pantheon Books., 2019.

Jaworski, Bernard J., and Ajay K. Kohli. “Market Orientation: Antecedents and Consequences.” *Journal of Marketing*57, no. 3 (1993): 53–70.

Jobvite. “Why Are Passive Candidates the Best Candidates?” Jobvite, September 10, 2014. https://www.jobvite.com/sourcing-and-nurturing/passive-positive/.

Johnson, Matthew, Jeffrey M. Bradshaw, Paul J. Feltovich, Catholijn M. Jonker, M. Birna Van Riemsdijk, and Maarten Sierhuis. “Coactive Design: Designing Support for Interdependence in Joint Activity.” *Journal of Human-Robot Interaction*3, no. 1 (2014): 43–69.

Kahneman, Daniel, and Amos Tversky. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica*47, no. 2 (1979): 263–91.

Kahneman, Daniel, Andrew M. Rosenfield, Linnea Gandhi, and Tom Blaser. “Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making.” *Harvard Business Review*, 2016.

Kaplan, Jerry. *Artificial Intelligence: What Everyone Needs to Know*. New York, NY: Oxford University Press, 2016.

Keeney, Ralph L., and Detlof Von Winterfeldt. “Practical Value Models.” *Advances in Decision Analysis*, 2007, 232–52.

Kleinberg, J., Mullainathan S., and Raghavan M. “Inherent trade-offs in the fair determination of risk scores.” In: *arXiv preprint arXiv:1609.05807*, 2016.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass Sunstein. “Discrimination In The Age Of Algorithms.” *Journal of Legal Analysis*101 (April 2019): 113–74.

Kogut, Bruce, and Nalin Kulatilaka. “Operating Flexibility, Global Manufacturing, and the Option Value of a Multinational Network.” *Management Science*40, no. 1 (1994): 123–39.

Korn Ferry. “Putting AI in Its Place' Artificial Intelligence Should Be Part of the Recruiting Process, But It Can't Replace the Human Touch.” Korn Ferry, June 28, 2018. https://www.kornferry.com/about-us/press/putting-ai-in-place-artificial-intelligence-should-be-part-of-the-recruiting-process-but-cant-replace-the-human-touch-korn-ferry-survey.

Kramer, A. D. I., J. E. Guillory, and J. T. Hancock. “Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks.” *Proceedings of the National Academy of Sciences*111, no. 24 (2014): 8788–90.

Lachman, Sheldon J., and Alan R. Bass. “A Direct Study of Halo Effect.” *The Journal of Psychology*119, no. 6 (1985): 535–40.

Lambert, Vickie A., Clinton E. Lambert, and Hiroaki undefined Yamase. “Psychological Hardiness, Workplace Stress and Related Stress Reduction Strategies.” *Nursing and Health Sciences*5, no. 2 (April 2003): 181–84.

Lapuschkin, Sebastian, Stephan Wäldchen, Alexander Binder, Grégoire Montavon, Wojciech Samek, and Klaus-Robert Müller. “Unmasking Clever Hans Predictors and Assessing What Machines Really Learn.” Nature Communications. Nature Publishing Group, March 11, 2019. https://www.nature.com/articles/s41467-019-08987-4.

Lawrence, Paul R., and J. W. Lorsch. *Organization and Environment*. Boston, MA: *Harvard Business School, Division of Research*, 1967.

Lazer, D. M., and Et al. “The Science of Fake News.” *Science*359 (2018): 1094–96.

Leivo, Joel Z., Cyprien De Masson D’Autume, Daniel Zoran et al. “*Psychlab: A Psychollogy Laboratory for Deep Reinforcement Learning Agents.*” arXiv:1801.08116, 2018

Leka, S., Griffiths, A., & Cox, T. 2003. Work organisation and stress: Systematic problem approaches for employers, managers and trade union representatives. Geneva: World Health Organisation.

Lent, Robert W., Steven D. Brown, and Kevin C. Larkin. "Comparison of Three Theoretically Derived Variables in Predicting Career and Academic Behavior: Self-Efficacy, Interest Congruence, and Consequence Thinking." *Journal of Counseling Psychology*34, no. 3 (1987): 293–98.

Lerouge, Cynthia, Anthony Nelson, and J. Ellis Blanton. "The Impact of Role Stress Fit and Self-Esteem on the Job Attitudes of IT Professionals." *Information & Management*43, no. 8 (December 2006): 928–38.

LinkedIn. *The Ultimate List of Hiring Statistics. For Hiring Managers, HR Professionals, and Recruiters.* (2015). https://business.linkedin.com/content/dam/business/talent-solutions/global/en_us/c/pdfs/Ultimate-List-of-Hiring-Stats-v02.04.pdf

Lorenz, Tamara, Astrid Weiss, and Sandra Hirche. "Synchrony and Reciprocity: Key Mechanisms for Social Companion Robots in Therapy and Care." *International Journal of Social Robotics*8, no. 1 (2016): 125–43.

Lynn, Monty L., and Richard N. Williams. "Belief-Bias and Labor Unions: The Effect of Strong Attitudes on Reasoning." *Journal of Organizational Behavior*11, no. 5 (1990): 335–43.

Manyika, James, Susan Lund, Michael Chui, Jacques Bughin, Jonathan Woetzel, Parul Batra, Ryan Ko, and Saurabh Sanghvi. "Jobs Lost, Jobs Gained: What the Future of Work Will Mean for Jobs, Skills, and Wages." McKinsey & Company, November 2017. https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages.

March, Christoph. "The Behavioral Economics of Artificial Intelligence: Lessons from Experiments with Computer Players." *BERG Working Paper Series,*154 (October 2019).

Marengo, Luigi, and Corrado Pasquali. "How to Get What You Want When You Do Not Know What You Want: A Model of Incentives, Organizational Structure, and Learning." *Organization Science*23, no. 5 (2012): 1298–1310.

Maskin, Eric, and Jean Tirole. "Unforeseen Contingencies and Incomplete Contracts." *Review of Economic Studies*66, no. 1 (1999): 83–114.

May, Rob. "How Artificial Intelligence Impacts HR: Incorporating Intelligent Assistants in the Workplace." HR.com, August 24, 2017. https://www.hr.com/en/magazines/hr_strategy/august_2016_hr_strategy_planning/how-artificial-intelligence-impacts-hr-incorporati_is8sjn4q.html.

McCarthy, J, M.L Minsky, N Rochester, and C.E Shannon. “The Machine with Randomness.” Hanover, NY, 1955.

Meehl, P. E. *Clinical versus statistical prediction: A theoretical analysis and review of the literature*. Minneapolis, MN: University of Minnesota Press, 1954.

Meister, Jeanne. “Ten HR Trends In The Age Of Artificial Intelligence.” Forbes. Forbes Magazine, January 9, 2019. https://www.forbes.com/sites/jeannemeister/2019/01/08/ten-hr-trends-in-the-age-of-artificial-intelligence/#7cca78783219.

Mendelson, Haim, and Ravindran R Pillai. “Information Age Organizations, Dynamics and Performance.” *Journal of Economic Behavior & Organization*38, no. 3 (1999): 253–81.

Mikkelsen, Aslaug, Torvald Øgaard, Preben H. Lindøe, and Odd Einar Olsen. “Job Characteristics and Computer Anxiety in the Production Industry.” *Computers in Human Behavior*18, no. 3 (2002): 223–39.

Milgrom, Paul, and John Roberts. *Economics, Organization and Management*. Englewood Cliffs, N.J: Prentice-Hall International, 1992.

Miller, Alex P. “Want Less-Biased Decisions? Use Algorithms.” *Harvard Business Review*, July 2018.

Montibeller, Gilberto, and Detlof Von Winterfeldt. “Cognitive and Motivational Biases in Decision and Risk Analysis.” *Risk Analysis*35, no. 7 (2015): 1230–51.

Mullainathan, Sendhil H., and Richard H. Thaler. “Behavioral Economics.” *SSRN Electronic Journal*, October 2000.

Nease, Anjanette A., Brad O. Mudgett, and Miguel A. Quiñones. “Relationships among Feedback Sign, Self-Efficacy, and Acceptance of Performance Feedback.” *Journal of Applied Psychology*84, no. 5 (1999): 806–14.

Nutt, Paul C. “How Decision Makers Evaluate Alternatives and the Influence of Complexity.” *Management Science*44, no. 8 (1998): 1148–66.

Okolie, Ugo Chuks, and Ikechukwu Emmanuel Irabor. “E-Recruitment: Practices, Opportunities and Challenges.” *International Knowledge Sharing Platform* 9, no. 11 (2017).

Önkal, Dilek, Paul Goodwin, Mary Thomson, Sinan Gönül, and Andrew Pollock. “The Relative Influence of Advice from Human Experts and Statistical Methods on Forecast Adjustments.” *Journal of Behavioral Decision Making*22, no. 4 (2009): 390–409.

Parnas, David Lorge. “Inside Risks: The Real Risks of Artificial Intelligence.” *Communications of the ACM* 60, no. 10 (October 2017): 27–31. https://doi.org/10.1145/3132724.

Plemenos, Dimitri, Georges Miaoulis , and Nikos Vassilas . “Machine Learning for a General Purpose Declarative Scene Modeller.” In *International Conference Graphicon* . Nizhny Novgorod, Russia, 2002.

Plemenos, Dimitri, W. Ruchaud, and K. Tamine. “Interactive Techniques for Declarative Modelling.” *International Conference, Limoges (France*), 28-29 of April 1998.

Porter, Theodore M. *Trust In Numbers: the Pursuit of Objectivity in Science and Public Life*. S.l.: Princeton University Press, 1996.

Powell, Walter W. “Hybrid Organizational Arrangements: New Form or Transitional Development?” *California Management Review*29, no. 4 (1987): 67–87.

Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr. “Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology.” *Administrative Science Quarterly*41, no. 1 (1996): 116.

Puranam, Phanish, Yash Raj Shrestha, Vivianna Fang He, and Georg Von Krogh. “Algorithmic Induction Through Machine Learning: Opportunities for Management and Organization Research.” *SSRN Electronic Journal*, August 2018.

PwC. “No Holding Back: Breaking down the Barriers to Diversity.” *No Holding Back: Breaking down the Barriers to Diversity*. 2017. https://www.pwc.co.uk/human-resource-services/assets/documents/real-diversity-2017-no-holding-back.pdf.

Rachlinksi, Jeffrey J., Shery Lynn Johnson, Andrew J. Wistrich, and Chris Guthrie. “Does Unconscious Racial Bias Affect Trail Judges?” *Notre Dame Law Review*, no. 84 (March 2009).

Rahwan, I., M. Cebrian, N. Obradovich, *et al.* “Machine behavior.” *Nature* 568**,** 477–486, 2019.

Rahwan, Iyad. “Society-in-the-Loop: Programming the Algorithmic Social Contract.” *Ethics and Information Technology*20, no. 1 (July 2017).

Raiffa H. Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Oxford: Addison-Wesley, 1968.

Randstad. *2018 Talent Trends Report*. (2018). https://www.randstad.it/hrsolutions/talent-trends-report-2018.pdf.

Raviprolu, Anjana. “Role of Artificial Intelligence in Recruitment.” *International Journal of Engineering Technology, Management and Applied Sciences* 5, no. 4 (April 2017): 115–17.

Rode, D. "Manageral Decision Making: Normative and Descriptive Interactions." Department of Social and Decision Sciences, Carnegie Mellon University, 1997.

Rosenbaum, Eric. "Silicon Valley Is Stumped: A.I. Cannot Always Remove Bias from Hiring." CNBC. CNBC, May 30, 2018. https://www.cnbc.com/2018/05/30/silicon-valley-is-stumped-even-a-i-cannot-remove-bias-from-hiring.html.

Russo, J.edward Husted, Victoria G. Medvec, and Margaret undefined Meloy. "The Distortion of Information during Decisions." *Organizational Behavior and Human Decision Processes*66, no. 1 (1996): 102–10.

Sayegh, Lisa P., William L. Anthony, and Pamela undefined Perrewé. "Managerial Decision-Making under Crisis: The Role of Emotion in an Intuitive Decision Process." *Human Resource Management Review*14, no. 2 (2004): 179–99.

Schwartz, Alan, and Robert E. Scott. "Contract Theory and the Limits of Contract Law." *The Yale Law Journal*113, no. 3 (2003): 541–619.

Sen, Amartya. "Well-Being, Agency and Freedom: The Dewey Lectures 1984." *The Journal of Philosophy*82, no. 4 (1985): 169–221.

Shadmehr, Reza, Helen J. Huang, and Alaa A. Ahmed. "A Representation of Effort in Decision-Making and Motor Control." *Current Biology*26, no. 14 (June 2016): 1929–34.

Shaffer, Victoria A., C. Adam Probst, Edgar C. Merkle, Hal R. Arkes, and Mitchell A. Medow. "Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System?" *Medical Decision Making*33, no. 1 (2012): 108–18.

Shavell, Steven. "On the Writing and the Interpretation of Contracts." *The Journal of Law, Economics, and Organization*22, no. 2 (2006): 289–314.

Shirado, Hirokazu, and Nicholas A. Christakis. "Locally Noisy Autonomous Agents Improve Global Human Coordination in Network Experiments." *Nature*545, no. 7654 (2017): 370–74.

Silberg, Jake, and James Manyika. "Notes from the AI Frontier: Tackling Bias in AI (and in Humans." *McKinsey Global Institute*, June 2019.

Simon, Herbert A. "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics* 69, no. 1 (1955): 99–118.

Simon, Herbert A. "Making Management Decisions: the Role of Intuition and Emotion." *Academy of Management Perspectives*1, no. 1 (February 1987): 57–64.

Simon, Herbert A. *The New Science of Management Decision*. New York: Harper & Row, 1960.

Simon, Herbert Alexander. *Models of Man Social and Rational; Mathematical Essays on Rational Human Behavior*. New York: Wiley, 1957.

Skyrms, Brian. *Evolution of the Social Contract*. Cambridge: Cambridge University Press, 2014.

Slovic, P, B Fischhoff, and S Lichtenstein. "Behavioral Decision Theory." *Annual Review of Psychology*28, no. 1 (1977): 1–39.

Smith, Aaron, and Monica Anderson. "Americans' Attitudes about a Future Where Robots and Computers Do Many Human Jobs." Pew Research Center: Internet, Science & Tech. Pew Research Center, October 4, 2017. https://www.pewresearch.org/internet/2017/10/04/americans-attitudes-toward-a-future-in-which-robots-and-computers-can-do-many-human-jobs/.

Smith, Chris, Brian McGuire, Ting Huang, and Gary Yang. "The History of Artificial Intelligence," 2006.

Smith, Derek, Justin De Passos, and Rafieqah Isaacs. "How IT Project Managers Cope with Stress." *Proceedings of the 2010 Special Interest Group on Management Information Systems 48th Annual Conference on Computer Personnel Research on Computer Personnel Research - SIGMIS-CPR 10*, May 2010.

Stanton, J. "Hazardous Work - Is the Risk Obvious?." *Manufacturers' Monthly*, 2006, 18–19.

Subrahmanian, V.s., Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. "The DARPA Twitter Bot Challenge." *Computer*49, no. 6 (2016): 38–46. https://doi.org/10.1109/mc.2016.183.

Sullivan, John. "Job Interviews Have Become Predictable and Ineffective – Here Are 10 Ways to Change That." LinkedIn Talent Blog. LinkedIn, March 28, 2017. https://business.linkedin.com/talent-solutions/blog/interview-questions/2017/job-interviews-have-become-predictable-and-ineffective-here-are-10-ways-to-change-that.

Sweeney, Latanya. "Discrimination in Online Ad Delivery." *Queue*11, no. 3 (January 2013).

Tandon, Lochan Sharma, Parikshit Joshi, and Richa Kinger Rastogi. "Understanding the Scope of Artificial Intelligence in Human Resource Management Processes - A Theoretical Perspective." *International Journal of Business and Administration Research Review*, 2017.

Thaler, Richard H. "From Homo Economicus to Homo Sapiens." *Journal of Economic Perspectives*14, no. 1 (2000): 133–41.

Thomaz, Andrea L., and Cynthia Breazeal. "Teachable Robots: Understanding Human Teaching Behavior to Build More Effective Robot Learners." *Artificial Intelligence*172, no. 6 (2008): 716–37.

Thorpe, Esther Kezia. "Gartner: By 2020, AI Will Create More Jobs than It Eliminates." IT PRO. IT Pro, February 5, 2018. https://www.itpro.co.uk/automation/30463/gartner-by-2020-ai-will-create-more-jobs-than-it-eliminates.

Tirole, Jean. "Cognition and Incomplete Contracts." *American Economic Review*99, no. 1 (2009): 265–94.

Trigeorgis, Lenos. "The Nature of Option Interactions and the Valuation of Investments with Multiple Real Options." *The Journal of Financial and Quantitative Analysis*28, no. 1 (1993): 1–20.

Turing, Alan M. "Computing Machinery And Intelligence." *Mind*LIX, no. 236 (January 1950): 433–60.

Tushman, Michael L., and David A. Nadler. "Information Processing as an Integrating Concept in Organizational Design." *Academy of Management Review*3, no. 3 (1978): 613–24.

Tversky, Amos, and Daniel Kahneman. "Judgment Under Uncertainty: Heuristics And Biases." *Science*, 1974.

Tversky, Amos, and Daniel Kahneman. "Rational Choice and the Framing of Decisions." *The Journal of Business*, no. 59 (1986).

Varian, Hal. "Artificial Intelligence, Economics, and Industrial Organization." *NBER WORKING PAPER SERIES*, July 2018. https://doi.org/10.3386/w24839.

Von Krogh, Georg "Artificial Intelligence in Organizations: New Opportunities for Phenomenon-Based Theorizing." *Academy of Management Discoveries*4, no. 4 (2018): 404–9.

Von Winterfeldt, Detlof, and Ward Edwards. *Decision Analysis and Behavioral Research*. Cambridge: Cambridge Univ. Press, 1993.

Westlund, Jacqueline M. Kory, Hae Won Park, Randi Williams, and Cynthia Breazeal. "Measuring Young Childrens Long-Term Relationships with Social Robots." *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 2018.

Williamson, Oliver E. "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives." *Administrative Science Quarterly*36, no. 2 (1991): 3–37.

Williamson, Oliver E. *Markets and Hierarchies*. New York: Free Press, 1976.

Wisskirchen, Gerlind, Blandine Thibault Biacabe, Ulrich Bormann, Annemarie Muntz, Gunda Niehaus, Guillermo Jiménez Soler, and Beatrice von Brauchitsch. “Artificial Intelligence and Robotics and Their Impact on the Workplace.” *IBA Global Employment Institute*, April 2017.

Witte, Eberhard. “Field Research on Complex Decision-Making Processes – The Phase Theorem.” *International Studies of Management & Organization*2, no. 2 (1972): 156–82.

World Economic Forum. *The Future of Jobs Report 2018. (2018). http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf.*

Yano, Kazuo. “How Artificial Intelligence Will Change HR.” *People & Strategy* 40, no. 3 (2017).

Yin, Robert K. *Case Study Research and Application: Design and Methods*. 6th ed. Los Angeles, Calif.: Sage, 2018.

Zeira, Joseph. “Workers, Machines, and Economic Growth.” *The Quarterly Journal of Economics*113, no. 4 (November 1998): 1091–1117.

Dipartimento di IMPRESA E MANAGEMENT

Cattedra: INTERNATIONAL MANAGEMENT

ARTIFICIAL INTELLIGENCE AS AN INNOVATIVE DECISION-MAKING TOOL

RELATORE

Prof. LUIGI MARENGO

CANDIDATO: GIULIA BATTISTA

MATR. 710551

CORRELATORE

Prof. ANDREA FILIPPETTI

ANNO ACCADEMICO 2019/2020

Table of Contents

Introduction

The current world finds itself in an era of significant technological advancements along with considerable amounts of economic instability. As the world is becoming increasingly unstable, an effective managerial decision-making process is becoming fundamental for organizations. However, because of the turbulent situation, the decision-making process is becoming more complicated and the weight attached to the taken decisions is increasing. As the process needs to be performed as efficiently as possible, this thesis proposes and evaluates the benefits and effects that artificial intelligence (AI) tools can have on the process, with a particular focus on recruitment process. The recruitment process has been selected as the main focus of the case study as it requires multiple levels of decisions both on the employee and on the candidates' side. Hence, the process is optimal to observe the practical implications that AI can have and its current technological level.

1.Presentation of the Theory

The first three chapter of the thesis aim at providing theoretical background on the topic of decision-making, artificial intelligence, and the impact that AI has on humans, organizations and on the economy.

1.1. Chapter I: Biases and Limited Rationality in Human Behavior

Chapter one, Biases and Limited Rationality in Human Behavior, focuses on providing background knowledge on the decision-making process and on the obstacles that can arise when performed by an individual. The chapter, after having provided a theoretical background on behavioral economics and on the decision-making process, analyzes the impacts of the current unpredictable environment on decision-making. Since the environment is unstable, it requires an increased pace in decision-making, thus creating difficulties in information gathering and analysis which in turn makes the process more difficult to perform. Moreover, uncertainty also threatens rationality, which is core in decision-making under uncertainty, as it prevents managers to distort reality under stressful conditions. Unfortunately, rationality is often expected by managers but hardly achieved as many executives base their decisions on judgmental responses arising from previous experience. As experience is largely subjective, different managers' approaches will vary widely and will lead to different outcomes, which can be positive in some environments and detrimental in others. In addition, decision-making in uncertain environments is also shaped by the managers' stress levels since stress is a powerful force that diverts behavior as it can create nonproductive responses and might lead managers to take decisions that will benefit personal comfort instead of the organization (Smith, Passos & Isaacs, 2010).

Another fundamental component to address when analyzing the impact that human agents have on decision-making is the presence of biases within the process. Individuals have limitations in information gathering and analysis, therefore decision-makers rely on heuristics and biases to simplify their decision-making process, especially under uncertainty (Tversky & Kahneman, 1974). Heuristics can impede rationality as they can create baseless deviations from rationality (Tversky & Kahneman, 1986), thus leading individuals to adopt alternatives that they consider satisfactory instead of adopting optimal solutions. In addition to heuristics, the major biases that have been found to hinder the decision-making process are the overconfidence bias, the anchoring bias, the halo effect and the confirmatory bias. Under uncertainty individuals make decisions following a decision tree (Raiffa, 1968). however, this process is highly influenced by biases. Because of biases individuals tend to evaluate only one alternative instead of all the possible ones, hence transforming the process in a binary choice. Moreover, biases prevent individuals from making rational decisions, as the outcomes are subject to the influence of biases, and also affect the final judgment of the possible alternatives. Furthermore, the decision-making process is also affected by noise which results in difficulties in information collection and can also interfere with the ability of professional judgment and consequently alter the decision-making process. Noise can be dialed down when decisions are immediately followed by clear feedback, however in unstable environments it is rarely the case, thus noise is rarely dialed down in uncertain settings. Briefly stated human emotions and experience can have a significant impact on the decision-making process overall by hindering optimal decision-making.

1.2. Chapter II: Organizational Relationships With Artificial Intelligence and Their Impact on the Economy

An alternative to reduce the impact of biases and noise could be to use artificial intelligence as an aid to the decision-making process. Chapter two, Organizational Relationships with Artificial Intelligence and Their Impact on Economy, provides background theory on what is meant with the term AI, while also underlining the impact that its adoption has on the organizations and on the economy. The term AI has multiple possible definitions but the essence of AI is clear and coherent between most scholars and it is that AI is the ability to make appropriate generalizations at the right time based on limited available data.

Because of its grate potential, AI is becoming increasingly introduced in daily tasks and is also affecting behavioral economics. The three major implications that AI has on behavioral economics are that it can be used as tool to identify variables that affect behaviors, that it can help individuals understand common limitations of human cognition through the difficulties of implementation, and lastly that it is essential to understand how human limitations can be overcome (Camerer, 2019). Through these implications AI affects human behavior, and studies (March, 2019) have shown that human players behave more selfishly when playing against computer players.

In addition to being increasingly used for daily tasks, organizations are also increasingly adopting AI. The increased adoption of AI from organizations is a direct consequence of the significant progress in technology in the past decades. Moreover, many technologies are provided under open-source licenses thus allowing to customize the technology thus increasing its adoption. In addition, the price of the hardware is decreasing, thus making the technology more affordable (Von Krogh, 2018).

Technological advancements, in addition to contributing to the creation of new types of firms such as the Information Age Organization, also allow for new opportunities to collect and leverage data. These opportunities are changing the decision-making process as managers are starting to rely less on intuition and more on data. Even though many organizations have started to adopt data-driven decision-making (DDD) as it allows for a more efficient decision-making process, its diffusion is uneven. DDD is most frequently adopted by organizations that show three key characteristics that are correct size, high levels of complementary investments and awareness of these practices within implementation (Brynjolfsson & McElheran, 2016). If companies are able to correctly implement DDD their productivity will increase by 3% (*ibid*), thus organizations should aim at successfully implementing this managerial practice. Moreover, DDD is always associated with improved performance since early adopters will be able to remain ahead of competitors who do not immediately realized the benefits of DDD, thus generating performance differences while allowing to gain a competitive advantage, at least in the short term.

Another key variable to observe in order to assess the impact of AI is its alignment. When designing AI agents, misalignment can cause economic losses associated with the delegation of decisional power. Designers are challenged with the task of achieving intended objectives, while acknowledging the limitations that arise from transforming those goals into implementable algorithms to guide an artificial agent's behavior. Another main cause of AI misalignment is incomplete contracting as complete contracting is routinely impossible and costly (Hadfield-Menell & Hadfield, 2019). A contract can be perceived an implemented reward structure for an individual, hence non-contractibility within the AI context can be thought of as a learning problem that cannot be solved with the currently known techniques. AI rewards might fail to address all the possible circumstances as algorithm designers are not able to think about all the possibilities. Moreover, contracts might be costly to enforce as problems with AI alignment occur because of the gap between the reward function and human actual values since there are engineering limitations. However, incomplete contracts might also be voluntarily implemented because of strategic behavior (*ibid*), as contract designers might want to protect private information.

In addition to impacting organizations, AI is also impacting the economy, especially the labor market. One of the concerns that people have with respect to AI and labor is that it might substitute capital for labor in prediction tasks. Being AI very efficient in prediction, many tasks, such as demand forecasting, are becoming increasingly performed by machines. Moreover, some tasks are being transformed into prediction ones in order to be performed by AI agents as it would reduce prediction costs when compared to that human agents. Secondly, AI is augmenting the relative returns to capital versus labor in complementary decision tasks as machines have a faster reaction time than humans. In addition, AI can make more accurate predictions than individuals since it can process a bigger quantity of data in a lower amount of time. Another impact that AI can have on human labor is to reduce uncertainty and enable new decisions that were not possible before as automated prediction technology can allow to perform tasks which would otherwise be unfeasible because of high uncertainty (Agrawal, Gans & Goldfarb, 2019)

AI is also impacting economic growth, and during the last decades it became a fundamental tool for progress. There are two major forces that affect the share of automated economy. First, as the fraction of automated goods increases, also the share of automated goods in GDP increase along with capital share. Secondly, as the ratio of aggregate capital to labor increases the capital share and the value of the automated sector as share of GDP increases. As more sectors become automated, the share of automated goods and capital will increase. However, since automated goods are subject to faster growth, their price will decline as will their GDP (Aghion, Jones & Jones, 2017). AI also has macroeconomic impacts which are dependent on the firms' behavior. Research highlights two contrary effects of competition and innovation-led growth. The first effect is that intense market competition, or imitation threat, will induce competing firms at the technological frontier to innovate to beat competition. However, the second effect states that because of more intense competition, firms that fall behind the technological frontier will be discouraged to innovate and will not catch up with frontier firms. The prevalent effect will depend on the level of advancement and competition of the economy. Other macroeconomic impacts that AI has,are sectoral reallocation and knowledge diffusion (Baslandze, 2016). Knowledge diffusion is a direct consequence of the IT revolution, which in turn caused sectoral reallocation from sectors that do not heavily rely on technological externalities to sectors that do. IT has two effects on innovative incentives. On one hand firms can learn from each other and therefore benefit from knowledge diffusion, On the other increased access to external knowledge can increase the scope for business stealing.

1.3. Chapter III: Artificial Intelligence As A Tool For Improved Decision-Making

Once theoretical knowledge on the decision-making process and on the impacts that AI has on individuals and on the economy has been presented, chapter three, Artificial Intelligence as a Tool for Improved Decision-Making, goes more in depth to assess the direct impacts of AI on the decision-making process. Managers make frequent deviations form perfectly rational behavior which impact the firm's value as well as the economy overall. Advancements in AI have allowed for potential to reduce these errors. Research has shown that the effectiveness of humans compared to that of algorithms is lower, in fact algorithms outperform human forecasters by 10% on average (Grove et al., 2000). In addition, algorithms are better predictors of decision-making as individuals are subjected to bias and judgments when making decisions. In many cases AI can contribute to reduce human subjective interpretation of data as algorithms are programmed to consider only those variables that contribute to improve the predictive accuracy based on the data used. In addition, AI can help to improve decision-making by making it fairer while also allowing for through examination if errors occur as machines can be easily opened and interrogated. Moreover, solutions can be hard to find and implement for human agents, but machine learning can be a great tool to alleviate pressure on human generation of solutions.

Although AI can have a variety of advantages, it also presents concerns linked to its diffusion. As of today, there are an increasing number of concerns related to AI's expansions which are slowing AI's adoption from both individuals and organizations while contributing to increasing algorithm aversion. Some of the most common causes of algorithm aversion include the wish for perfect forecasters, the inability of algorithms to learn, and the notion that algorithms are incapable of considering specific targets and ethical decisions. Other causes that contribute to reduce faith in AI are linked to the fact that individuals tend to rely more on human forecasters than on algorithmic ones even though algorithmic models are more correct than humans even if mistakes are made. Moreover, when individuals witness another individual make a mistake, their tendency to rely on human agents does not decrease. However, when mistakes made by machines, they are weighted more heavily, thus creating a trend to abandon algorithms faster when mistakes are made (Dietvorst et al., 2015).

Another reason why individuals might not completely rely on algorithms is because they can be subject to bias which stems from the underlying data (Kleinberg et al., 2019). The data on which algorithms are trained and designed to work can contain human decisions and might reflect social disparities, which are then transferred to the algorithm and applied when used. Since algorithms can reflect social concerns, individuals do not perceive algorithm as being completely reliable. In order to minimize ethical concerns, framework such as the Society In The Loop (SITL) framework (Rahwan, 2017), have been developed in order to integrate the value of society within the algorithm. However, although such a framework has been established, it is not straightforward to apply it since society presents some barriers. One of the obstacles is represented by the cultural discrepancy between engineers and humans as it is easy to identify in which ways constitutional rights are violated, but it is not straightforward to code this knowledge through engineering. Moreover, difficulties in implementation also arise because of the existence of negative externalities resulting from algorithms.

In addition to ethic concerns also safety concerns are increasing as individuals are scared of what machines can do. However, differently from ethical concerns, there are a variety of approaches to build safe and robust machine learning systems thus decreasing this concern. The other major concerns that society has with AI are cognitive conflicts, which arise since the designer is not always able to full specify the objective to the machine thus causing misalignment between the machine and the organization, along with fear that machines will take over. The latter concern however has been proved to be unrealistic as individuals perceive AI as the one seen in movies when in reality AI is not even close to that level of technological advancemen, therefore excluding the possibility of a takeover in the near future.

The last section of the chapter focuses on machines behavior as there exist three scales of inquiry for it, which are namely individual machines, collective machines, and groups of machines with are part of a social environment along with human individuals in hybrid or heterogenous systems. The most interesting scale of inquiry is the last one as it allows for human machine co-behavior. The majority of AI systems work in environments in which humans and machines co-exist in hybrid systems Within these systems it has been observed that machines can either enhance human efficiency or directly replace humans as they are more efficient. Because of the uncertainty of what the real impact of machines would be in human-machine environments, scholars started to examine these interactions in laboratories and observed that bots are able to increase human coordination and that machines can cooperate with humans as much as other humans could.

2. Chapter IV: Case Study: Artificial Intelligence And The Recruitment Process

After having provided knowledge and theory on the topics of decision-making and AI, chapter four's objective is to apply the previously mentioned literature to a practical case: the adoption of AI within the recruitment decision-making process.

The recruitment industry is facing a significant issue as the traditional hiring process, where resumes and interviews are used, has found to be ineffective because of the growth of the internet which could yield unreliable results. To improve the recruitment process, many companies have started to incorporate recruitment software within their process in order to scan more candidates faster and to acquire more talent. Some recruitment software also adopt AI within their technology, but being a novel implementation not much research has been performed to understand whether or not AI is the best tool to simply the recruitment process. Hence, the study conducted aims at understanding the state of AI that is available to companies and consequently the use that companies can make of AI within Human Resource Management Services, with a particular focus on recruitment, and to understand its practical implications on decision-making. In addition, the study aims at providing an encompassing overview on the topic by delivering insights on both the benefits and the detrimental effects that AI can have on the recruitment process.

In order to develop a qualitative analysis on the topic, thematic analysis was used. The themes were taken from arguments presented in literature and touch upon the influence that AI can have on recruitment and the impact that is has on candidates and employees. The analyzed themes are as follows: Limitations and opportunities to overcome them; Evolution of the recruiter's role; Bias reduction.

2.1. Literature Review

The analyzed themes are further discussed by presenting a detailed literature review. With respect to Limitations and opportunities to overcome them, the literature emphasizes how the topic of AI is surrounded by worries and concerns. Studies (Pew Research Center, 2017) have found that individuals tend to reflect more worries than positive feelings when questioned about AI, mainly because of the limited availability of data to train the algorithms (Raviprolu, 2017). The shortage of available data is caused by increased data protection and higher costs of acquiring it, along with policies such as GDPR which can prevent access to some datasets. Moreover, concerns are also shared by supporters of AI such as Wisskirchen et al. (2017), who argue that even the most advanced technologies can make mistakes. In addition, some concerns are raised specifically for the recruitment process since AI agents do not have strong communications abilities, therefore are incapable of automating the entire process. Despite many concerns however, a study from Korn Ferry (2018) observes that, even though professionals have less trust in AI than in human recruiters, nearly 72% argues that AI should be used within the recruitment process.

Contrarily to the literature on constraints, not much literature exists on how the recruiter's role will be affected by AI as the literature mostly focuses on the general workforce. According to the Future of Job Report 2018 (World Economic Forum, 2018), by 2020, 75 million of jobs will be displaced as AI takes over the routine aspects of work and substituted with different versions of the original jobs. Moreover, employees' tasks will change since it is becoming fundamental to have an AI ready workforce by upskilling non-AI workers (Meister, 2019). Moreover, Meister also argues that, specifically for the HR field, it is fundamental to change performance management and to allow HR roles to understand how to use AI across the employee life cycle.

A literature review on bias reduction is also addressed in chapter four. Campolo et al. (2017) observe that ML programs use existing data points in order to make decisions and that if those datapoints are biased, it could result in transferring cultural, gender and other types of discrimination to the algorithm. Further literature (Florentine, 2016) also highlights how discrimination could arise when using algorithms as they might pick up on gender related clues, even involuntarily. A contrasting view however is provided by Wisskirchen et al. (2017) who suggest that using AI in recruitment will remove all types of biases as it allows to purely focus on data without emotion and sympathies interfering. Moreover, according to Randstad (2018) approximately 70% of human capital and C-suite leaders suggests technology is improving hiring decisions.

2.2. Research Aim and Objectives

The aim of the study is to perform a qualitative research to understand the current state of the AI technologies and to explore the extent to which it is used within the recruitment process and the changes it can bring to the recruiter's role. Moreover, the study aims at understanding which factors are promoting or impeding the adoption of AI from recruiting companies.

The cognitive objectives of the study are as follows:

5. Provide insight on how companies have overcome the concerns and successfully implemented AI (Theme 1)
6. Observe which concerns are raised by those companies that do not use AI within their recruitment software (Theme 1)
7. Investigate whether AI changes the recruiter's role once adopted (Theme 2)
8. Investigate if AI can be a valid tool for bias reduction (Theme 3)

2.3. Methodology

In order to conduct the study, 6 recruitment software providers were interviewed, both users and non-users of AI technologies, in order to provide a full spectrum of opinions on the technology. Participants were subjected to semi-structured, shorter case study type interviews, in order for the interviewees to provide insightful commentary on the topic while allowing for deviations in those areas that were brought up but not yet researched on the topic of AI and recruitment. Interviews have been used as they are a useful tool to gather insights on the participants' perspective and each respondent gave permission to be featured in this research. The interviews with the representatives of Altamira, Allibo, Cornerstone and In-Recruiting were conducted in Italian and later translated to English.

Semi-structured interviews however can present pitfalls. The first is the possible emergence of the interviewer's bias because of the lack of structure along with the possible problem of inaccurate articulation of the questions (Yin, 2018). Because of this, extreme care has been adopted in order not to introduce bias before a responded could answer (Bradburn et al, 2004). The questions posed to the participants were structured in such a way that the respondents could answer based on their expertise on the topic. In addition, to further counter bias, a final review of the questions and answers was performed and any answer that might have been a consequence of the question's phrasing was removed from the findings of the study.

2.4. Participants

Interviews were conducted with 6 experts via phone and meeting platforms. Stefano Michetti, CEO at Allibo, having directly experienced the benefits and disadvantages of adopting AI. Emanuele Rifaldi, Account Manager at Altamira, having contact with many professionals within the HR field and participating at events with a focus on AI. Markellos Diorinos, CEO and Co-Founder at Bryq, having knowledge in providing an efficient recruiting software without the aid of AI. Francesca De Cristofaro, Senior Solution Consultant EMEA at Cornerstone, having relevant insights on the benefits that AI has on the recruiting process. Matteo Cocciardo, CEO and Co-Founder at In-Recruiting, having know-how with respect to the implementation of AI within the recruitment process. Tim Peltier, Head of Marketing and Communication at Manatal, having insights on AI adopting software and how it is perceived by their customers. These companies provide recruitment software that have significantly contributed to change the recruitment processes, therefore could provide knowledgeable insight and experience on the topic of AI adoption within the recruiting process and its potential benefits. The participants have been selected based on Capterra's (2020) list of recruitment software along with a through Google search. Moreover, the chosen software reflect the features that the current recruiting systems can offer to the companies that adopt them.

2.5. Findings and Results

To analyze the transcripts of the interviews, coding was used. The three themes of Limitations and opportunities to overcome them (Theme 1); Evolution of the recruiter's role (Theme 2); and Bias reduction (Theme 3) that were highlighted in the literature review have been used indirectly to code the interviews. The codes have been inspired from the original themes and therefore have been linked to them. The coding process has enabled a more effective analysis of the insights and knowledge provided by the interviewees (Yin, 2018). For these interviews the considered codes and the linked themes are as follows:

CODE	DISCUSSION	THEME
Capability	Which parts of the recruitment process are performed by AI?	Theme 1
Benefits	What are the advantages of using AI?	Theme 1, Theme 2
Concerns	Have your customers experienced concerns with the use of AI?	Theme 1
Limitations	What are your concerns with respect to the adoption of AI?	Theme 1
Bias	Can an algorithm reduce bias?	Theme 3
Supervision	Does an algorithm require human supervision?	Theme 2, Theme 3
Recruiters	How did the recruiter's role change with the adoption of AI?	Theme 2, Theme 3
Data	Is there an adequate amount of available and credible data?	Theme 1; Theme 3

Moreover, a summary table of the interviews is provided below

Questions	Software					
	IN-RECRUITING	BRYQ	ALLIBO	MANATAL	ALTAMIRA	CORNERSTONE
Which tasks are perfomed by AI?	CV acquisition phase and canidates' pre-screening phase	None	Had previous experience but as of today none	Candidate screening; suggestion of candidates; scoring; data mapping	No use of AI within the platform	Automatic identification of skills within resumes; CV passing; identification of skills based on interests and experiences
Any prejudice from clients?	Yes, because of AI's probabilistic nature and because of other company's ineffective use of AI	N/A	N/A	Not until now	N/A	Doubts on data privacy and security
Is an algorithm prejudice-free?	Depends on its training as it requires unbiased datasets	No since it is trained on biased data	Depends on its training	Depends on training: the larger the dataset on which it is trained, the less likely for a small minority to influence the model	Depends on who trains it and on what datasets it is trained	Yes ,provided it is sophisticated enough and is trained on non-biased data
Can an algorithm can reduce human bias?	Absolutely but the algorithm needs to be trained correctly	No because it is trained on biased datasets	Depends on how the models are trained	Definitely, if trained on an appropriate dataset	Not always, as it depends on datasets	Can help recruiters to have a more objective view on the process and helps them to make objective decisions, provided it is well trained
Is human supervision required?	As of today, yes	AI cannot replace humans	Yes, so it defeats its purpose	Yes	Yes	Yes, and is desirable
Main benefits of using AI	Economic; image; time, information availability; helps to overcome the recruiter's lack of multidomain knowledge	Can help in some recruiting phases where there is substantial data processing to be done provided it is trained on a bias-free dataset and is explainable	No benefits within the hiring process, but can be used to analyze the information on the uploaded Ids and is successfully used by job-posting websited (Ex. LinkedIn)	Accuracy in identifying skills; more efficiency in talent acquisition; professional can work faster	Can help during the pre-screening phase to reduce time taken; can help to understand candidate's language skills	Increases employee engagement; provides a wider view; automates time consuming activities; speeds the process; increases the talent pool
Why is AI is not a good tool for bias correction?	N/A	Trained on biased datasets; no explainability; lack of scientific easerch	Output validity cannot be assessed; can create "dangerous" outcomes; can be used inappropriately	N/A	Cannot be developed in-house (most of the time); high costs and time consuming to adopt	N/A
Are there any limitations or concerns with AI?	There should not be any software that allows AI to take decisions alone at the moment	Trained on biased datasets; lacks explainability; lacks scientific foundation	Leads to hiring "clones"(ethic concern); insufficient quantity of data to tailor algorithms; GDPR concerns; difficult to understand if AI is working properly	Needs to be carefully trained	The market is not thirsty for AI; needs training otherwise is useless; explainability concerns; ethical concerns	N/A
What is missing to ensure a successful implementation of AI?	N/A	Data and science	Clear idea on behalf of the employers of what candidate they are looking for	N/A	Financial means; effective training	N/A
How did the recruiter's role change after AI?	Recruiters can now focus on value adding activities while AI takes care of low value adding ones	By either using AI or science-backed systems it would improve	Did not change	Recruiter is now able to focus more on human oriented tasks	Did not change	Recruters are more efficient and faster

2.6. Discussion

Three themes have been analyzed in order to gain an insight on what impacts AI can have within the process. The analyzed themes: Limitations and opportunities to overcome them, the Evolution of the recruiter's role and Bias reduction, have led to interesting considerations with respect to the topic. From the literature and the conducted interviews, it emerged that there are multiple concerns linked with AI as it can be costly to develop, it has no explainability and as it needs to be trained efficiently in order for it to work. Although some of the concerns cannot be bridged, as demonstrated by the fact that AI is not able to carry out the entire recruitment process by itself, others, such as the concern of machines taking over human agents, can be minimized as a human recruiter will always be responsible for making the final decision. In addition, the crucial role played by the human recruiter within the process ensures that the recruiter's role will still be present in the future and that it will only slightly change since AI models can take care of tedious tasks for them but the recruiter will always have to focus on human related tasks. On the other hand, the required presence of a human agent can be detrimental when looking at AI as a bias correction tool. The initial phases of the recruitment process such as screening and resumes acquisition, will be completely bias-free, if the algorithm is correctly trained, but since the final decisions are taken by human agents their implicit biases could affect the process thus ensuring a bias-free process only at the initial stages.

Furthermore, another important factor to guarantee a bias-free assessment is efficient training. If an algorithm is trained on improper datasets or by companies that have implicit prejudices, the algorithm will turn out to be completely biased and therefore damage the integrity of the process and also the company's image. Therefore, as of today, AI can be a valuable tool to reduce bias in the initial steps of the recruitment process, as it can simplify the recruiter's decision-making, but the technology is not advanced enough to ensure an entirely automated recruitment process. Therefore, the adoption of AI within the decision-making process cannot guarantee a bias-free and efficient process from start to finish.

Moreover, the conducted study provides insights on the current level of the technology and on how it is perceived by experts in the recruitment field, but could be generalized to many more areas as the basic benefits and limitations of the technology are the same independently of the field. However, the study has some limitations such as the fact that it is based solely on literature and interviews and not on practical observations. Hence, for further studies it would be helpful to directly observe the effects that this technology has on companies through first-person observations.

Conclusion

This thesis aims at analyzing whether or not the currently available AI tools can be valid instruments to improve the decision-making process. After examining literature concerning the main reasons why human agents divert from optimal decision-making, a study was conducted by interviewing 6 experts of the recruitment field. The recruitment filed was chosen since the recruitment process requires different levels of decision-making on both the recruiters and the applicants' part, hence it is of great use to understand the practical implications of the technology. By interviewing experts of the field, some supporting AI and other against it, a full spectrum of perspectives was gathered. The interviewees were asked questions with the aim of obtaining answer to questions related to the following research objectives: understanding whether or not the concerns linked to AI adoption can be bridged; gaining insights on how the recruiter role will evolve with the adoption of AI tools; and assessing if AI can be a valid tool to reduce bias within the decision-making process. Within the case study, three themes have been analyzed in order to understand the impacts AI can have within the process. The analyzed themes: Limitations and opportunities to overcome them, Evolution of the recruiter's role and Bias reduction, have led to interesting considerations with respect to the topic.

The literature and the conducted interviews highlighted a variety of concerns related to AI adoption such as its high development costs, its lack of explainability and the importance of algorithm training, which if done wrong will jeopardize the efficacy of the tool. Although some of the presented concerns cannot be bridged, as highlighted by the fact that AI is used as an aid for human agents and not as an independent component, AI can still be valuable within the recruitment process. Some of the presented concerns can be minimized by the presence of human agents who are supervising the process and since the human recruiter will always be responsible for making the final decision. In addition, the crucial role played by the human recruiter within the process will ensure the durability of the recruiter's role which will not be eliminated, but will be improved, by the presence of AI. Since AI models will take care of tedious tasks, the recruiters will have more time to carry out human-related ones, thus being more able to focus on a lower number of repetitive tasks while speeding-up the process.

On the other hand, the required presence of a human agent can diminish the potential of AI as a bias correction tool. Even though the initial phases will be completely bias-free since AI will solely focus on data, provided that the algorithm is correctly trained, the final decisions will always be made by a human agent. As highlighted both in the literature and through the interviews, human agents will always engage in subjective decision-making and are often subject to implicit biases which will affect the process. Thus, even with AI, a completely bias-free process is not entirely possible as it will be free of bias only in the initial stages. Moreover, another important factor to guarantee a bias-free assessment is efficient training. If the used datasets are compromised, also the algorithm will be, thus making the entire process inefficient.

Since AI is not technologically advanced enough to completely automize the entire recruitment process, it can be a valid tool to reduce bias only in the initial phases of the recruitment process. In the starting stages, in which AI is efficiently used, the algorithm will base its decision on objective data and consequently will suggest candidates based on unbiased terms. However, the need to delegate the final decisions to a human agent does not allow for an entirely bias-free assessment as the individual will make the final decisions which will be affected by subjective components, thus re-inserting bias within the process.

The study also gives a brief insight on the impact that the technology can have on the candidates. Since the topic is not frequently addresses by the literature, further studies could address the effects that the adoption of AI has on candidates and how they perceive the recruitment process when being aware that they are being assessed by a machine.

References

Aghion, Philippe, Benjamin Jones, and Charles Jones. "Artificial Intelligence and Economic Growth." *NBER Working Paper Series*, October 2017.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction." *NBER Working Paper Series*, 2019.

Baslandze, Salome. "The Role of the IT Revolution in Knowledge Diffusion, Innovation and Reallocation." EIEF manuscript, 2016.

Bradburn, Norman M., Seymour Sudman, and Brian Wansink. *Asking Questions: the Definitive Guide to Questionnaire Design*. San Francisco: Jossey-Bass, 2004.

Brynjolfsson, Erik, and Kristina Mcelheran. "Data in Action: Data-Driven Decision Making in U.S. Manufacturing." *SSRN Electronic Journal*, 2016.

Camerer, C.F. "Artificial Intelligence and Behavioral Economics." *The Economics of Artificial Intelligence*, 2019, 587–610.

Campolo, Alex, Madelyn Sanfilippo, Meredith Whittaker, and Kate Crawford. "AI Now 2017 Report," 2017.

Capterra. "Software per Selezione Del Personale." Capterra, 2020. https://www.capterra.it/directory/30602/recruiting/software.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err." *Journal of Experimental Psychology: General*144, no. 1 (2015): 114–26.

Florentine, Sharon. "How Artificial Intelligence Can Eliminate Bias in Hiring." CIO. CIO, December 22, 2016. https://www.cio.com/article/3152798/how-artificial-intelligence-can-eliminate-bias-in-hiring.html.

Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson. "Clinical versus Mechanical Prediction: A Meta-Analysis." *Psychological Assessment*12, no. 1 (2000): 19–30.

Hadfield-Menell, Dylan, and Gillian K. Hadfield. "Incomplete Contracting and AI Alignment." *SSRN Electronic Journal*, 2019

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass Sunstein. "Discrimination In The Age Of Algorithms." *Journal of Legal Analysis*101 (April 2019): 113–74.

Korn Ferry. "Putting AI in Its Place' Artificial Intelligence Should Be Part of the Recruiting Process, But It Can't Replace the Human Touch." Korn Ferry, June 28, 2018. https://www.kornferry.com/about-us/press/putting-ai-in-place-artificial-intelligence-should-be-part-of-the-recruiting-process-but-cant-replace-the-human-touch-korn-ferry-survey.

March, Christoph. "The Behavioral Economics of Artificial Intelligence: Lessons from Experiments with Computer Players." *BERG Working Paper Series,*154 (October 2019).

Meister, Jeanne. "Ten HR Trends In The Age Of Artificial Intelligence." Forbes. Forbes Magazine, January 9, 2019. https://www.forbes.com/sites/jeannemeister/2019/01/08/ten-hr-trends-in-the-age-of-artificial-intelligence/#7cca78783219.

Miller, Alex P. "Want Less-Biased Decisions? Use Algorithms." *Harvard Business Review*, July 2018.

Rahwan, Iyad. "Society-in-the-Loop: Programming the Algorithmic Social Contract." *Ethics and Information Technology*20, no. 1 (July 2017).

Raiffa H. Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Oxford: Addison-Wesley, 1968.

Randstad. *2018 Talent Trends Report*. (2018). https://www.randstad.it/hrsolutions/talent-trends-report-2018.pdf.

Raviprolu, Anjana. "Role of Artificial Intelligence in Recruitment." *International Journal of Engineering Technology, Management and Applied Sciences* 5, no. 4 (April 2017): 115–17.

Smith, Aaron, and Monica Anderson. "Americans' Attitudes about a Future Where Robots and Computers Do Many Human Jobs." Pew Research Center: Internet, Science & Tech. Pew Research Center, October 4, 2017. https://www.pewresearch.org/internet/2017/10/04/americans-attitudes-toward-a-future-in-which-robots-and-computers-can-do-many-human-jobs/.

Smith, Derek, Justin De Passos, and Rafieqah Isaacs. "How IT Project Managers Cope with Stress." *Proceedings of the 2010 Special Interest Group on Management Information Systems 48th Annual Conference on Computer Personnel Research on Computer Personnel Research - SIGMIS-CPR 10*, May 2010.

Tversky, Amos, and Daniel Kahneman. "Judgment Under Uncertainty: Heuristics And Biases." *Science*, 1974.

Tversky, Amos, and Daniel Kahneman. "Rational Choice and the Framing of Decisions." *The Journal of Business*, no. 59 (1986).

Von Krogh, Georg "Artificial Intelligence in Organizations: New Opportunities for Phenomenon-Based Theorizing." *Academy of Management Discoveries*4, no. 4 (2018): 404–9.

Wisskirchen, Gerlind, Blandine Thibault Biacabe, Ulrich Bormann, Annemarie Muntz, Gunda Niehaus, Guillermo Jiménez Soler, and Beatrice von Brauchitsch. "Artificial Intelligence and Robotics and Their Impact on the Workplace." *IBA Global Employment Institute*, April 2017.

World Economic Forum. *The Future of Jobs Report 2018. (2018). http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf.*

Yin, Robert K. *Case Study Research and Application: Design and Methods*. 6th ed. Los Angeles, Calif.: Sage, 2018.