Master (MSc) in Economics and Business Administration

Management of Innovation and Business Development (MIB)

Master's Thesis (CMIBO1000E) - Contract no: 22923

Exploring the Adoption of Artificial Intelligence in Venture Capital

Current Status and Opportunities

Student: Federica Scabbio (148657)

Supervisor: Francesco Di Lorenzo

Date of submission: September 4th, 2022

Number of characters: 157,980

Number of pages: 70

Abstract

Venture Capital firms have long invested in Artificial Intelligence startups, contributing to economic development and innovation in various sectors. However, it is only recently that investors started to understand Artificial Intelligence's potential for supporting their daily business operations. Following the example of technological adoption rollout that has characterized organizations in other industries, Venture Capital firms are starting to approach data usage to improve efficiency. Nonetheless, the implications of Artificial Intelligence software usage in Venture Capital are still unclear to most investment funds, which struggle to adopt new technology and integrate them with traditional practices. As it is expected that numerous Venture companies will start utilizing data approaches in the following years, this study aims at investigating Artificial Intelligence adoption in Venture Capital. This research uses grounded theory methodology to provide a qualitative analysis of technology implementation in the industry. By leveraging online sources and interviews with investors, this research provides an overview of the current adoption status of Artificial Intelligence and its future opportunities. Through statistics and real-life examples analysis, the dissertation suggests that the market is still at the early stage, characterized by few adopters. The most frequent adoption use cases currently leverage Machine Learning and Natural Language Processing models in the deal sourcing and screening stages for investing in early-stage startups. Even though the market is still characterized by low transparency, it was possible to develop a model that integrates traditional Venture Capital decision-making with data-driven approaches. The results present a best-case scenario for Artificial Intelligence implementation in Venture Capital which can be useful as a benchmark for incumbents and future new adopters in the industry. This research concludes that Artificial Intelligence can play a relevant role for Venture Capital firms, increasing operational efficiency and eventually granting a competitive edge in the market. However, due to limitations related to data scarcity and to the human nature of the Venture Capitalists' job, this dissertation also acknowledges the impossibility of Venture funds to rely solely on technology for investment decisions. Notwithstanding technological developments which will guarantee the diffusion of data-driven operational models in the future, Artificial Intelligence adoption will hardly disrupt the Venture Capital industry. However, its integration with traditional practices will create a new standard of data-augmented Venture firms that will be able to create more value.

Table of Contents

List of Abbreviations

VC = Venture Capital
AM = Asset Management
AuM = Assets under Management
PE = Private Equity
LP = Limited Partner
GP = General Partner
LPA = Limited Partnership Agreement
CRM = Customer Relationship Management
VCaaS = VC-as-a-Service
AI = Artificial Intelligence
ML = Machine Learning
DL = Deep Learning
ANN = Artificial Neural Network
ANI = Artificial Narrow Intelligence
AGI = Artificial General Intelligence
ASI = Artificial Super Intelligence
NLP = Natural Language Processing

List of Figures and Tables

Chapter 1: Introduction

1.1 Background

The Venture Capital (VC) industry forms the backbone of innovation, contributing to financing high-growth companies that have the potential to disrupt markets and industries. VC has become an essential driver of today's economic development thanks to its predominant role in helping create new products, technologies, and processes. Despite a recent setback influenced by the Ukrainian war and the Covid-19 pandemic, the VC industry has constantly been growing, deploying billions in startup funding every year. VC investment has reached $650 billion in 2021, maintaining a strong focus on high-tech sectors (Teare, 2022). Venture capitalists fuel novelty by especially deploying funds to companies leveraging cutting-edge technologies. Out of all the investment themes, Artificial Intelligence (AI) stands out thanks to its recent technological advancements.

For many years, VCs have successfully capitalized on and supported AI startups in different sectors, including healthcare, transportation, education, and financial services. AI companies attracted the largest number of VC investors in 2021, and total VC funding in AI has increased dramatically in recent years (Geronimo, 2022). AI investment grew from less than $ 3 billion in 2012 to about $ 90 billion in 2021, reporting a rise of 20% from the previous year alone (Tricot, 2021; Zhang et al., 2022).

However, it is only recently that VC investors have started to shift their focus from merely investing in AI companies to understanding AI's potential of application to their own industry (Fairview Capital, 2018). As many organizations in other economic sectors have started integrating AI within their companies to automate business processes, VCs are starting to follow their lead. Venture capitalists are timidly beginning to incorporate data-driven approaches in their operations, but few companies are consistently innovating their processes (Foy, 2021). Gartner has estimated that only 5% of VCs and early-stage investors were using AI technologies for decision-making in 2021. Nevertheless, they predict that VCs and Private Equity (PE) investors will change their traditional operations by exploiting data insights given by AI-enabled models. They forecast that more than 75% of VCs, PE companies, and early-stage investors will use AI in their decision-making process by 2025 (Rimol and Costello, 2021).

The topic has raised further interest among VCs and the general public by recently hitting the headlines with the announcement of the creation of an AI-led fund by the billionaire co-founder of Revolut, Nik Storonsky (Martin, 2022; Isaacs, 2022). According to its founder, QuantumLight Capital will revolutionize the VC industry that is currently "still stuck in the 20th century". The fund will use its proprietary quantitative decision engine that guarantees more accurate and faster decision-making (Martin, 2022; QuantumLight, 2022). Nik's fund follows the lead of existing VCs that have already started setting the path to AI adoption in the industry. Funds such as SignalFire, InReach Ventures, and EQT Ventures have been marketing themselves as data-driven companies, fueling the trend of AI usage in VC. Despite the moderate mediatic coverage achieved, there is still uncertainty about technology application in the industry. The implications of AI usage in VC are still unclear to most investment funds, which struggle to understand its benefits of implementation. However, it is expected that numerous other companies will hop on the trend over time and that adoption of this technology will gradually take place in the market. What is still dubious is how AI adoption will occur, whether technology applications will change the state-of-the-art VC investing, and if they will be able to revolutionize the industry.

1.2 Literature Review

Considering the innovative character of the topic, research on the application of AI technologies by VCs is limited and circumscribed to scarce areas of study. Literature mainly focuses on technical aspects of AI application, proposing AI and ML prediction quantitative models. Instead, the qualitative and business perspective of AI adoption in VC has not received much attention.

The salient study area in the field has concentrated on developing quantitative models to support VC investors in their decision-making process. A handful number of researchers have contributed on the technical level, constructing AI algorithms that aim to predict VC-backed venture opportunities and future events. Arroyo et al. (2019) studied the performance of different Machine Learning (ML) algorithms in predicting the progress of ventures in a three-year time window using data from Crunchbase. Based on their positive results, they concluded that data could help baseline screening of VC investors. Mishra et al. (2017) proposed a model that expects to improve VC decision-making by building a causal map of the investment decision using Bayesian networks. Ross et al. (2021) developed a ML model to predict the upcoming unicorns to invest in, drawing on

Crunchbase data and reporting an 89% accuracy rate. Caruso et al. (2015) predicted valuation increases within financing rounds for startups through a ML model using data from Pitchbook.

The still underdeveloped research field connected to VC draws upon a set of more generic research studies on the development of AI models to predict success. The following contributions have committed to proving the feasibility of predicting small companies' success using AI and ML algorithms by considering different variables. Even though not directly connected to the theme of VC, investors can still leverage these studies. Ragothaman et al. (2003) presented an expert system capable of predicting the acquisition of companies with a success rate of 70%. Yankov et al. (2018) compared different ML methods to predict the success of startups from a questionnaire. Böhm et al. (2017) used cluster analysis and Support Vector Machines to classify business model performance, indicating the survival of a small company with 83.6% accuracy. Xiang et al. (2012) used a ML framework to analyze people profiles and news articles from Crunchbase to predict acquisitions. Wei et al. (2009) used patent data to support M&A prediction. Antretter et al. (2019) showed that it is possible to predict company success with 76% accuracy by analyzing Twitter content.

Even though the studies mentioned above signal the possibility of predicting ventures' success using AI algorithms, they are only valid on the theoretical level. As the proposed models have not been further tested under changing assumptions, nor have they been practically experimented by VCs, the concrete application remains uncertain. Moreover, some of the analyses are performed on a too-small sample (Ragothaman et al., 2003; Yankov, 2018; Böhm et al., 2017), while others might be affected by information availability biases (Wei et al., 2009; Xiang et al.,2012; Arroyo et al., 2019). Finally, since the VC industry is a dynamic environment not only defined by a few companies' variables but also influenced by greater economic forces, it is impossible to practically generalize outcomes to ascertain unconditioned AI predictions of venture success from a limited set of factors. Nonetheless, these studies are still relevant because they highlight the researchers' interest in the topic and can still be used as inspiration for practical, real-life AI algorithms development in VC.

The qualitative studies that have focused on the topic under investigation are very scarce. Schmidt (2018) studies opportunities for AI application in VCs' decision-making process by using information gathered from interviews with stakeholders in the sector. The findings reveal that AI can be applied to all the steps of the decision-making process but that it is currently primarily implemented at the beginning of the value chain, during deal sourcing and deal screening. Moreover, researchers have attempted to map the implementation status of AI technologies in the VC market. Corea (2018) reports that, to his knowledge, there were only 13 VCs actively using AI in 2018. A broader approach is taken by Trocha (2019), who identifies a sample of 83 VCs, accelerators, and PE firms that used data analysis in their processes. However, due to the fast pace of change in the venture landscape, these mappings could already be outdated or imprecise.

The vast focus on technical topics in literature has overlooked the importance of studying this phenomenon from the business perspective, leaving a gap in research. In particular, no attempt has been made to perform a comprehensive study on the current status of the adoption of AI by VC firms. As of now, no research has been conducted on the market- and firm-level business implications of new technology applications in VC. The question of how AI is currently being implemented within VCs' operating structure remains unsolved, and it is still unclear whether AI will change VCs' traditional operations and impact standard practices.

1.3 Problem Statement and Research Questions

This research aims to broaden the current knowledge on the topic of AI implementation in the VC industry from a business perspective. This dissertation's goal is to explore the current status of AI adoption in the VC industry and the opportunities and potential limitations embedded in new technology implementation. This paper's research statement can be summarized with the following question:

How does AI technology adoption unfold in the VC industry?

The research problem is approached by providing a market overview of existing data-driven VCs, aiming at describing emerging significant trends and characteristics, and a firm-level study of practical AI technology application to VC practices. Its purpose is to provide a current view of the industry, as well as to reflect on future trends and possible developments. Thus, this study addresses the following research questions:

RQ1: What is the current market status of AI adoption in the VC industry?
RQ2: How can AI be practically applied in the VC sector?
RQ3: Which are the benefits and opportunities of AI implementation in VC?
RQ4: Are there any, and if so, which are potential limitations of AI implementation in VC?

1.4 Delimitation

An initial discussion on the technical aspects of the application of AI and ML algorithms in VCs (see *Section 1.2*) has been performed to provide a complete review of the literature on the topic. However, this research's intent is not to elaborate on nor enrich the technical literature. This paper's aim is not to propose new AI algorithm models to be applied in VC. Instead, this dissertation's purpose is to provide an overview of AI adoption in VC, taking a qualitative perspective. As such, only the theoretical concepts related to AI that will be functional in performing the analysis will be explained and discussed without exploring technical matters in depth. Furthermore, this paper is not restricted in space or in time. The study proposed takes on a global geographical scope and is not time-bound.

1.5 Thesis Outline

This paper is organized into six chapters. *Chapter 1* introduced the subject matter by providing a general contextual overview of the concept and literature review on the topic, finally presenting the research questions. *Chapter 2* provides a theoretical background on the notions of VC and AI to allow the reader to fully understand the analysis performed. In the first section of this chapter, the VC industry is placed in the asset management space, highlighting how it has recently gained relevance within the financial sector. Then, the structure and the functioning of the industry are explicated. Finally, the operational breakdown of traditional VCs and the challenges faced by investors are described. The AI section covers theory on the origins and definition of AI and deepens into relevant AI technologies knowledge. It also provides an overview of AI adoption by firms across industries. *Chapter 3* explains the methodology applied in this paper and overviews the data sources and data collection process chosen to answer the research questions. *Chapter 4* elaborates on the data gathered and analyses AI adoption in VC by focusing on the concepts of market adoption, technology applications, and VC data strategy. This section proposes a database of data-driven VCs autonomously developed and an analysis of interviews performed. *Chapter 5* discusses the findings providing a market-level overview of the current status of AI adoption, proposing a model for practical AI application in VC, and discussing the opportunities and challenges of new technology implementation. Finally, *Chapter 6* concludes by summarizing the study findings and debating limitations and future research possibilities in this field.

Chapter 2: Theoretical Background

2.1 Venture Capital

2.1.1 Asset Management and Venture Capital

Venture Capital (VC) is a subset of Asset Management (AM), which is a sector of the financial services industry that comprehends companies whose aim is to professionally manage assets for individuals, families, and institutions (Stowell, 2018). VC pertains to the alternative asset class and can be considered a form of Private Equity[1] (PE) that focuses on investments in early-stage ventures (Cote, 2021). In recent years, VC has become a prominent branch of the PE sector, positively impacting its growth and, in turn, gaining importance in the AM space, attracting investors' interest.

The AM industry has developed exponentially in the previous years, reaching a total of $112 trillion AuM[2] globally in 2021, growing 12% from 2020 and outpacing the average growth rate of 8% experienced between 2010 and 2020 (McIntyre et al., 2022). It is estimated that the industry will reach more than $145 trillion AuM globally by 2025 (Alexander et al., 2020) and that alternative assets will play a key role in AM growth over the next five years. In 2021 they represented only 20% of the global AuM but accounted for more than 40% of total AM revenue, and projections expected them to reach more than half of all global revenues by 2026 (McIntyre et al., 2022). PE, which represents 37% of all alternative assets, stands out as one of the primary growth drivers within this class, accounting for $6.3 trillion AuM in 2021 (Heredia et al., 2021, McIntyre et al., 2022, Asaftei et al., 2022). Concerning PE, buyouts are currently the largest sub-asset class. However, growth and venture assets have expanded their capital funding and AuM share at a faster rate than traditional buyouts and will probably outpace them in the future (Sheth et al., 2022). VC has been the fastest growing strategy within PE in terms of AuM in recent years, recording an unprecedented year-over-year growth of 43% in 2021, 12 percentage points above buyout, and 22 percentage points above growth equity. VC has also outperformed the other strategies' returns, confirming itself as the top-

---

[1] Private Equity refers to capital investments made into private companies. PE can be differentiated among VC, which focuses on investments in early-stage ventures, Growth Equity, which refers to capital investments in established and growing companies and Buyouts, which refer to the action of acquiring a controlling stake in a mature, typically public company, taking it private.

[2] Assets under Management (AuM), which is the total market value of the assets managed by the players in the industry, can be used to monitor the growth of the Global Asset Management market.

performing PE sub-asset class (Asaftei et al., 2022). In 2021, VC reached $ 1.68 trillion of global AuM, tripling its value from 2016 (Preqin, 2022). Moreover, global venture investment totaled more than $650 billion in 2021, reporting a 92% increase from $335 billion in 2020 (Teare, 2022). Regarding the geography of investment, VC fundraising is mainly concentrated in North America, followed by Asia, in which China takes the lead, and Europe (Statista Research Department, 2022b).

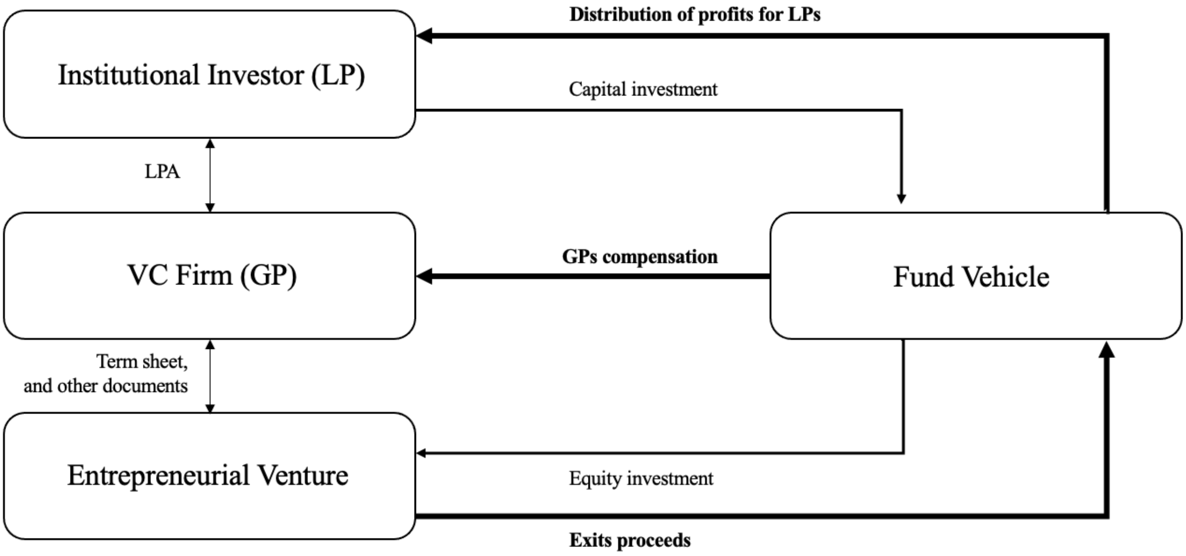
Despite the recent success, the VC sector has undergone a minor setback in the previous months. Venture investment has resented the recent world crises, leading to a quarter-over-quarter decrease of 23% in venture funding in Q2 of 2022 (CB insights, 2022). Late-stage and growth investment have been the most affected sectors by the pullback, while early-stage and seed funding have been less impacted. Even though the current startup funding does not reach the all-time-high records of 2021, it still exceeds pre-pandemic levels (Lehot, 2022). This insight foreshadows VC's predominant role in the AM space and predicts that it will remain relevant in the future.

2.1.2 VCs Structure and Investment Strategies

VCs can be defined as organizations that professionally invest funds in new privately held ventures with high-growth potential (Sahlman, 1990; Da Rin and Hellman, 2013; Gompers and Lerner, 2001). Strömberg and Kaplan (2001) argue that VCs are the intermediaries that connect entrepreneurs with good ideas but no funding and investors who have money but lack attractive ideas. As VCs act as middlemen between entrepreneurs and investors, it is mandatory to understand the contractual relations established between these parties to grasp the functioning of the industry.

VCs are typically structured as limited partnerships (Gompers, 1994; Gompers and Lerner, 2001). As exemplified in *Figure 1*, institutional investors, called Limited Partners (LPs), provide money to VCs, called General Partners (GPs), through a fund. The fund is the legal vehicle through which GPs invest LPs' money on their behalf in entrepreneurial companies. This contractual setup usually lasts ten years, in which GPs select, mentor, and monitor ventures to lead them to a profitable exit, finally redistributing returns to LPs. Once over, this cycle renews itself with the VC raising additional funds and investing in new companies (Gompers and Lerner, 2001; Torres, 2020).

Figure 1 - The Venture Cycle



Source: Adapted from Da Rin and Hellman, 2020

Thus, VCs have contractual relations with both LPs and venture companies. The relationship between LPs and GPs is governed by the Limited Partnership Agreement (LPA), which is a contract that sets out the rights and duties of both groups (Da Rin and Hellman, 2020; Litvak, 2009, Gompers and Lerner, 1999). As the industry is characterized by a high level of information asymmetry and uncertainty of payoffs (Amit et al., 1998; Sahlman, 1990), the LPA sets out covenants and rules concerning fund structure and GPs compensation to align the interests of the parties, limiting potential conflicts (Gompers and Lerner, 2001). Even though there is variance in the amount and distribution timing in the industry (Litvak, 2009), GPs usually receive a management fee which corresponds to 2% of LPs' capital contribution and a carried interest, which accounts for 20% of final profits (Gompers and Lerner, 1999; Metrick and Yasuda, 2010). The first is paid overtime to cover the VCs' operational costs, while the latter is a performance-based payment functional to incentivize GPs' good performance (Da Rin and Hellman, 2020). Usually, GPs are also required to make an upfront contribution of about 1% of committed capital to the fund to put something at stake, which disincentivizes excessive risk-taking. The relationship between GPs and venture companies is instead governed and managed by a series of agreements and legal documents drafted after the investment decision.

Focusing on GPs structure, VC firms usually have a similar internal organization and differentiate themselves based on the investment strategy adopted. VCs are generally small companies with simple corporate structures. They are organized as a three-layers hierarchy, with partners at the top, venture partners and associates in the middle, and support staff at the bottom[3] (Da Rin and Hellman, 2020). Gompers and Lerner (2020) report that ventures do not scale up much and that the average VC has 14 employees, among which five senior investment professionals, one venture partner, and three associates. Thus, VCs differentiate by specializing along three main dimensions: industry, geography, and stage of investment (Sahlman, 1990). First, VCs can choose the type and number of industries in which they focus their investment. Generally, they concentrate on high-growth and technology-intensive sectors with strong consumer interest and opt to differentiate in the number of focus industries. When investing in only one industry, they are called "specialists", while when investing in multiple, they are defined as "generalists". Second, when considering geography VCs can take a local or global focus. The benefit of investing close to their location is maintaining familiarity with the market and the regulatory environment, which reduces investment risk. However, the downside is that local business opportunities are limited, and VCs could miss several profitable deals abroad. Third, VCs determine their preferred stage of investment. Opting for early ventures entails higher risk and requires profuse investor domain expertise but demands lower financial requirements. Instead, later-stage ventures necessitate higher capital investments and superior financial expertise, but venture success is more easily foreseeable, reducing risk. Each VC fund philosophy aligns with the investment strategy, which will finally dictate the rules for capital deploying decisions (Da Rin and Hellman, 2020).

2.1.3 VCs Operations Breakdown

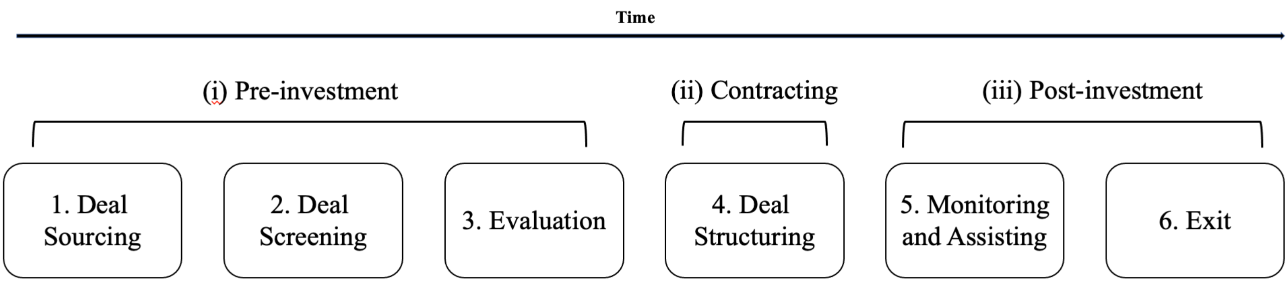
As compensation is partly performance-based, it is in the VCs' interest to select and invest in the most promising ventures in the market. However, it is not a straightforward task considering that up to 90% of new ventures end up making no financial returns, with half of them failing in the first five years (Patel, 2015; Da Rin and Hellman, 2020). VCs must rely on their processes, skills, and expertise to select flourishing startups. They aim to bring added value by providing assistance to enhance ventures' survival chances (Gupta and Sapienza, 1992). Even though VCs cannot eliminate risk as failure greatly depends on uncontrollable factors (Sahlman, 1990), they perform before- and after-investment decision operations to mitigate the risk of not realizing any profits.

[3] Support staff includes back-office and logistics personnel, analysist, researchers, and entrepreneurs in residence (Gompers et al., 2020)

The VC investment decision-making process has been widely studied in the literature (Wells, 1974, Tyebjee and Bruno, 1984; Silver, 1985; Hall and Hofer, 1993, Fried and Hirsch, 1994; Wright and Robbie, 1998; Gompers et al., 2020). Even though researchers have agreed upon taking a process perspective, there is no clear definition of the steps, as every author refines the stages by adopting diverse nomenclature and slightly different configurations (Hall and Hofer, 1993; Fried and Hirsch, 1994). For the sake of simplicity, the process presented in this paper (*Figure 2*) represents a generalization, proposing a six-stage partitioning that tries to reconcile most of the literature on the topic.

Leveraging the framework proposed by Strömberg and Kaplan (2001), it is possible to classify VCs practices in three salient moments of the lifecycle of a deal: (i) pre-investment, (ii) contracting, and (iii) post-investment. Pre-investment refers to operations of (1) deal sourcing, (2) deal screening, and (3) evaluation; contracting entails (4) deal structuring; post-investment comprehends (5) monitoring and assisting the venture towards (6) exit. The following sections will overview all these practices in detail.

Figure 2 - VC Decision-making Process



Source: author, adapted from Wells, 1974; Tyebjee and Bruno, 1984; Silver, 1985; Hall, 1989; Hall and Hofer, 1993; Fried and Hirsch, 1994; Wright and Robbie, 1998; Gompers et al., 2020.

1. Deal sourcing

Sourcing refers to the process by which VCs start considering deals as prospect investments (Tyebjee and Bruno, 1984). Deals can be originated through outbound or inbound sourcing (Charafeddine, 2016). Outbound sourcing refers to the active search for startups by leveraging VCs' internal networks and investors' referrals, participating in events, and performing proactive research aligned with the fund's investment strategy. Inbound sourcing refers to the VC activities that are functional in attracting startup applications through marketing and branding and by building an online presence. A recent survey by Gompers et al. (2020) emphasizes the importance of outbound deal sourcing, reporting that 90% of deals are generated actively by VCs, while only 10% of deals come from inbound activities. In particular, 30% of deals are generated through professional networks, 28% are from referrals, and almost 30% are proactively self-generated. This step is crucial in the investment decision process as all investors strive to find high-quality proprietary deals to gain a competitive advantage over competitors (Gompers et al., 2016).

2. Deal Screening

Before deciding which ideas and teams to support, VCs perform screening procedures on potential new investments (Sahlman, 1990). As the market is characterized by adverse selection (Amit et al., 1998), screening is essential to understand the fit between the VC and the startups. This step is relevant to preliminarily assess the idea's potential and the entrepreneurs' ability to ultimately reduce post-transaction monitoring issues (Wright and Robbie, 1998). First, VCs apply a filter that concerns the investment strategy, considering only the ventures that are a good fit with the firm (Fried and Hisrich, 1994). At this stage, venture factors such as investment size, geographic location, financing stage, product technology, and market sector are analyzed (Tyebjee and Bruno, 1984; Kaplan, 2010; Kaplan and Strömberg, 2004). Second, VCs investigate the management team's capabilities and the business plan's validity (Fried and Hisrich, 1994).

Zacharakis and Meyer (2000) highlight the importance for VCs to implement actuarial models of decision-making to set up an ad-hoc process for selecting ideas, striking a balance between the number and quality of deals proposed for further consideration and the ones that get discarded at this stage. Extensive literature has focused on studying which factor provides VCs' most prominent decision-making criterion. Some researchers argue that the most critical ingredients relate to the team and management capabilities (MacMillan et al., 1985; Gompers et al., 2020). Others, instead, support the supremacy of market and business components, highlighting the importance of competition and environmental threats (Khanin et al., 2012, Tyebjee and Bruno, 1984) and business model (MacMillan et al.; 1987; Zacharakis and Meyer, 1998; Kaplan et al., 2009).

3. Evaluation

Companies that pass the screening stage are then subject to due diligence, which is performed to reduce the information asymmetry between the parties (Torres, 2020). The deal originator presents the opportunity to other members of the VC firm who revise and scrutinize it. During this stage, VCs employ resources to further understand the business and ensure entrepreneurial capabilities through industry analyses and meetings with the company (Sahlman, 1990; Gompers et al.; 2020). Moreover, before committing to investing capital, VCs assess the potential risks and returns of the investment by performing the company's financial valuation (Tyebjee and Bruno, 1984). Up to this point, the VC has committed plenty of resources, undertaking an intensive process that can take up to months (Kaplan, 2010). Furthermore, Gompers et al. (2020) report that VCs, on average, screen more than 200 deals per year, making only four final investments.

4. Deal Structuring

When an investment decision has been made, the VC considers the results from preliminary analysis to negotiate the conditions with the venture to close the deal. The parties sign a term sheet, a contract that defines the rights and obligations by allocating cash flow rights and control rights and defining employment terms, the final valuation, and the investment staging (Strömberg and Kaplan, 2004; Kaplan, 2010). The agreed-upon terms are then converted into a set of final legal documents that regulate the relationship between the VC and the entrepreneurs (Da Rin and Hellman, 2020).

5. Monitoring and Assisting

After closing the deal, VCs can leverage their expertise and vast network to support ventures (Sørensen, 2007; Hellman and Puri, 2002). Out of all the value-adding activities performed by VCs, Gorman and Sahlman (1989) report that it is commonplace for VCs to provide monitoring, find management resources, and give strategic advice. Furthermore, VC partners often sit on the ventures' board after investing and provide strategic guidance and business advice (Amornsiripanitch et al., 2019; Gompers et al., 2020). Even though there is variance in the level of VCs' provision of support depending on whether they take on the role of lead investors (Gorman and Sahlman, 1989), it is reported that 60% of VCs interact at least once per week with their portfolio companies (Gompers et al., 2020).

6. Exit

Due to the industry's compensation structure (see *section 2.2.2*), VCs receive the carried interest only if capital gains are realized, making the timing and type of exit crucial to investment success. VC support to companies is functional to increase the chance of successful exits, most commonly through Initial Public Offerings (IPOs) or acquisitions by other companies or financial buyers (Da Rin and Hellman, 2020). However, due to the high market risk and uncertainty, up to 32% of ventures close down before reaching this stage (Gompers et al., 2020). In support of VCs' value-adding role, evidence shows that VC-backed startups have lower failure rates than other ventures (Puri and Zarutskie, 2012) and that VCs with experienced investors are more likely to go public (Sørensen, 2007).

2.1.4 VCs Investment Challenges

VC's decision to invest in young and innovative companies with a high probability of failure is characterized by uncertainty. As new ventures are experimenting with novelty, VCs face investment risks related to the development of the new technology, the relative market adoption, and the concrete execution of the idea. While the latter is more easily controllable by VCs, the first two types of risks are harder to mitigate. In fact, due to a general lack of data and information in the early-stage venture landscape (Kaplan and Lerner, 2016), the potential of technologies under development is still unknown, and the market reaction and consumer adoption levels are difficult to predict. Thus, VCs often must rely on their intuition, or "gut feel", when making investment decisions (Hisrich and Jankowicz, 1990).

Rather than relying on external information, VC investors leverage their expertise and cognitive frameworks to evaluate the attractiveness of a deal. Huang (2015) reports that decisions attributed to gut feel and based on the chemistry between the investor and entrepreneurs can prove to be highly effective. However, interposing subjective information in the decision-making process adds complexity by introducing biases in the decision. VCs' reliance on experience can lead to cognitive errors and problems of overconfidence and overfitting (Shepherd and Zacharakis 2003). Moreover, due to the large and varied pool of deal options, VCs can suffer from information overload, similarity bias, and availability bias (Franke et al., 2006; Shepherd et al., 2003; Zacharakis and Meyer, 1998).

Strictly linked to the venture selection decision, VCs face the crucial challenge of determining how to allocate time and resources internally. VCs source thousands of deals and screen more than 200 opportunities per year. Since they are small companies with limited working capacity, VCs must maximize their efforts to create the highest value possible with their own resources. It is reported that out of all the decision-making steps, deal scouting and post-investment activities profoundly matter in VC value creation. Sørensen (2007) suggests that while being both extremely relevant, deal sourcing and investment selection are relatively more important for driving VC returns than activities aimed at adding value after the investment. Building on his theory, Gompers et al. (2020) claim the importance of deal selection for VC value creation, which is reflected in time allocation choices. They report that out of the average working week of 55 hours, VCs spend 22 hours sourcing and screening deals and 18 hours per week working with portfolio companies. They also state that the average deal takes about 83 days to close, which can represent much time in fast-changing markets such as the ones in which VCs are operating.

Notwithstanding all these challenges, some VCs are trying to mitigate risks by relying more on data and processes to speed up decisions, creating value more efficiently. A recent trend called quantitative sourcing has emerged in the VC industry. VCs are starting to study ways to use data from multiple sources to identify investment opportunities (Gompers et al., 2020). With advancements in technology and the development of new AI applications, quantitative sourcing might be a powerful tool in the hands of VCs.

2.2 Artificial Intelligence

2.2.1 The Origins of AI and its Definition

The origins of Artificial Intelligence (AI) can be traced back to the 1830s, when Ada Lovelace and Charles Babbage developed the Analytical Engine (Chalmers et al., 2021). Even though the machine was the first operating program to compute Bernoulli's numbers, offering enormous computational power, it could not produce original ideas as incapable of performing creative acts (French, 2000). This idea was referenced as the "Lady Lovelace Objection" and was disputed by Alan Turing in his seminal article "Computing Machinery and Intelligence" (Turing, 2009). Turing theorized how computers could automatically learn and perform actions not anticipated by their programmers, introducing the Turing Test as a measure of machine intelligence.

These historical foundations led to the first formal definition of AI at Dartmouth College in 1956. Marvin Minsky, John McCarthy, and Claude Shannon first described the concept of AI, establishing it as an organized field of research and academic discipline (Pan, 2016; Chalmers et al., 2021). The notion of AI was outlined as *"making a machine behave in ways that would be called intelligent if a human were so behaving"* (McCarthy et al. 1955, p. 11), indicating machines' ability to understand, learn and think similarly to human beings (Pan, 2106). Likewise, Marvin Minsky defined AI as *"the science of making machines do things that would require intelligence if done by* *men"* (Minsky, 1968, p. v). Since then, the concept of AI has evolved, and different denotations have emerged, leading to the lack of a singular agreed-upon definition for AI (Allen, 1998; Wang, 2019). Usually, delimitations of the idea comprise a reference to human intelligence, describing AI as technology that performs and automate actions that require intelligence when performed by people (Taddy, 2018; Norvig, 2002).

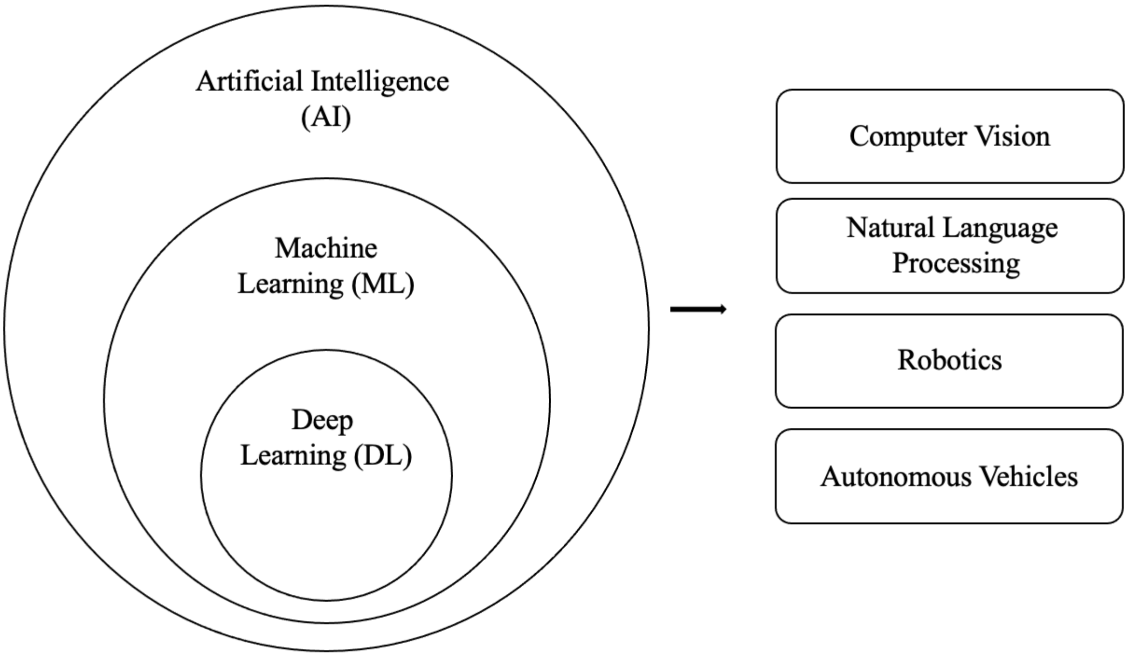
The extent to which AI can interiorize human behavior to create conscious machines has been conceptualized by Kurzweil (1999) with the distinction between "Narrow AI" and "Strong AI" (Gobble, 2019; Kurzweil, 1999). Narrow AI, also referenced as "Weak AI" (Norvig, 2002) or "Artificial Narrow Intelligence" (ANI), refers to the capability of AI to replicate specific intelligent behaviors in particular contexts. Thus, this type of AI is focused on narrowly defined structured tasks and requires some level of human reconfiguration to keep operating. In contrast, Strong AI, or "Artificial General Intelligence" (AGI) (Goertzel and Pennachin, 2005), refers to the application of generalized AI to any domain and is not restricted to a specific context as the former. Unlike ANI, AGI creates machines that think and reason like human beings, capable of self-adapting to situations.

Although AI development is moving quickly, the current status of technology has mostly focused on Narrow AI, concentrating on automating and making everyday tasks more efficient (Manyika et al., 2017, Vorholes, 2016). However, the evolution towards AGI is gradually happening, and the technology paradigm is shifting to this next frontier, which over time would allow to satisfy early definitions of AI and assimilate technology with human intelligence (Goertzel, 2014). Eventually, researchers argue that AGI will evolve into "Artificial Superintelligence" (ASI) and surpass human intelligence in all its aspects (Bostrom, 1998).

2.2.2 AI Technology

Shifting focus from a prevalently theoretical dissertation of the concept of AI to a description of its technical aspects, this section is centered on explaining the technology behind AI. In fact, AI is an umbrella term encompassing subfields that relate to various technologies and applications (Manyika et al., 2017). *Figure 3* proposes a graphical representation of the technological concept of AI, overviewing the most common applications, which will be discussed in the following paragraphs.

Figure 3 - Artificial Intelligence Technologies and Applications



The capability of learning from data is at the core of AI. Machine Learning (ML), a subset of AI, is a method of data analysis that allows to create and improve mathematical models by continuously learning from data to infer the future and make predictions (Bonaccorso, 2017). ML comprehends three major paradigms: supervised learning, unsupervised learning, and reinforced learning (Jordan and Mitchell, 2015).

In supervised learning, the algorithm is trained with an externally supplied dataset that takes the form of a collection of *(x,y)* pairs, where *x* is the input and *y* is the output, containing the correct answers. Then the algorithm generalizes from training data and forms its predictions generating a function *f(x)* that maps all possible inputs *x* to desired outputs *y* (Jordan and Mitchell, 2015; Ayodele, 2010; Marsland, 2014). Many different forms of mapping f exist and can be grouped into regression and classification. Examples are linear and logistic regression, decision trees, decision forests, support vector machines, kernel machines, Bayesian classifiers, and K nearest neighbors (Bonaccorso, 2017; Jordan and Mitchell, 2015; Snigh et al.,2016; Marsland, 2014).

Instead, unsupervised learning is a learning approach involving the analysis of unlabeled data that relies on assumptions of structural properties of the data (Jordan and Mitchell, 2015). It differs from supervised learning because there is no explicit target outputs but only input patterns *x* (Dayan, n.d.). It is useful when it is necessary to understand how a set of elements can be grouped according to their resemblance (Bonaccorso, 2017). Thus, unsupervised learning exploits similarities between inputs to automatically cluster data points close to each other without being told which class they pertain to (Marsland, 2014). Unsupervised learning algorithms comprehend dimension reduction methods such as principal components analysis and factor analysis and clustering methods, such as K-means algorithms, hierarchical clustering, and density-based clustering (Mitchell and Jordan, 2015).

Reinforcement learning is an intermediate approach between supervised and unsupervised learning. An agent interacts with the environment via a set of actions to learn a strategy via trial and error to select the sequence of actions that grant the maximum return over time (Kaelbling et al., 1996). Thus, training data only indicates whether an action is correct or not through a reward. In this case, the algorithm knows the current input (*state*) and the possible set of actions that it can perform (*actions*). The algorithm computes the value of its actions by learning through feedback, trying to maximize the expected reward over time (Mitchell and Jordan, 2015). This learning paradigm is proper when the environment is unknown and very dynamic (Kaelbling et al., 1996). Even though these above-mentioned paradigms help organize ideas, current developments in the field involve blends across the three categories, such as semi-supervised learning (Mitchell and Jordan, 2015, Ayodele, 2010).

Deep Learning (DL) is a subset of ML that allows computational models of multiple processing layers to learn from complex and high-dimensional data with multiple levels of abstraction (LeCun et al., 2015). Deep Learning uses Artificial Neural Networks (ANN) to simulate the structure of neurons and biological neural networks in the human brain, loosely modeling their functioning. Researchers McCulloch and Pitts (1943) have studied neurons and created a mathematical model, which has now become the primary building block of ANN, permitting the development of different DL methods (Alexander, 2020; Marsland, 2014). DL algorithms cover all three learning paradigms and present a better version of ML. Although DL requires more training time and computational power, it offers more effective processing models with better learning ability and prediction accuracy, increasing performance with the amount of data used for training (Lai, 2019; Kraus et al., 2018).

Within AI, ML, and, more specifically, DL algorithms have emerged as the method of choice for developing practical software due to their ability to learn and predict through diverse inputs such as images, texts, video, and audio. Thus, ML and DL play a crucial part in AI applications, such as Computer Vision and Natural Language Processing (LeCun, 2011; Manyika et al., 2017, Jordan and Mitchell, 2015).

Computer Vision is the field of AI that allows computers to take actions and make recommendations based on information gathered from visual inputs (IBM, 2022). It leverages DL, Convolutional Neural Networks[4] and Recurrent Neural Networks[5]to perform tasks such as object detection, image classification, object tracking, and content-based image retrieval (IBM, 2022). Real-life examples of these tasks are face recognition, action and activity recognition, and human pose estimation (Voulodimos et al., 2018).

Natural Language Processing (NLP) stands at the intersection of AI and linguistics. It refers to the ability of a computer to understand and process human languages in the form of text or voice data (Hirschberg, 2015). NLP has two distinct focuses: the first one refers to analysis for representation purposes and is referred to as natural Language Understanding (NLU), and the second one relates to the production of language and is referred to as Natural Language Generation (NLG)[6] (Liddy, 1998; Khurana et al., 2012). ML and DL serve as an essential value addition to all the analysis layers involved in NLP and allow to perform tasks such as speech recognition, sentiment analysis, natural language generation, text categorization, and information extraction (Khurana et al., 2012, IBM, 2022). Practical use-cases of NLP include machine translation, social media sentiment analysis, text summarization, spam detection and filtering, and conversational agents such as chatbots and virtual agents (IBM, 2022; Liddy, 1998; Hirshberg, 2015).

Applications of AI are not limited to Computer Vision and NLP but also include Robotics and Autonomous Vehicles (Manyika et al., 2017) and are continuously developing thanks to advances in ML and DL.

2.2.3 Global AI adoption: Statistics, Trends, and Challenges

The development of AI technology is currently at an evolutionary state (Rimol, 2021), signaling the importance for organizations to adopt AI solutions and integrate them seamlessly into processes to deliver higher business value. While some enterprises have implemented specific AI programs, others are scrambling to figure out how to integrate AI within business workflows to realize the full potential of these new technologies (Dataiku, 2022). Although the complexity of AI implementation is high, reports show that AI adoption by companies is increasing yearly. In 2021, the global total corporate

[4] CNN: ANN in which the connections between neural layers are inspired by the organization of the animal visual cortex, the portion of the brain that processes images, well suited for perceptual tasks.
[5] RNN: ANN whose connections between neurons include loops, well- suited for processing sequences of inputs
[6] NLU includes: phonology, morphology, pragmatic syntax and semantic analysis, while NLG includes natural language text generation.

investment in AI reached 93.5 billion U.S. dollars, an increase of 38% from the previous year (Zhang et al., 2022, Statista Research Department, 2022a). Furthermore, the McKinsey State of AI 2021 report signals a 56% rate of AI adoption by companies in at least one function, increasing by 50% from 2020 (Chui et al., 2021).

From a macro perspective, AI adoption differs vastly across geographies, industries, and companies. At the geographical level, global adoption of AI in 2021 was driven by India, followed by the Developed Asia-Pacific region and Developing Markets, including China (Chui et al., 2021). AI rollout in 2022 accelerated by an average of 53% globally compared to 2021, mainly driven by China, Latin America, and India (IBM, 2022). At the industry level, high tech, telecom, and financial services sectors are driving AI adoption (Zhang et al., 2022). These sectors, which have a long history of digital investment, are advantaged in qualifying as first movers and early adopters (Bughin et al., 2017). Moreover, at the company level, large firms have a higher rate of AI awareness than smaller firms across all sectors because they typically have better-structured data and business processes and more technically and digitally skilled employees. Furthermore, they report a greater adoption rate due to their economic capacity to commit to significant fixed-cost investments required for AI.

From a micro perspective, firms hardly adopt AI along the complete value chain. Business functions that report the highest level of AI application are product and service development, service operations, and marketing and sales (Chui et al., 2021; Zhang et al., 2022). As the activities with the highest automation potential involve predictable physical activities, processing data, and collecting data (Bughin et al., 2017), the top AI capabilities embedded in standard business processes are reported to be NLU, Robotics, Virtual Agents, and Computer Vision (Zhang et al., 2022).

The firms that have not yet or only partially adopted AI find internal barriers to adoption and external risks. Internally, firms have limited AI skills and knowledge from the workforce, face a too high price for implementation, lack tools or platforms to develop models, and find too much project and data complexity (IBM, 2022). Externally, there is a concern for cybersecurity and regulatory compliance (Zhang et al., 2022). Despite challenges, companies that have set out proactive strategies for AI adoption have implemented successful mitigation ways to overcome barriers. Apart from hiring skilled employees and carrying out technological change, mutual learning between AI and humans is a pivotal matter to strategize upon to facilitate organizational learning (Ransbotham et al., 2020).

Chapter 3: Methodology

3.1 Research Approach

This research aims to study the adoption of new technology, namely AI, in the VC industry. The phenomenon under study has a novel character, both in terms of the real-life development status and literature record. As AI technology adoption is happening and evolving at the same time as this dissertation is written, its implementation and development path is still uncertain and might take unexpected turns in the future. Accordingly, considering the novelty of the matter and the continuous concrete technological and business developments, the topic has not yet been extensively studied in the literature.

Due to the innovative nature of the research purpose, this paper aims to perform exploratory research. Exploratory research is the ideal method to investigate un-explored and un-interpreted phenomena and better understand the issues under inquiry (Cavana et al., 2001; Marshall and Rossman, 2006; Khan, 2014). It is the optimal design to adopt when there is no past data or few studies for reference, such as in this case. Exploratory studies are used to investigate undefined problems and provide in-depth descriptions of the unknown, potentially identifying future research directions to elaborate on. Thus, this type of research helps conduct initial explorations of complex phenomena and sets the basis for future investigations (Sofaer, 1999).

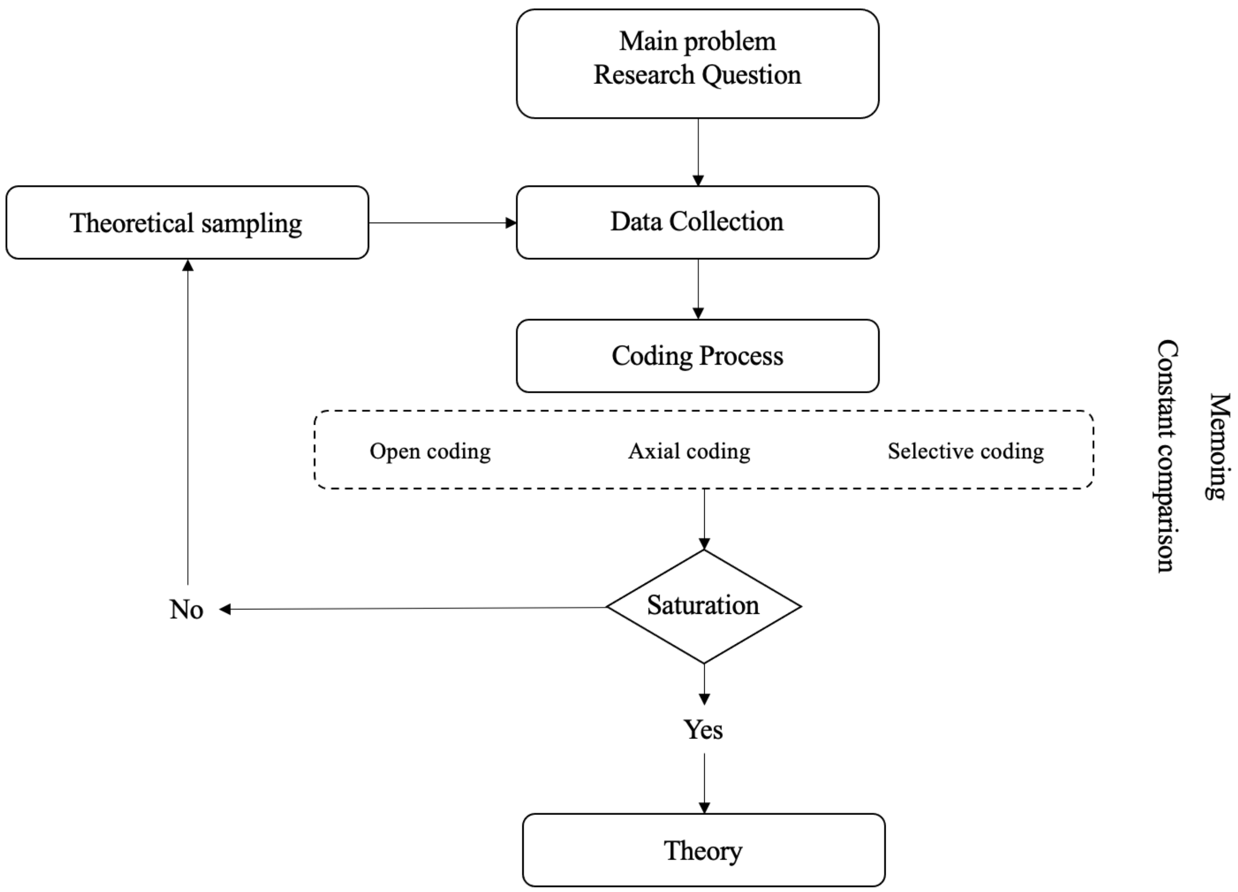
In line with the exploratory character of research, this paper will use qualitative methods to examine the topic and answer the research questions. This choice is driven by the fact that qualitative techniques are frequently used to inform on new concepts and theories. Moreover, qualitative studies mainly focus on answering "what" and "how" types of questions as they offer a dynamic approach to research, which is needed in highly uncertain and not transparent contexts, such as the one under analysis (Khan, 2014). This type of research describes the potential antecedents and factors that aim to explain matters about which little has been explored (Strauss & Corbin, 1998). Moreover, apart from observing and narrating facts, qualitative research can be based upon the interpretations of people's perceptions of different situations and events (Guba & Lincoln, 1994; Khan, 2014).

The chosen qualitative design is grounded theory, as the study aims to derive a general theory grounded in the data collected (Charmaz, 2006, Strauss & Corbin, 1998, Strauss and Corbin, 1990). Grounded theory allows to overlook experiences of participants and many data sources to develop an objective understanding of the topic. The grounded theory process involves using systematically collected data to identify relationships of categories of information and develop a model or explanation of the meaning of the study.

3.2 Research Design

The grounded theory method is represented in *Figure 4*. The process involves data collection from different sources that can answer the original research questions. Sources can pertain to interviews, observation, surveys, focus groups, and other types of data collection techniques. The data gathered is then analyzed systematically through three coding levels that aim to investigate the issue in detail. The first level is called open coding and refers to the process of breaking up the collected data and developing categories of information. The second level is axial coding and aims to establish relationships and interconnections between categories. The third level is called selective coding and refers to identifying the core categories and building up a storyline that can link the previously identified categories (Chun Tie et al., 2019). During the coding process, continuous comparative analysis and memoing are performed. Memoing, which refers to generating detailed memos recording the researchers' thoughts and interpretations during the process, can be subsequently leveraged in the theory formulation and narration phase (Birks and Mills, 2015). The theory is formulated if the coding process provides a complete and detailed explanation of data. Instead, if open investigation possibilities still exist, theoretical sampling is performed again to include further data in the analysis. New theoretical sampling kickstarts an iteration between data collection and coding, which is performed until theoretical saturation is reached. Theoretical sampling follows cues from coding and analysis and fills gaps, clarifies doubts, and tests the researchers' interpretations as the study progresses. Finally, the findings drawn from data collected and coding are reported in a structured way to inform and create knowledge about the topic under research.

Figure 4 - Grounded Theory Process



Source: adapted from Chun Tie et al. (2019) and Jiang et al. (2019)

3.3 Data Sources and Data Collection Method

Extensive data collection from multiple sources is the backbone of qualitative research, specifically of grounded theory methodology (Khan, 2014). For this study, data has been collected through both primary and secondary sources (*Figure 5*).

Following the grounded theory research process, this study has progressed and iterated through three stages (summarized in *Table 1*) before reaching theoretical saturation.

Figure 5 - Data Sources Classification

Primary Sources

- Interviews with VCs
- Informal talks with experts in the field

Secondary Sources

- Newspaper articles
- Blogs and newspapers
- Social networks (e.g. Twitter, LinkedIn)
- VCs websites
- Online VC databases (Crunchbase)

The first round of data collection used secondary sources to gain an overview of the VC market concerning AI adoption. Data has been gathered online through newspaper articles, blogs and forum discussions, VC websites, online databases such as Crunchbase, and social networks such as Twitter and LinkedIn. This round of data collection aimed at building a database of data-driven VCs. Data-driven VCs can be defined as VCs that use technology and data to improve their decision-making processes and to increase the chances of finding and investing in successful opportunities. In particular, data-driven VCs include firms that use AI, different types of automation software and platforms that heavily rely on data. The choice of focusing on a broader technological group that is not strictly related to AI has been driven by the relatively scarce information available on the topic, which imposes constraints on research specificity. Moreover, due to the structure and competitiveness of the VC market, maintaining confidentiality for VC firms is key to achieving success. Thus, not all VCs are keen on outing their internal processes and discussing the technologies they are currently using because of fear of competitors' imitation. For these reasons, it was deemed nearly impossible to discern VCs that are using data from the ones fully leveraging AI. Thus, all the VCs that are heavily using data have been considered for building the database under the assumption that any type of intensive data usage in the VC market can represent AI application in the present or might lead to AI adoption in the future.

The database has been built using a bottom-up design, starting from reports of singular cases available online and abstracting them to create a generalized database. The first coding phase has led to the identification of three categories of different sources reporting cases of data-driven VCs. The results show that startup cases can be reported by (i) newspapers or blog articles, (ii) employees' reports, and (iii) on the VC website. This classification can be utilized as a proxy for signaling data commitment. It is expected for the VC firms which openly declare their data strategy to be inclined to invest in AI development. Publicly admitting a data approach suggests that the company has taken concrete actions to back up the claims made, at least to a certain extent.

Building on the first step, a second round of data collection to gain more accurate firm-level knowledge on the topic has been performed based on primary data. Out of the first classification, the most relevant VCs for data and AI usage have been contacted, asking for an interview. Interviews were conducted in a semi-structured way following an interview guide (reported in *Appendix 1*). Interviews aimed to uncover technology uses other than the ones reported online and to investigate the experience of venture capitalists currently working with this new technology. Moreover, various informal consultations with experts in the field have been performed to gain further insights. The novel acquired data has permitted to uncover momentous themes linked to the research topic and expand the database with new cases.

Table 1 - Data Collection and Iteration Steps

Step	Source Type	Source	Description	Coding
Step 1	Secondary Sources	Newspaper articles, blogs and forum discussions, VC websites, online databases, and social networks	Online desk-based research provides an overview of data usage in the VC market.	Classification by reporting source
Step 2	Primary Sources	Interviews and informal conversations with experts in the field	Contacting the most relevant data-driven VCs to investigate their practices and the application of AI to their internal processes closely.	Theme coding to investigate the topic
Step 3	Secondary and Primary Sources	VC websites, online databases, social networks, interviews, and informal conversations with experts in the field	Enrich the database by adding new cases and bringing more detail	Classification by presence of skilled data employees

The third step of data collection consisted in incorporating the updated knowledge about data-driven VC cases in the database, expanding it accordingly. Additionally, the database has been reviewed in light of valuable insights received during interviews. Hiring data scientists in the context of AI adoption in VC has emerged as a relevant theme. Thus, the database has been enriched with a new classification, considering whether each VC firm in the list had hired data scientists, eigengenes, or employees covering similar roles. Information has been collected by reviewing each firm's website and their profile pages on Crunchbase and LinkedIn. This evidence brings new facets to the analysis, contributing to the final aim of understanding AI adoption in VC.

This iteration has led to reporting a total of 92 cases of data-driven VCs, four interviews, and two informal conversations (*Table 2*). Data collection to build the VC database has been performed as continuous research in the period between January and June 2022. Regarding interviews, it was possible to talk with four VC firms out of 33 companies that had been contacted, representing a response rate of 12%. Considering the total number of VCs in the database and the level of secrecy in the market, this response rate is deemed appropriate to guarantee a solid analysis. The interviews were conducted online through video calls between April and May 2022 and ranged between 25:14 and 35:44 minutes in terms of maximum duration[7]. Finally, informal conversations have been held on multiple occasions with two stakeholders, Francesco Corea from Balderton Capital and Lucrezia Lucotti from 360 Capital. Francesco Corea has provided support through online meetings in trying to delineate the research study and identify data-driven VCs. Lucrezia Lucotti has helped depict VCs' traditional operations, raising interesting questions to be considered in the topic analysis.

Table 2 - Interviews and Conversations Overview

Type of data source	Participant name	VC	Role
Interview	Giovanni Calabrese	Redstone	Investment & Data Intelligence
Interview	Moreno Bonaventura	InReach Ventures	Lead Machine Learning Engineer
Interview	Amit Kaistha	Hatcher+	Head of Investor relations
Interview	Neil Callahan	Pilot Growth Equity	Founder and Managing Partner
Informal conversations	Francesco Corea	Balderton Capital	Research Lead
Informal conversations	Lucrezia Lucotti	360 Capital	Associate

[7] Interview transcriptions are available in *Appendix 5*.

Chapter 4: Analysis

This section presents the data collected during the research. It provides an overview of the VC database that has been built and elaborates on insights received from interviews and conversations with essential stakeholders in the industry. The analysis chapter has been organized by thematic areas that have emerged during the study. This choice is functional in giving this paper a clear and understandable structure and facilitating the comparison among different sources. The themes analyzed cover the topics of market status, technology application, and data strategy considerations.

4.1. Market Status

An information collection exercise has been performed to get an overview of the status of data usage and AI application in the VC market, resulting in the construction of a database of data-driven VCs (*Table 3, Appendix* 2). The database totals 92 VCs which have been selected after meticulous research triangulating numerous sources (see *4.3*). The final list includes data-driven VCs in a broad technological sense, comprehending firms that use data-intensive platforms to support entrepreneurs and investors, companies that implement automation algorithms to improve efficiency, and VCs that leverage AI in their internal practices.

To analyze the market landscape, VCs have been classified using standard indicators such as their geographic location, stage of investment, and number of employees. However, due to the low level of information disclosure that characterizes the VC sector, it was nearly impossible to categorize VC firms into strict types of data-usage classes without relying on insider information for all the firms listed. Nonetheless, an alternative classification has been created based on the type of reporting source from which information was gathered to provide insights into the transparency of the market. *Table 3* provides an overview of the mentioned categorizations and additionally presents case-by-case notes on particular data applications and technology usage. This information adds a further level of specificity, which can be helpful for readers interested in investigating distinct cases.

Table 3 - VC Database

Name	Location	Stage	Size	Source	Notes on data-driven VC type
500 Startups	US, San Francisco, CA	Seed, Early	100+	Newspaper or blog articles	Reported in articles to be data-driven.
645 Ventures	US, New York, NY	Seed, Early	21-50	Employee's reports	It has developed 645 Voyager, which is a comprehensive software platform that powers the entirety of the firm's operations to Automate VC manual tasks. It is reported to have a data-driven approach to sourcing.
Accel	US, Palo Alto, CA	Seed, Early, Late	100+	Newspaper or blog articles	It uses a 3rd party data and AI software for VCs called Spectre.
Accelerated Ventures	EU, London, UK	Early	2-10	Newspaper or blog articles	It has developed a matchmaking platform for startups and investors to reduce fundraising time.
Akkadian Ventures	US, San Francisco, CA	Late	11-20	VC website	It reports that it has built its custom analytics software to spot companies early.
Amaranthine	US, San Francisco, CA	Early	2-10	Newspaper or blog articles	According to articles, the fund uses the information it gathers on the startups and attendees who come to WebSummit to support its investment decisions.
Andressen Horowitz	US, Menlo Park, CA	Early, Late	100+	Newspaper or blog articles	It has hired engineers and data scientists over the years and mentioned "proprietary data sources" in some of the publicly available materials, but there are not a lot of public details.
AngelList VC	US, San Francisco, CA	Seed, Early, Late	100+	Newspaper or blog articles	They define themselves as data-driven and use their dataset to leverage investment decisions.
Ardian	EU, Paris, FR	Late, debt, PE	100+	Newspaper or blog articles	It has been cited in an article as VC that leverages data.
Astorya	EU, Paris, FR	Early	11-20	VC website	They have an automated scouting technology that supposedly creates Europe's biggest insurance-related deal flow.
Atomico	EU, London, UK	Early, Late, PE	21-50	Newspaper or blog articles	Reference to data usage is present in one of the employee's job descriptions, but no further public data is available.
Backed.vc	EU, London, UK	Seed, Early	21-50	Newspaper or blog articles	It has developed a platform for founders to find the best team members and match talent.
Bain Capital Ventures	US, San Francisco, CA	Seed, Early, Late	51-100	Newspaper or blog articles	It uses a 3rd party data and AI software for VCs called Harmonic.
Balderton Capital	EU, London, UK	Early	51-100	Newspaper or blog articles	It uses a 3rd party data and AI software for VCs called Spectre.
Bessemer Venture Partners	US, Redwood City, CA	Early, Late	100+	Newspaper or blog articles	It had built its proprietary tracking systems when Chris Farmer (currently CEO of Signalfire) had been working there, but details about current technology are confidential.
Bloomberg Beta	US, San Francisco, CA	Early	2-10	Newspaper or blog articles	It partnered with People.co to create algorithms capable of predicting "future founders".

Name	Location	Stage	Size	Source	Notes on data-driven VC type
Blossom Capital	EU, London, UK	Early	11-20	Employee's reports	It follows a data-led approach to sourcing companies, helping to identify companies that are not on the radar.
Clear Ventures	US, Palo Alto, CA	Seed, Early	11-20	VC website	The firm has developed a software called “Clear Ecosystem Advantage”, which uses data and automation to aggregate data from many sources and allows the entrepreneurs to take advantage of the firm's relationships with customers, talent and investors.
Connectic Ventures	US, Covington, KY	Seed, Early	11-20	VC website	It has replaced the standard pitch with a data-driven process. Its mission is to become the first VC Firm in America to use data from sourcing all the way through to investment decisions.
Correlation Ventures	US, San Diego, CA	Seed, Early, Late	21-50	VC website	Thanks to data, they claim to be able to take investment decisions in under two weeks, removing biases.
Daphni	EU, Paris, FR	Seed, Early	21-50	Newspaper or blog articles	IT created a digital platform that allows people to co-invest and brings transparency into startup applications.
Deep Knowledge Ventures	ASIA, Hong Kong	Seed, Early	51-100	Newspaper or blog articles	The firm used software called VITAL to spot and assess investment opportunities. It is unsure if it still uses this software today.
Dorm Room fund	US, Philadelphia, PA	Seed	21-50	Newspaper or blog articles	It developed a matchmaking platform for founders and investors.
Draper Fisher Jurvetson	US, Menlo Park, CA	Early, Late, PE	11-20	Newspaper or blog articles	The founders have publicly acknowledged the power of AI and ML in VC.
Earlybird VC	EU, Berlin, DE	Early, Late	51-100	Employee's reports	They are not vocal about AI, but an employee posted an article on medium explaining their AI usage for deal sourcing and screening.
Entrepreneur First	EU, London, UK	Seed, Early	100+	Newspaper or blog articles	It has developed a platform that matches founders with cofounders and investors and uses AI to identify the most promising founders.
EQT Ventures	EU, Stockholm, SE	Early, Late	21-50	VC website	It has an AI software called "Motherbrain", which is highly advertised.
Fairview Capital	US, West Hartford, CT	Early, PE	21-50	Newspaper or blog articles	They shared an article stating that data matters in VC, but currently, there are no further insights.
FF Venture Capital	US, New York, NY	Seed, Early	51-100	Newspaper or blog articles	It has developed Totem, an operating system for VCs which was later spun off.
First Mark Capital	US, New York, NY	Seed, Early, Late	21-50	Newspaper or blog articles	The firm has built a platform that connects backed entrepreneurs with talent, customers, and expertise, running hundreds of public and private events. It leverages proprietary software to track how they impact its portfolio.
First Round Capital	US, San Francisco, CA	Seed, Early, Late	100+	Newspaper or blog articles	It has developed a platform for founders to share their knowledge.

Name	Location	Stage	Size	Source	Notes on data-driven VC type
Floodgate	US, Palo Alto, CA	Seed, Early	21-50	Newspaper or blog articles	The firm has publicly acknowledged using data mining techniques to uncover the typical characteristics of the most successful companies in 2013, but since then, the efforts in the data space have remained undisclosed.
Fly Ventures	EU, Berlin, DE	Seed, Early	2-10	Newspaper or blog articles	It is reported to use AI for deal sourcing.
Follow[the]Seed	AUS, Sydney	Seed, Early	11-20	VC website	It defines itself as a global data-driven VC and has developed the Raving Fans algorithm to identify applications and technologies that are likely to become viral hits.
Force Over Mass	EU, London, UK	Early	2-10	Newspaper or blog articles	It has created an online portal that allows the fund's LPs to easily review deals currently in the pipeline and easily participate in them.
Future Ventures	US, San Francisco, CA	Early	2-10	Newspaper or blog articles	It uses a third-party software called Hum Capital for fundraising matchmaking.
Fyrfly	US, Menlo Park, CA	Early	2-10	Newspaper or blog articles	It claims to be data-driven as a principle, but no further insight.
General Catalyst	US, Cambridge, MA	Seed, Early, Late	21-50	Newspaper or blog articles	A former employee is the creator of Signalfire. They supposedly started using data, but there is no clear information.
Georgian Partners	CA, Toronto, ON	Late	100+	VC website	Claims to have a data-driven approach, it has a proprietary platform that uses AI to help founders scale up and grow their business.
Gradient Ventures	US, Mountain View, CA	Seed, Early	21-50	Newspaper or blog articles	Supposedly uses some data tools, as its parent company is Google.
Greycroft	US, New York, NY	Seed, Early, Late, PE	51-100	Newspaper or blog articles	The firm sees the potential in using data to allow human investors to make better decisions and has been known to invest significant resources into its data team to help build this capacity.
Greylock Partners	US, Burlington, VT	Early, Late	100+	Newspaper or blog articles	Cited in various articles as data-driven, but there are no further insights as of now.
GV	US, Mountain View, CA	Seed, Early, Late	100+	Newspaper or blog articles	It uses an investment approval process aid called "The Machine", which is a ML algorithm that opines on potential investments—no detailed information on the software is available.
Hatcher+	ASIA, Singapore	Early	21-50	VC website	It has developed a venture-as-a-service technology platform called Hatcher+ VAAST, which has Deal Scout and Deal Screen features.
Headline	US, San Francisco, CA	Seed, Early, Late	51-100	Newspaper or blog articles	Former e.ventures, they claim to have a proprietary deal sourcing software, which apparently works with AI. However, the website is not vocal about it.
HOF Capital	US, New York, NY	Early, Late	11-20	Newspaper or blog articles	It is reported to have built a proprietary tool to help them with sourcing and pre-screening deals.

Name	Location	Stage	Size	Source	Notes on data-driven VC type
Hone Capital	US, Palo Alto, CA	Early, Late	2-10	VC website	It is vocal about having a data-driven approach to analyze deal-specific information in its proprietary model to identify start-ups with the greatest potential for superior returns.
Hummingbird Ventures	EU, Gent, BE	Early	21-50	Employee's reports	A former employee says that he was using data for sourcing collected through third-party sources.
IA Ventures	US, New York, NY	Seed, Early	2-10	Newspaper or blog articles	It is reported to be data-driven thanks to its investment in a data scientist, no further insight is provided.
Index Ventures	US, San Francisco, CA	Seed, Early, Late	21-50	Employee's reports	Ex-employee at Index Ventures, that has now founded Blossom Capital, claims that he built one of the first data-driven deal sourcing practices in VC when employed at Index Ventures in 2012. No further details are available.
Initialized Capital	US, San Francisco, CA	Early	21-50	Newspaper or blog articles	It has developed internal software for voting decisions.
InReach ventures	EU, London, UK	Seed, Early	11-20	VC website	It has developed a proprietary deal flow platform called DIG, which allows to discover and evaluate the most promising startups in Europe.
Insight partners	US, New York, NY	Late	100+	Newspaper or blog articles	The firm uses data to invest, as reported in an article, but no clear information is available.
Ironstone Group	US, San Francisco, CA	Early	2-10	Newspaper or blog articles	Mentioned in an article to be among the VCs that perform semi-automate investing decisions, but no further information is available.
Joyance Partners	US, San Francisco, CA	Seed, Early	11-20	VC website	Their vision is that "data science and AI create the opportunity for a new, transformational approach to early investment with the potential to deliver stronger, more consistent returns, greater exit optionality, and more precise response to changes in key factors driving entrepreneurial success than ever before possible."
Kima Ventures	EU, Paris, FR	Seed, Early	11-20	Newspaper or blog articles	It is reported to have developed an in-house deal flow watcher tool.
Kleiner Perkins	US, Menlo Park, CA	Early, Late	100+	Newspaper or blog articles	It is reported to have developed a social monitoring software called Dragnet in 2013, but there have not been insights since then.
Labx Ventures	US, San Diego, CA	Early	11-20	VC website	It has developed a tool for venture funding called RubX, which allows making scientifically-based investment recommendations.
Lightspeed venture partners	US, Menlo Park, CA	Early, Late, PE	100+	Newspaper or blog articles	The firm has recently become increasingly vocal about using data to drive investment decisions, but there is no specific information.
Lux Capital	US, New York, NY	Early, Late, PE, Debt	21-50	Newspaper or blog articles	According to familiar sources, the firm has invested in its software stack and has publicly acknowledged hiring software engineers and data scientists, yet specifics about its undertakings are not public.
Menlo Ventures	US, Menlo Park, CA	Seed, Early	51-100	Newspaper or blog articles	The firm has developed its proprietary software, Menlo Signals, which is used as a sourcing engine and benchmarking tool. Its model tracks a range of metrics

Name	Location	Stage	Size	Source	Notes on data-driven VC type
					from daily and monthly users to chart position, page views, and appearances across the web.
Montane Ventures	ASIA, Mumbai, IN	Early	2-10	Newspaper or blog articles	It is reported to have recently started applying data to VC, but there is not a lot of data available.
Nauta Capital	EU, London, UK	Seed, Early	11-20	Newspaper or blog articles	They reported building a prediction engine and a dead flow engine, but not much information is available in the public domain.
Next View Ventures	US, New York, NY	Seed, Early	2-10	Newspaper or blog articles	It publicly acknowledged using data mining software to evaluate investments and discover market trends back in 2013, but there have not been any other updates.
NFX	US, San Francisco, CA	Seed	21-50	VC website	It has developed Signal, an investing network for founders, investors, and VCs.
Omers Ventures	CA, Toronto, ON	Early, Late	21-50	Employee's reports	It sees data as a competitive advantage for ventures, and it is working on developing its own software.
OpenView Venture Partners	US, Boston, MA	Early, Late, Debt	51-100	Newspaper or blog articles	They reported using data analytics in 2012, but no other statement has been made.
Origin Ventures	US, Chicago, IL	Early, Late	11-20	Newspaper or blog articles	Since 2015, the firm has been developing its proprietary software, which allows them to ingest various data, filter and apply themes against data sets, and use various calculations to determine potential investment opportunities.
Pilot Growth Equity Partners	US, San Francisco, CA	Late, PE	11-20	VC website	They have developed NavPod, which is proprietary software that automates 90% of deal sourcing and helps analyze the firms' investments.
Project A	EU, Berlin, DE	Seed, Early	100+	Newspaper or blog articles	It defines itself as an "operational VC" and has made public its data warehouse infrastructure.
Quake Capital	US, Austin, TX	Seed, Early	11-20	Newspaper or blog articles	It uses a hybrid approach that leverages a Web Interface to collect information from applicants and collect it with information scraped from the web to then apply predictive analytics on opportunities.
Quantum Light Capital	EU, London, UK	Early, Late	/	VC website	This new VC founded by Revolut's founder is built as a technology company and employs AI scientists and engineers. They claim to have a proprietary quantitative decision engine, but there is no evidence yet.
Redstone	EU, Berlin, DE	Seed, Early	21-50	VC website	They are vocal about using AI-driven deal sourcing and deal execution, due diligence, and portfolio management.
Right Side Capital Management	US, San Francisco, CA	Seed	2-10	Newspaper or blog articles	They developed custom-built software to streamline operations.
Rocketship.vc	US, Los Altos, CA	Early	2-10	VC website	They are vocal about investing in companies using models built through data science. Their proprietary Escape Velocity algorithm identifies companies with a Sustainable Growth Engine at their core.

Name	Location	Stage	Size	Source	Notes on data-driven VC type
Root Ventures	US, San Francisco, CA	Seed, Early	2-10	Employee's reports	The firm's partners mentioned building custom software tools on Twitter, but unfortunately, no specifics are available.
Scale Venture Partners	US, Foster City, CA	Early	21-50	Newspaper or blog articles	It is reported to take use-cases of hundreds of successful startups to benchmark companies' "Vital Signs".
Sequoia	US, Menlo Park, CA	Early, Late	100+	Newspaper or blog articles	It is reported to be using programmatic sourcing systems to improve deal flow, but no publicly available information exists.
Signal Fire	US, San Francisco, CA	Seed, Early, Late	21-50	VC website	It has developed Beacon Talent, which is a proprietary AI-based recruiting platform that tracks the world's top engineers, data scientists, product managers, designers, and business leaders.
Social Capital	US, Palo Alto, CA	Seed, Early	21-50	VC website	They claim to leverage ML to build companies and help them operate excellently to drive long-term success. They have built an automated system to invest in startups without meeting them.
SOSV	US, Princeton, NJ	Seed, Accelerator	100+	Newspaper or blog articles	The firm has developed proprietary tools aimed at matchmaking, knowledge sharing, and streamlining operations.
Switch Ventures	US, San Francisco, CA	Seed, Early	2-10	VC website	They claim to use data to identify talent, leaning on a unique dataset to identify the founders who will outperform the venture industry. There is not much information about the dataset.
Telstra Ventures	US, San Francisco, CA	Early, Late	21-50	VC website	They claim to be sourcing 15% of deals from data science recommendations.
Tribe Capital	US, Menlo Park, CA	Early, Late	21-50	VC website	It is reported to use "the 8-ball", an automated diligence tool to assess potential new investments and help the portfolio companies. It allows quantitatively measuring product-market fit and growth trajectory. Former Social Capital employees launched it.
Two Sigma Ventures	US, New York, NY	Late	21-50	Newspaper or blog articles	They publicly admitted exploring technology and data usage at their firm, but they still claim that their process is still very human-driven.
Ulu Ventures	US, Palo Alto, CA	Early	2-10	VC website	They are vocal about having a data-driven decision analysis process. They have created a proprietary, disciplined investment process that aims to reduce risk and produce better, more consistent returns.
Union Square Ventures	US, New York, NY	Seed, Early, Late, Debt	11-20	Newspaper or blog articles	According to insiders, the firm has invested heavily in its technology stack. Unfortunately, there is hardly any data publicly available to verify and expand on this claim.
Unusual Ventures	US, Menlo Park, CA	Seed	21-50	Newspaper or blog articles	They have reportedly been working on a proprietary platform called Get Ahead to help entrepreneurs pool information and access firm resources.
Venture/Science	US, San Francisco, CA	Late	2-10	VC website	They use principles, models, computation, and AI to make investment decisions and deploy capital to recognize and reduce biases. They have built a scoring system that uses AI and decision theory to assess every opportunity's attributes and determine the associated risks.

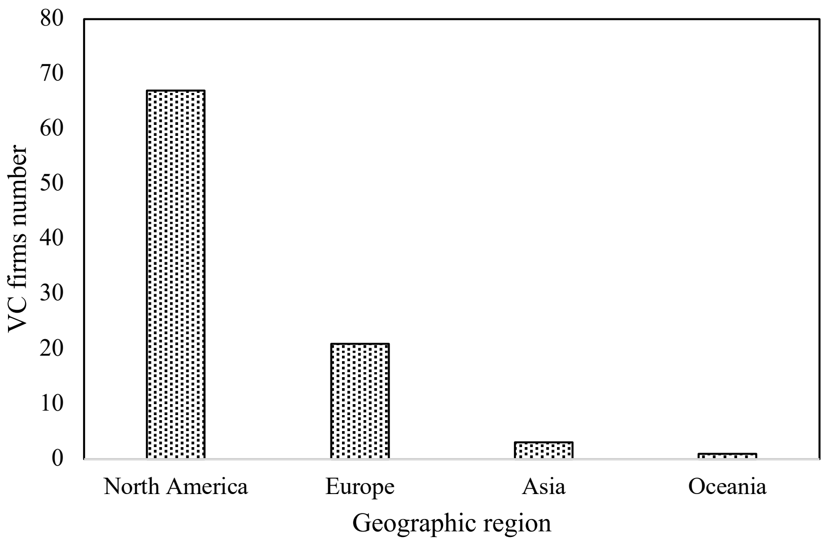
Name	Location	Stage	Size	Source	Notes on data-driven VC type
Venturerock	EU, Amsterdam, NL	Early	11-20	Employee's reports	The firm claims that thanks to its proprietary software platform, it was able to digitize and standardize the entire investment process. The firm has reportedly also built a tool to allow its portfolio companies to connect to its extended network of experts and advisors easily.
WRH Hambrecht Ventures	US, San Francisco, CA	Early, Late	21-50	VC website	It employs a hybrid strategy combining humans and technology. It has developed the Market Exaptation Simulation Engine ('MESE®'), an advanced computing and big data platform that provides unique insights into private markets and can pinpoint early disruptive companies around the globe.
Y Combinator	US, Mountain View, CA	Seed, Early, Accelerator	51-100	Newspaper or blog articles	This accelerator uses data and AI to automate some of its processes. It has software called HAL to help screen applications.

The complete database, including the data staff variable, is available online as an interactive version and accessible through the link presented in *Appendix 2*.

4.1.1 Classification by Geography, Stage, Size, and Reporting Source

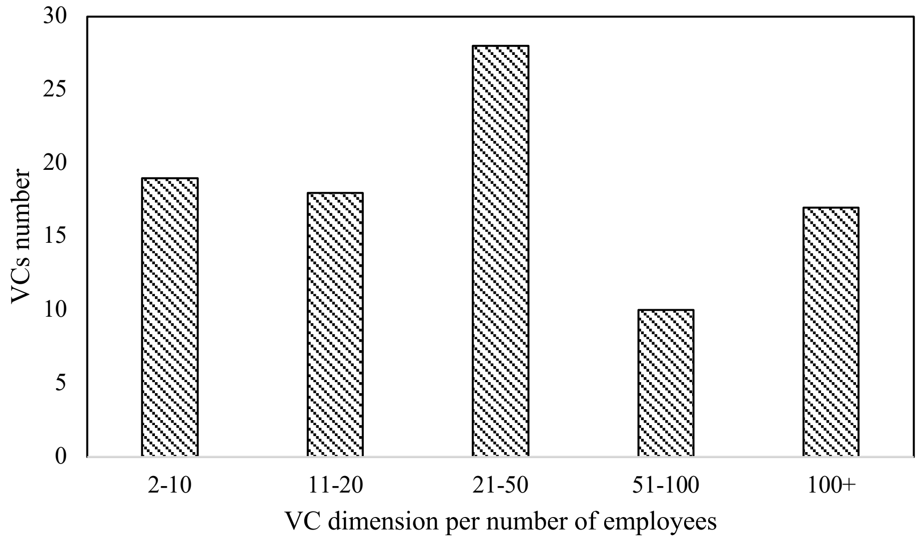
Regarding geographic location, most data-driven VCs operate in North America, followed by Europe and Asia (*Figure 6*). The supremacy of US VCs follows the distribution of VC activity at the global level, considering that the US attracts the largest part of VC fundraising globally. Surprisingly, only three funds in Asia are reported to be data-driven, even though it is the second region by VC capital financing globally. In each geographical area, data-driven VCs concentrate close to technology hubs, namely the Silicon Valley in the United States and London, Paris, and Berlin in Europe. However, the database also shows evidence of a small number of data-driven VCs that have established headquarters outside of innovation-intensive areas. Examples are Connectic Ventures, which is located in Kentucky, and Hummingbird Ventures, which operates in Gent, Belgium.

Figure 6 - Number of Data-driven VC Funds by Geographic Location



Considering the fund investment strategy, there is evidence of data and AI usage at all stages of investment. Out of the entire list of data-driven VCs, 55% operate at the early or seed stage only, 37% invest both at the early and late stage, and 8% opt only for late-stage investment. In terms of dimensions of VCs adopting AI, information gathered demonstrates that firms of different sizes can leverage data. *Figure 7* shows the number of data-driven VCs per number of VC employees, highlighting that VCs of every size are starting to implement data in their operations.

Figure 7 - Number of Data-driven VC Funds by VC Dimension



After assessing the standard indicators of geography, size, and investment stage, a classification in terms of reporting source has been performed to evaluate the transparency of the market. While carrying out the investigation, it has been observed that there is a high variance among the communication methods of VCs' data strategy. Some firms are vocal about their technological implementation practices, while others prefer to maintain a certain level of secrecy. To quantify the level of transparency in the market, three classes of sources have been identified (*Table 4*): (i) VC website, (ii) Employee reports, and (iii) Newspaper or blog articles.

For each class of reporting source, a level of transparency has been assigned. The VCs that are vocal about data or AI usage on their website are contributing to a high level of transparency in the market. One of the most salient examples is EQT Ventures which has an entire section of its website dedicated to "Motherbrain", their proprietary AI software that claims to redefine Venture Capital. However, even though these VCs are advertising their data strategy, they still lack in providing details. Thus, improvement is still possible to provide a higher level of transparency. The second class refers to employees' reports. Investors' and partners' statements are a reliable source of information but are considered less undisputable compared to the first class. As personal experience might provide a subjective view of reality and employee statements might not be legitimized by the employer itself, this category has a medium level of transparency. The third and last class refers to data gathered from newspapers or blog articles published by third-party sources external to the VC. It is considered the category with the lowest transparency, as information is not originated directly from the company or its affiliates. Nevertheless, these VCs are frequently referred to in articles on the topic; as such, they deserve to be further investigated and monitored for possible data strategy announcements in the future. The data gathered on the total sample reveals that transparency within the data-driven VC sector is relatively low, considering that 63% of VCs in the database are reported from newspapers and blog articles sources, which are characterized by the lowest level of transparency.

Table 4 - VCs Transparency Level Classification

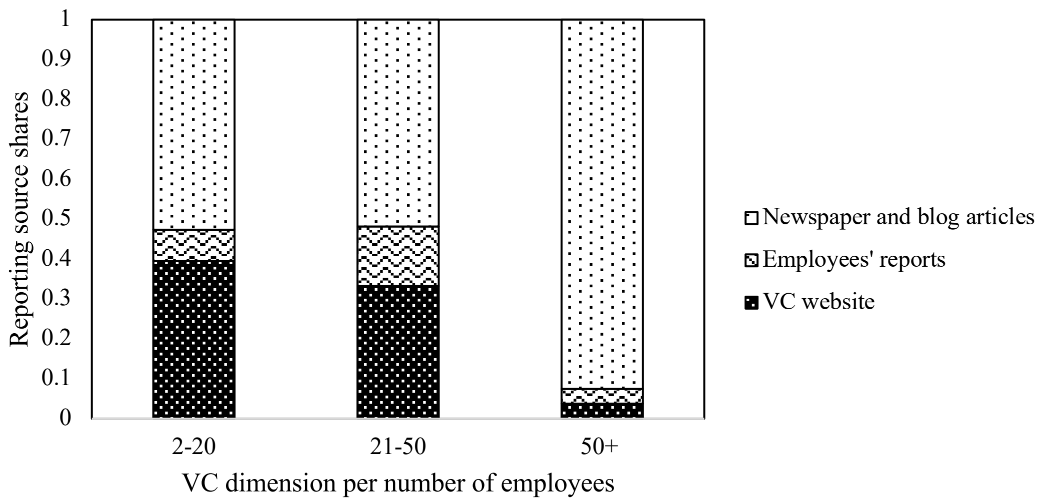
Reporting Source	Location	Reporter	Transparency	Number of Cases
VC website	"Home" or "About" sections of the VC's website or landing page	VC as an institution	High	26
Employees' Reports	Blogs, social networks, articles written by VC employees, references in interviews by VC stakeholders	VC employees	Medium	8
Newspaper or blog articles	Articles and posts found online	Journalists, researchers, third parties	Low	58

4.1.2 Market Trends

When analyzing the data collected, two market trends have emerged as significant and worth mentioning. One is related to employee mobility and the creation of new data-driven VCs, and the other is related to the lower transparency of larger VCs.

The first trend refers to the market tendency to create new data-driven VCs from spin-offs of pre-existing data-focused venture firms. In particular, three instances of employees leaving their company to found new VC firms have been identified in this database. The first case refers to SignalFire, which was founded in 2013 by Chris Farmer, a former partner at General Catalyst and Bessemer Venture Partners. The second case refers to Tribe Capital, founded in 2018 by three former Social Capital partners, Arjun Sethi, Jonathan Hsu, and Ted Maidenberg, who were interested in developing a VC that could strongly rely on data analysis and prediction. Finally, the third case refers to Blossom Capital, founded by Candice Lo, Imran Ghory, Mike Hudack, and Ophelia Brown in 2018. Two out of the four founders were former employees at Index Ventures. In each of these three cases, partners working on data usage and technology implementation in their former company have decided to further investigate data science applications in VC by founding their own VC firm. The three newly created companies have the common trait of being vocal about data usage, having it as one of their value propositions. Both SignalFire and Tribe Capital are reported as data-driven on the VC website. Blossom Capital information comes from an employee's report and various internet sources that corroborate its extensive data commitment.

Figure 8 - Type of Reporting Source per VC Size



The second market trend refers to the tendency of larger VCs to be less transparent than smaller VCs. In this context, where the average firm dimension is reported to be 14 employees, VCs with more than 50 employees can be classified as large VCs firms. In the sample collected, there are a total of 27 large VCs, 10 that range between 51 and 100 employees, and 17 that exceed the 100-employees threshold. By cross-analyzing reporting source and dimension data, it has been discovered that large organizations generally do not report data usage directly. Instead, it is third-party claims in newspaper or blog articles that inform their technology strategy.

From *Figure 8*, it is noticeable how smaller size (2 to 20 employees) and medium size (21 to 50 employees) VCs' data strategy is reported in 53% and 52% of the cases respectively by newspaper and blog articles, with the rest of cases fulfilled by employees' reports and VC website reporting sources. This evidence points to an average medium level of transparency provided by small- and medium-sized VCs. Instead, larger VCs' data usage is reported by third-party sources in 93% of the cases. This result is 41 percentage points higher than the one that emerged when considering VCs with fewer employees. Out of 27 cases of large VCs, only two are evidenced from the VC website and Employees' reports: Canadian-based Georgian Partners and Berlin-based Earlybird VC. These results point to relatively low transparency about internal data application practices in large VCs, signaling the propensity of bigger organizations to operate under the radar to retain competitiveness in the market.

4.2 AI Technology Applications

The second theme that has emerged in the analysis is related to the technological aspect of AI. As evidenced from the database, there is not a one-size-fits-all AI implementation scenario to follow for VC firms. Specifically, there are various factors to take into consideration when deciding to adopt AI. VCs must decide whether to develop the software in-house or rely on third parties. They additionally must agree on the functionalities of the technology to be implemented and set out a technology maintenance plan.

First, firms need to take a make or buy decision, evaluating the benefits and constraints of each option. While developing software in-house allows for greater control and a higher possibility of customization, it represents a consistent investment both in terms of time and financial resources. For these reasons, some VCs have opted for using third-party software. However, different technologies require different outsourcing strategies. For standard technologies embedded in the creation of digital Customer Relationship Management software (CRM), many VCs have opted for using third-party developed options such as Affinity[8]. Since CRMs have become commonplace even among traditional VCs, the software options that exist on the market are highly efficient and allow companies to reap the benefits without investing large amounts of capital in developing their own version. Instead, when considering AI software, the decision is not as straightforward. Most VCs have opted for building their own AI models, while others have started subscribing to specific AI-led deal scouting platforms, such as Specter and Harrmonic.ai[9]. Third-party AI software options for VC are still limited in number and not very diffused.

Second, VCs must decide which type of AI to implement and at which stage of the investment process. Developing proprietary software leaves more freedom of decision. As evidenced in the database, there is a high variance in AI application practices among VCs. Generally, companies have mainly decided to implement AI in two phases of the VC process, either pre-investment or post-

[8] Affinity is a CRM software that allows to manage deals in a more efficient way. Find it at https://www.affinity.co/solutions/venture-capital

[9] Harmonic.ai is a platform that gathers millions of datapoints on startups, find it at: https://www.harmonic.ai; Specter tracks data points and provides data insights to identify opportunities in the market. Find it at: https://tryspecter.com

investment, or in both. Pre-investment AI usage focuses on deal sourcing and screening. Post-investment implementation is centered on the development of platforms that leverage AI to support companies' growth and development by providing investors and entrepreneurs matchmaking or knowledge sharing among communities.

Finally, VCs need to set up a strategy to maintain data platforms and software over time, possibly hiring employees for this task. Specific data capabilities and skills should be internalized to spot future opportunities and implement further developments to become a data-driven VC.

4.2.1 AI Systems Development and Practical Functioning

When VCs decide to adopt AI, they need to understand the technology-related capabilities and how they can be of practical value. As publicly available information shared by the VCs who are vocal about AI implementation is not detailed enough to provide in-depth explanations of technology applications to VC practices, interviews have been conducted to gain a perspective on this matter. A summary of the technological insights from Redstone, InReach Ventures, Hatcher+, and Pilot Growth Equity is presented in *Table 5* and explained in the following paragraphs to provide examples of current practical AI applications in the VC industry.

Table 5 - Technology Application Cases from Interviews

VC	Type of VC	AI usage	Use-cases	Technology used	IP
Redstone	VC-as-a-Service, VC investing in seed and early stage	Deal Sourcing and Screening	Deals search, competitive analysis, market mapping, document generation	Unsupervised learning NLP	Proprietary software, not licensed
InReach Ventures	VC investing in seed and early stage	Deal Sourcing and Screening	Deals recommendation	Supervised learning	Proprietary software, not licensed
Hatcher+	VC-as-a-Service	Deal Sourcing and Screening	Deals search, deals recommendation, document generation	NLP	Proprietary but licensed to clients
Pilot Growth Equity	VC investing in late-stage, PE	Deal Sourcing and Screening	Deals recommendation	Unsupervised learning	Proprietary, not licensed

1. Redstone

Redstone was founded eight years ago in Berlin, Germany. It was born as a VC-as-a-service company advising large corporations approaching the VC ecosystem through Corporate Venture Capital (CVC). Over time, the VC twisted its business model primarily becoming fund managers, reducing the importance of the VC-as-a-Service provision to companies. Redstone currently operates in six countries and invests in the seed and early stages. The VC focus is on raising specific sector funds to permit the promotion of startups in a highly-effective way and to generate greater returns from specialization. Up to now, they have created ten funds focused on digital industries such as FinTech, Digital healthcare, and RetailTech and have raised a total of € 421 million.

Redstone was originally founded with the vision of combining VC and data analytics to understand, disrupt, and reinvent methods of Venture Capital. From early CVC experience, Redstone has developed market intelligence and due diligence tools powered by data to support corporate efforts. Over the years, the VC has created a proprietary database called Phoenix that collects data from various providers, including structured and unstructured databases. The raw data has been cleaned and standardized, resulting in a database composed of more than 2.2 million companies. For each company, details on the industry, financial transactions, founders, and team are gathered and categorized creating taxonomy values and company tags. The database is leveraged through a proprietary system of tools called Sophia, which is used for two main use-cases: deal sourcing and market mapping. In the first case, an NLP algorithm was developed to search for startups and competitors in the database. After providing a string input, the algorithm decomposes each company into a multi-dimensional vector based on its taxonomy values taken from the database. Then, the algorithm matches the input given with cases in the database through K-Nearest Neighbors and Average Nearest Neighbors ML models. Results are ranked in terms of accuracy and similarity, providing recommendations on the best deals or most similar competitors that match input criteria. In the second case, Sophia uses a similar NLP algorithm that allows to auto-generate lists of companies that match the firms' investment focus, providing a continuous market mapping. Redstone also claims to use AI in deal screening thanks to an algorithm that can partially automate the production of investment one-pagers and recommendations. However, it is recognized that human intervention is still needed for this specific use case. The VC is also working on a new system called Hermes, which is aimed at automatically tracking companies. The new tool will give real-time updates on the "hot companies" in terms of rounds raised, technologies developed, and salient events in the lives of founders and team members that might signal startup success. Developers are currently at the data collection and training stage, but they aim to finalize the project by the end of the current year. They believe that it can contribute to creating a competitive edge in the market, allowing them to beat the average investment timing of present investors.

The algorithms were developed internally, and the systems are proprietary and only accessible to the investment teams. The VC does not license to competitors nor provide access to eventual corporate clients in case of CVC advisory. The interviewee also reports that to build Sophia, the most time-consuming part of developing the AI software was collecting and standardizing the data. After that, it took about six months to train and develop the first version of the model, which is currently still updated frequently.

2. InReach Ventures

InReach Ventures is a London-based VC founded by Roberto Bonazinga, Ben Smith, and John Mesrie in 2015. They advertise themselves as the "AI-powered" VC. The founders' mission is to create and use proprietary software to scale early-stage investment across Europe and develop a new investing model that ensures complete alignment between investors and entrepreneurs. InReach invests in seed and early-stage European companies, and so far, they have closed a € 53 million AI-led fund focused on SaaS consumer application, internet, and marketplace companies.

From a technology perspective, InReach's proprietary AI architecture is based on three distinct layers: data, intelligence, and workflow. The data layer collects data from more than 300 sources and aggregates and enhances information to create a large set of original data. The intelligence layer makes sense of millions of data points by classifying companies according to different business styles and features. Then, an ensemble of ML algorithms combines information to provide companies recommendations through a raking system. The algorithm works on various types of information, such as founders' and team members' backgrounds, company location, funding configuration, recent financial transactions, and app downloads. Then, it matches the results with the VC investment strategy. The ML algorithms are trained every week with supervised learning. The VC believes that by leveraging past investors' decisions, it can recommend to them the companies that closely match their interests. The interviewee suggests that this approach allows for maintaining a flexible recommendation system that moves with investors' current focus in terms of industry, geography, and stage. The last layer is deal flow. The software recommendations enter the company's traditional deal-flow funnel, and a proprietary CRM system supports investors in the process of contacting and reviewing deal proposals manually. Thus, AI at InReach is leveraged only in the deal sourcing and deal screening investment stages, while the other operations in the VC value chain are approached traditionally.

InReach Ventures states that they started developing the AI software in 2016 and that it took about three years to have it fully operating. Finally, the software is under continuous improvement and, as such, requires frequent maintenance and overview. The ML algorithms are proprietary and not licensed to third parties. The interviewee was also reluctant to provide specific information on the types of ML algorithms used as it is the company's policy to maintain secrecy around the proprietary software functioning.

3. Hatcher+

Hatcher+ is a VC-as-a-Service (VCaaS) company based in Singapore that was founded in 2016 by Dan Hoogterp, John Sharp, and Wissam Otaky. The company offers a data-driven venture firm that uses AI and ML-based technologies to identify early-stage opportunities. VCaaS works by empowering other venture firms, accelerators, or companies by providing them with a pre-sourced and screened deal pipeline ready for direct investment to create a venture fund that follows a defined strategy. Hatcher+ has so far helped invest in 140 companies, focusing on early-stage startups.

Hatcher+ defines itself as Venture 3.0 because it has built a platform called Hatcher+VAAST that automates some steps of the VC process. The platform leverages a standardized dataset with more than 220,000 companies collected from different data sources such as Crunchbase, Pitchbook, and in general, the web. The database is also enriched with companies' applications through partnerships with various accelerators. The platform leverages this database for deal sourcing and deal screening. The first feature is called Deal Scout and allows fund managers to identify emerging business opportunities through NLP algorithms by providing search inputs such as stage, location, and industry. The second feature of the platform is deal recommendation. Working with more than 60 APIs, the software can score deals by comparing them to the past performance of companies in the market. The final score is built on eight main metrics: return potential, funding potential, venture trend, geographic potential, CEO exit potential, CEO fundraising, impact readiness, and exit potential. The software can additionally identify companies' SDG goals compliance and web traffic across their website, app, and social networks. Moreover, it provides an entirely AI auto-generated investment memo for fund managers giving insight into singular metrics scores. Finally, the platform also provides a CRM for clients to manage their deal flow and visualize performance analytics through dashboards.

The interviewee was not well-prepared to discuss specific technical aspects of AI algorithms training because of lack of data science background. However, the interviewee disclosed that it took almost one year and a half to collect and clean up data of about 600,000 VCs transactions in the time span of 22 years and to use it to train the AI algorithms subsequently. The interviewee also mentioned that they currently have an AI project in the pipeline to forecast founders' success based on specific success metrics.

4. Pilot Growth Equity

Pilot Growth Equity is a Private Equity and Venture Capital firm based in San Francisco that was founded in 2011 by John DeLoche, Neil Callahan, Rob Walker, and William Lee. It claims to differentiate itself from competitors thanks to its AI-driven deal sourcing practices and portfolio company value creation from networks and experience in the field. The firm invests in late-stage and growth stage B2B technology companies specializing in industries such as Cybersecurity, FinTech, Healthcare Technologies, and AdTech. The VC invested in 16 companies through two funds and realized eight exits.

The company leverages AI through a platform called NavPod, commonly referred to by investors and partners as "Chief Deal Sourcing & Workflow Officer". NavPod is proprietary software that provides deal recommendations through automated deal sourcing and screening. It actively tracks more than 3,000 companies looking for a point of inflections of growth, signaling the potential of investment. The AI algorithm is based on unsupervised learning and uses NLP models for breeding company data gathered from different sources. NavPod analyses information collected from the company website, public marketing materials, social networks, and existing companies' databases from data providers. To provide recommendations, it leverages financial data, state filings, patent filings and grants data, product releases, events participation, customer data, founders and team information, job postings, and competitors' data. Every day, NavPod sends an email to partners and investors signaling the top 10 recommended deals, which are automatically added to the proprietary CRM for the investment team to review and manually proceed with due diligence. The algorithm is built on the VCs investment strategy and only proposes deals relevant to the investors.

The interviewee has reported the success of the AI software, sharing that it has helped scout and initiate eight deals. Overall, it took about two to three years to develop NavPod considering data collection, model building, and training.

The interview cases reported show varied evidence of AI application and represent a good sample that allows for investigating different aspects of the topic. The AI strategies presented display common traits in technology usage, such as the consistent leveraging of NLP models and in broader data strategy, considering the prevalence of application in the first stages of the VC process.

4.2.2 AI Systems Maintenance

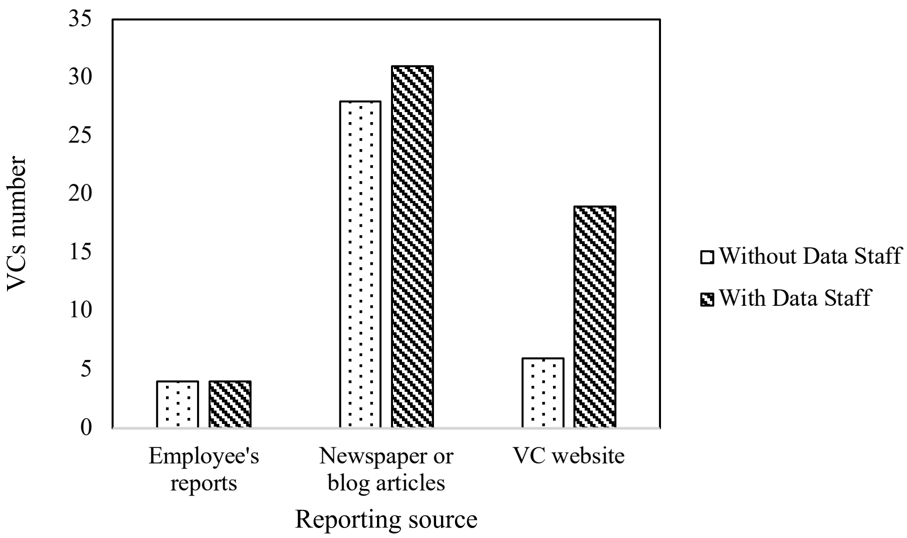
Once a new technology has been implemented, it needs maintenance and assistance. All the interviewees and stakeholders involved in this research have highlighted the importance of having someone in the organization tasked with software updates and improvements. Three out of four VCs interviewed have also reported having hired data scientists in their teams to manage the VC technology stack.

Redstone hired one full-stack developer, one data scientist, and one software engineer to develop their AI software platform, Sophia. They have also just welcomed a ML and NLP expert to their team to help them with the current AI projects in the pipeline. InReach Ventures, which is composed of a team of 13 people, boasts five employees focused on data analysis, representing 38% of the total workforce. The data team includes a lead ML engineer, a data scientist, a freelance data engineer, a senior frontend engineer, and a senior ML/backend engineer. The interviewee has claimed that the unusual employee specialization balance at InReach Ventures allows them to be structured as a software company and provide continuous improvement by releasing daily updates. Hatcher+ reports to have hired four data scientists. Moreover, Dan Hoogterp, a VC founder, has a software development and technology architecture management background. Pilot Growth Equity is the only company out of the four interviewed that does not have a data science member in the team. However, William Lee, founding general partner at the firm, was a computer scientist at Carnegie Mellon and has built the AI platform himself. The interviewee agrees that having a software expert in the team is a rare instance and that, in the absence of William, they would have to hire skilled employees to have in-house knowledge. The interviewee has also expanded, arguing that, considering the trend of data application in VC, he expects to see more and more data teams in PE and VC in the upcoming years.

After realizing the importance of having data scientists in the company, a check on all the companies in the database has been performed to uncover the presence of data experts in each VC. Information has been gathered from each VC website's team page, the corresponding LinkedIn page, and the company’s Crunchbase profile. The data collected is expected to be reliable due to checks performed across different sources. However, a small percentage of error must be taken into consideration due to the possibility of VCs not disclosing support team members to the public and employees potentially not having a LinkedIn or a Crunchbase profile. However, it is unlikely that VCs which are spending financial resources on data hires are not signaling it in any way. Thus, the data gathered is considered to be rather solid.

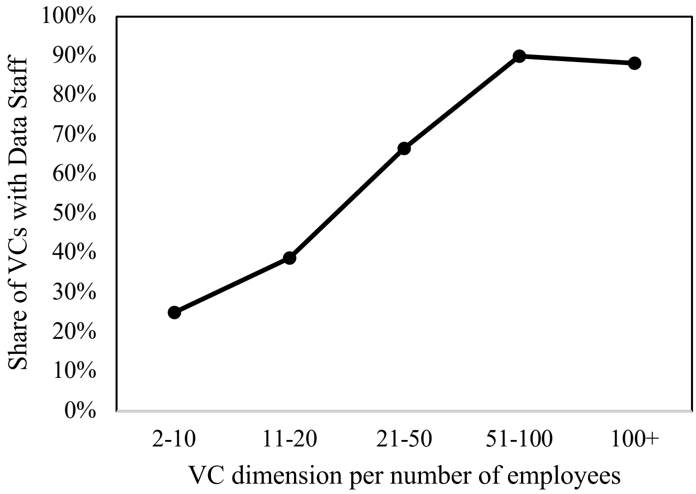
The analysis results show that out of 92 data-driven VCs, 54 have hired data staff, which corresponds to 59% of the total. The professional roles which are more common across companies are data scientists and software engineers. Some VCs also have research partners that, aside from performing research on market trends and new technologies, help the company develop or acquire AI software. An example is Francesco Corea, Research Data Lead at Balderton Capital.

Figure 9 - Data Staffed VCs Comparison per Rereporting Source



Diving deep into this topic, it has been observed that VCs that report their strategy with higher transparency are more likely to employ data-skilled employees than other firms. As depicted in *Figure 9*, 73% of the total number of VCs that report to be data-driven from their own website have hired data staff. Instead, only 50% of VCs reporting through employees' reports and 53% of companies reported through newspaper or blog articles praise to have data staff in their team. As the relative difference between the categories is consistent, it can be concluded that more transparent VCs are strongly committed to their data strategy vision, willing to invest money in competent resources to onboard in their company.

Figure 10 - Share of VCs with Data Staff per VC Dimension



Moreover, larger VCs seem to be keener to hire data scientists in their teams, as shown in *Figure 10*. This result is plausible, bigger VCs have a higher financial capacity for allocating capital to recruit data scientists and researchers than smaller VCs.

4.3 Data Strategy

The third relevant theme that emerged during the interviews refers to data strategy. As Venture Capital investment is traditionally considered a people business, AI implementation can be a controversial topic. This section will investigate the advantages and limitations of data usage in VC, as perceived by venture capitalists and stakeholders in the industry. Moreover, an exploration of opinions about the extent to which data could substitute humans' role in this sector is provided.

4.3.1 Benefits of AI in VC

Many VCs are starting to leverage AI technology in their investment process, hinting at the existence of benefits associated with this new practice. When multiple companies actively invest in a new trend, they signal their belief in its potential to support value creation or build an advantage above competitors in the market.

All the interviewees agree that the primary purpose of AI application to VC is to increase operational efficiency. Leveraging data allows VCs to reduce the focus of investors on a restricted number of companies that, through prequalification, are considered potentially good investments.

Redstone states that in the sourcing phase, "*instead of 1000 companies, AI allows you to look at only a few, narrowing down the scope and becoming more operationally efficient*" (Interview 1, min 20:45). On the same line, Hatcher+ argues that AI can bring consistent time savings to investors, stating that *"(AI) is going to give them (investors) more latitude to spend more time on the top 25 companies, rather than 100 companies. They (investors) can dig deeper and can get more efficient in deploying capital across those 25 (companies) or less"* (Interview 3, min 28:57).

Another benefit that AI provides to VCs is the ability to recommend deals that align with the VC strategy. By building an AI model internally and setting up the rules for its functioning, VCs define which are the relevant metrics that should contribute to a good investment. Practically, the deals recommended will be in line with the industry, geography, and stage of investment of interest of the VC. Pilot Growth Equity argues that their AI software has allowed them to stick successfully to their fund strategy: "*I think it keeps us focused on what we do because it is very easy to get distracted*" (Interview 4, min 17:49). All the deals proposed by the AI software, fall automatically within the company's investment criteria and technology allows them to keep their specialization. Moreover, AI technology permits VCs to specify which are the company's highly valued factors in decision-making. For instance, if the VC believes that founders and team structure are relevant indicators for startup success, it is possible to give a higher weight to such factors in the model. This increases the chances of finding startups that are a good fit for the investment firm.

All the benefits mentioned by interviewees currently pertain to the deal sourcing and deal screening stages. InReach Ventures argues that AI's goal is to find good leads that investors can later investigate in depth. Hatcher+ and Pilot Growth Equity also agree that AI aims to get them in contact with founders of promising startups. After this stage, all the companies interviewed rely on manual due diligence to continue toward the final investment decision. The main reason is that by that stage, the number of companies under consideration has considerably shrunk, and it is easier for investors to manage the volume of work by themselves.

Even though AI's benefits are currently predominantly leveraged in the early part of the VC value chain, success results associated with these practices are starting to emerge. Pilot Growth Equity reports that their AI platform has spotted eight companies in which they ended up investing. Redstone also shares that their AI software discovered three companies in the last 12 months in which they ultimately invested, which would have passed under the radar without the AI technology.

Aside from operational advantages, the question of whether AI can provide VCs with a competitive advantage in the long term is also valid. In this sense, VCs building their own proprietary software must be distinguished from VCs relying on third-party software. On one side, for the first category, it will be possible to build a competitive advantage thanks to data. Redstone believes that internally developing AI models early on will allow to collect data as time passes and provide the algorithms with better training. Moreover, as the company thinks that differentiation will happen in the way that data is collected and structured, first movers in AI implementation will be advantaged in the competitive landscape. On the other side, it will be more challenging for the second category to rely on data to outperform competitors. As third-party software is standardized by nature, the benefit will reside in how VCs will use them. Thus, it will be a matter of how well they will be integrated into the company and to which extent VCs will value data insights.

4.3.2 Data Limitations

Even though AI has been successfully applied at the organization level in various industries, including in the financial sector, its application to VC is limited due to data availability issues. The environment in which VCs operate is characterized by data scarcity. As a matter of fact, VC firms are dealing with investments in young private companies, which have reduced publicly available information. Besides, they can gather limited data points on deal opportunities, and due to the volatile nature of the market, information might not be updated or precisely mirror reality. For these reasons, AI implementation in VC risks providing a reduced and inaccurate prediction capability.

To mitigate the situation, companies gather data drawn from many different sources and then merge it to provide a comprehensive overview of each startup. General information on the company can be easily automatically extracted from companies' websites and the web. Information about the founders gets mined from social networking profiles and online records on their background. Data on financing rounds, patent grants, and other state filings can be gathered from public databases. Data on customers' perceptions and market trends can be collected from market research or by engaging directly with the final consumers through surveys.

However, it is not always possible to accumulate all the information mentioned for every single existing company. First, some startups might decide to develop their product or service in stealth mode, making themselves invisible to investors. In this case, the only possibility of spotting the deal early is to have personal connections with the founders. Second, the level of information availability varies depending on the companies' stage. Early-stage companies produce less information overall, which makes prediction exercises more challenging. Instead, late-stage and growth companies produce more signals in the market, including patent applications and financial information. Therefore, using data might result easier for late-stage investment VCs. Instead, early-stage focused VCs need to exploit the few startup data points they can gather. Redstone, which invests in the early stage, acknowledges the difficulty of data collection but suggests that it is possible to leverage founders' information. By using AI to discover the essential traits of founders' success, it would be possible to look for people with similar characteristics to make reliable future predictions. However, more research is needed to reach a point where accurate decisions on limited data sets can be made.

4.3.3 The Human Factor

Interviewees in the research have agreed that data is helpful for providing support to VC decision-making but that it cannot substitute humans completely, or at least not as of now. VCs are majorly using AI for deal sourcing and deal screening practices because they still believe that investors need to personally meet the founders and have conversations to get to know them. Hatcher+ states that "*you need to eyeball the founder, you need to ask those valuable questions, you need to feel the connection*" (Interview 3, min 26:03). Specifically for VCs that base their decisions on personal factors, it is vital to building relationships to evaluate the characteristics of entrepreneurs and estimate their resilience in business. As of today, AI cannot gauge the founders' sincerity, passion, and attitude. Nevertheless, Hatcher+ believes that with new technology developments in the future, it will be possible to approximate human-decision making to include personal and creative factors. Moreover, it will always be difficult to have data substitute investors' "gutfeel" for specific investment opportunities.

Even though AI technology is not equipped to make an investment decision autonomously, it is of great support in the early discovery and screening of deals, as evidenced by interviews. InReach Ventures argues that AI can reduce human inefficiency embedded in the daily screening of multiple companies, both in terms of time wasted and decision biases. As human judgment is by nature imperfect, data provides insights and suggestions to reduce noise and error in decision-making. In support of this view, Pilot Growth Equity thinks that AI in the deal sourcing and screening phases reduces recency and confirmation biases usually entailed in human decision-making. The interviewee believes that AI technology can be used as a tool to improve diversity in the industry. By forcing pre-determined rules in the AI algorithms, consideration of entrepreneurs pertaining to minorities could be increased, making capital deployment more equitable. Even though Pilot Growth praises AI as a valuable resource that can reduce human errors and biases, the VC believes that traditional practices are still relevant. Redstone has the same opinion and advocates for VCs' complementary usage of traditional processes and data: "*People think that VC is still a people business, that it is about network, which is, of course, true but to some extent. You have to merge the two things*" (Interview 1, min 22:19). Thus, VCs agree that it is still impossible to remove the human factor from venture capitalists' decision making.

Chapter 5: Discussion

The analysis has touched upon three pillars of AI application in VC, namely market, technology, and data strategy. Based on the findings, the discussion section will provide a condensed overview of market and firm implications of VCs' usage of AI. Apart from providing business recommendations and discussing practical implications of leveraging data for investment purposes, this section aims at enriching the literature on the topic under analysis.

The discussion primarily contributes to emergent studies in the area of AI usage in VC. By delivering a market and firm analysis of AI adoption, it sets the basis for continuing qualitative research in the field. The database created allows expanding literature on the adoption mapping of data-driven VCs. Furthermore, the research confers academic integrity to a subject that has only been explored in informal settings, such as online articles and blog posts. Thanks to extensive data collection, re-elaboration, and the application of a rigorous method for investigating the topic, the analysis performed allows to generate a picture of AI adoption in VC that is comprehensive and theoretically valid.

By studying the application of new technology in this industry, the discussion also contributes to the broader literature on the topics of VC and AI. Concerning VC, it adds to research focused on deal scouting. This dissertation suggests that data implementation in all the stages of VC investment is possible. Particularly, it advances an AI-led solution that aims at reducing VC challenges arising specifically in the first investment stages, which have been widely discussed in the literature. The considerations developed suggest ways for improving operational efficiency in VC daily operations, aiming at solving problems that concern the effective allocation of resources for value creation. Moreover, results also expand knowledge on VC decision-making, proposing a new model that includes traditional processes and data approaches to evaluate investment opportunities. Finally, the interviews presented relevant insights that corroborate literature findings on the importance of gut feel and human factors in this industry.

Regarding the AI research area, the analysis findings enrich studies on sectorial technological adoption. By providing practical examples of AI implementation by VC firms through database examples and interviews, it is possible to propose new AI use cases specific to the VC industry, which has not been extensively investigated thus far.

All in all, the following paragraphs will provide a condensed overview of the status of the current adoption of AI in the VC industry which is valid from the business and theoretical point of view. The discussion inspects the present market condition, debating adoption trends and characteristics. Moreover, it develops a best-practice model of AI technology adoption based on the evidence provided by stakeholders interviewed during the research. In particular, the model addresses how to realistically apply AI within the investment processes of a standard VC firm and proposes how to integrate traditional decision-making and data analytics. Finally, future opportunities and limitations for the application of AI in VC are discussed.

5.1 Current Market Status of AI Adoption

The task of quantifying the current adoption of AI technology in the VC industry is not straightforward and has not been explored before in literature. As such, the analysis has approached the research question by taking a broad perspective. First, the sample has been constructed analyzing data usage within the industry, relying on the assumption that VCs highly leveraging data are either currently using AI in their practices or could start implementing it shortly due to their predisposition to data analytics. Second, the cases in the sample have been used to examine and discover patterns across the variables concerning VCs' characteristics. Finally, trends of adoption have been observed.

The findings of the analysis highlight that AI adoption in VC is currently at an early stage. The research has resulted in collecting a total of 92 data-driven VCs worldwide. Considering the assumption mentioned above, it is reasonable to presume that only a fraction of these VCs is currently actively applying AI technologies to internal operations. Thus, only an appreciably small part of VC firms in the industry has hopped on the trend of AI adoption. Moreover, transparency in the market is deemed low. Information is still difficult to gather, and VCs generally prefer to maintain secrecy about details of technological practices at this stage of market development. Even interviewed investors were sometimes reluctant to share specific details of technology usage within their firm. Furthermore, as evidenced through the analysis of VCs' reporting sources, it is rare for firms to directly report their AI application strategy on official channels, such as their website. This market-wide attitude hints at a general willingness to maintain confidentiality during technology experiments at the current adoption stage.

In terms of general characteristics, geography, size, and stage of VC investment were analyzed during the research. Considering geography, technology diffusion has generally followed the global trends of VC activity. Data-driven VCs have emerged according to VC relative market sizes, with North America and Europe prevailing over other regions. The focus of activity has been registered close to innovation hubs, like the Silicon Valley. However, cases outside of VCs-intensive areas have also been mentioned, signaling that technology innovation is not restricted to specific areas. Instead, even isolated VCs are starting to implement AI thanks to its capabilities of spotting deals that overcome locational barriers. Asia is the only geographic area currently lagging behind, with a lack of data-driven VCs recorded in China. However, considering the vast size of the VC market and the fast pace of AI adoption in the region, this result seems odd. The absence of data-driven VCs in the region could be explained by the stricter information disclosure rules, which contribute to the difficulty of gathering testimonies in this specific area. In terms of size, there is no specific pattern of AI adoption from firms of different dimensions. Regarding the stage of investment, evidence shows that VCs investing in the early stage are the ones that are majorly adopting AI. Half of the companies listed in the database focus on funding early-stage companies only. As AI permits to spot companies at their outset, using this technology in the early stage when it is more difficult for humans to find patterns and discover new possibilities of investment gives VCs an advantage over competitors.

Thus, the VC firms presented in the database are considered the market early adopters who will set sector practices and drive future adoption. Even though, as of now, it is still not feasible to precisely quantify the speed of AI technology diffusion in the VC industry, it is possible to infer the characteristics of the forthcoming implementation. First, by observing real-life cases, one can predict that more and more firms will start adopting AI relatively soon. The topic has been recently put under the spotlight by newspapers announcing the creation of new data-driven VCs, such as in the case of QuantumLight Capital. Additionally, VC investors have started to update their medium blogs and write articles on the importance of data usage in VC. Second, it is possible to predict that the future adopters will likely be new funds created by external stakeholders or funds created by former employees of data-driven VCs who decide to leave their organization to pursue the data-empowering vision. As shared by interviewees, it could be difficult for traditional VCs to approach the AI paradigm. Thus, it is expected that it will take more time for them to join this trend, while it will be easier for data-minded investors. Finally, despite the current adoption predictions, interviewers have highlighted that data could take up to 10 years to become a commodity in VC. This insight once again emphasizes the initial state of development of AI adoption in VC.

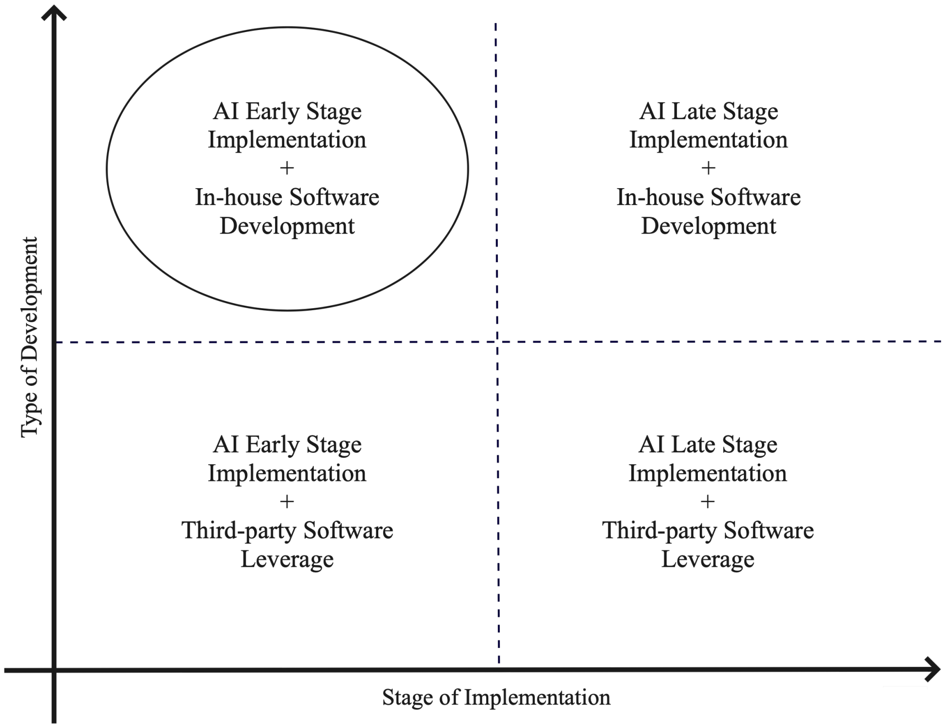
5.2 A Technology Model of AI Adoption in VC

Technology application at the firm level is highly varied in the VC industry, which makes it difficult to provide a general and exhaustive taxonomy of types of technology practices. Firms have not yet aligned on a standard AI application model in the market, as they are all experimenting to find the best setup that can provide for the highest business value added.

To facilitate the current discussion, VCs that are implementing AI were categorized across two variables. The first one concerns the source of software development, while the second one refers to the stage at which AI is implemented. Considering the first variable, it has been noticed that VCs can either apply AI by developing their software in-house or by relying on third-party software providers. This type of organizational decision comes down to evaluating the costs and benefits in terms of investment of resources and potential value creation. The second variable categorizes VCs by stage of implementation. Most VC firms are using AI in the pre-investment stage, particularly at the deal sourcing and screening stages. Others are leveraging AI in platforms aimed at helping companies in the post-investment stage. These variables create different configurations of VC firms applying AI (*Figure 11*). Some companies use third-party platforms to support startups' development, while others leverage them for deal sourcing. At the same time, some VCs decide to develop their own AI-led matchmaking platform while others develop their proprietary scouting and screening algorithms.

The categorization created allows for structuring the discussion of the analysis' findings by comparing them to the theoretical review of VC and AI. The following paragraphs highlight the investors' preferred AI adoption practices in terms of implementation stage and source of technological development. Finally, this section proposes a model of AI application in VC that considers best practices.

Figure 11 - VCs Classification Matrix

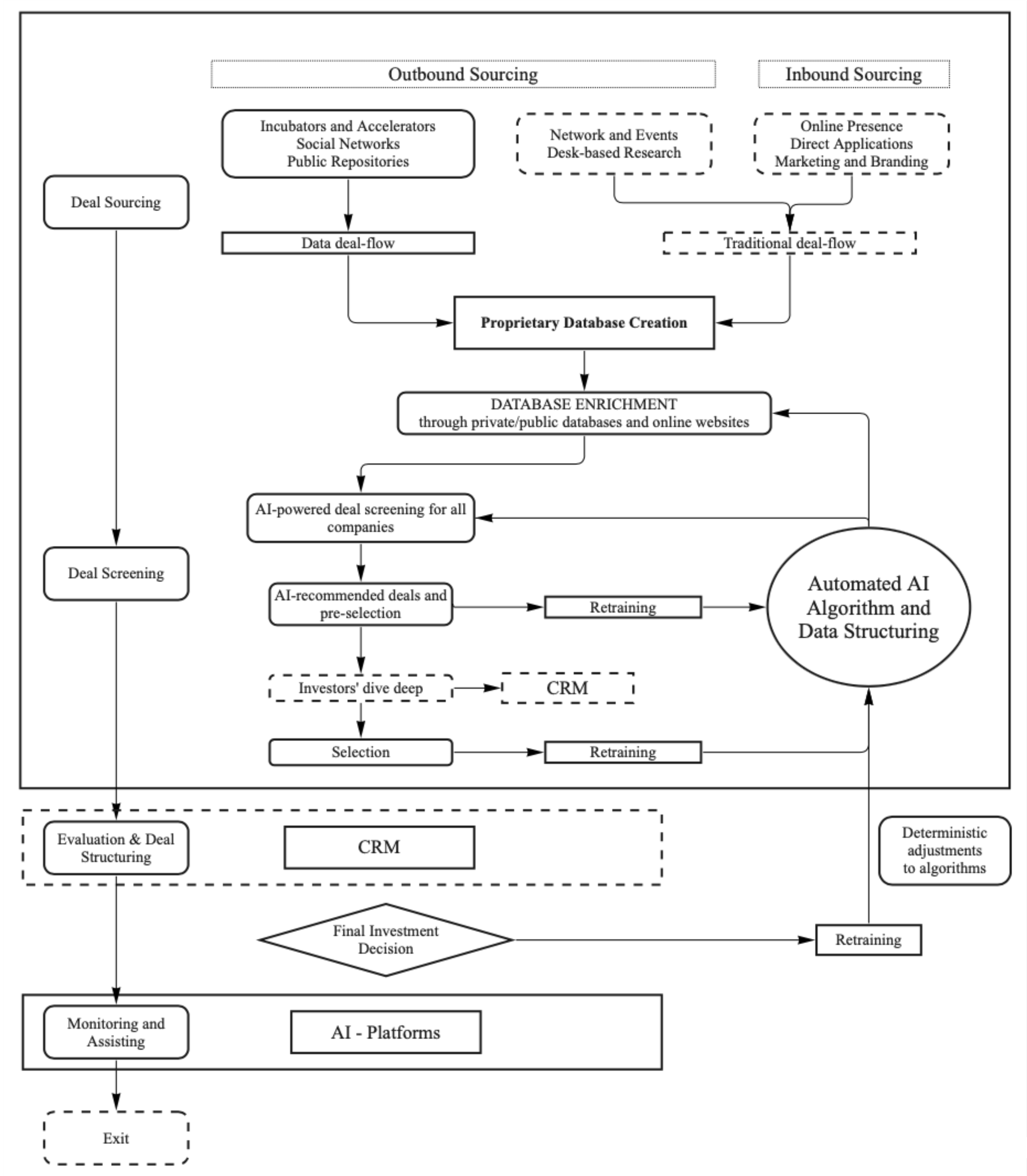


Both literature review and investors' opinions signal the superiority of AI applications at the earliest steps of the investment process. From the literature review performed on the theoretical background of VC, it has emerged that the first stages of the VC value chain are the ones that create the most value for management funds. Due to their relative importance, deal sourcing and screening take up most of the time of investors' work week. VCs employees have reported that almost 40% of their time is spent searching for new opportunities and reviewing them to evaluate their investment potential. As AI capabilities allow to improve and automate business processes, applying these technologies to the most labor-intensive steps of the process would reduce the average time required for these tasks, ultimately improving VC efficiency. As evidenced by the research performed, consistently with literature review expectations, AI is majorly used in the deal screening and deal sourcing phases of the VC investment process. Even if to a lower extent, AI is also used in the monitoring and assisting phase, which is the second most important VC value-added activity. Stakeholders interviewed believe that using AI specifically at the beginning of the investment process allows improving operational efficiency by spotting the most promising deals at an early stage. AI benefits investors because it allows them to automatically source deals and screen them by matching features with preset criteria, allowing them to look at a reduced number of proposals if compared to traditional VC operations.

The decision to develop software in-house rather than relying on third parties seems to affect the possibility of creating a competitive advantage in the VC market. Interviewees believe that competitive edge can be built only if the software is developed directly by the VC, without relying on external platforms. Even though purchasing third-parties services allows for faster and cheaper software implementation, it restricts customization possibilities. By subscribing to the same service, all the VCs would search for deals from the same pool of opportunities and benefit from an identical algorithm. This setup reduces personalization opportunities and thus precludes the creation of competitive advantage from technology. However, third-party AI technology would still improve operational efficiency, and VCs could still differentiate by specializing in industries, stages, and investment strategies. Thus, the research has highlighted the superiority of the in-house software development decision, as it would give VCs a competitive edge in terms of algorithm creation. Companies can create models that follow rules that strictly match the firm's investment strategy and value the features of startups and founders. Additionally, it would provide a seamless implementation within the organization, efficiently merging deals from web sources to those spotted from personal connections through the VC network.

Based on the experiences and beliefs reported by interviewed investors and the data gathered throughout the research, a model that summarizes AI practices in VC has been built. Drawing from the above-mentioned considerations, the model focuses explicitly on VCs that pertain to the first quadrant of *Figure 11*. Thus, it specifically theorizes on VCs that develop their AI software in-house and that mainly leverage it in the first stages of the VC process, as this configuration is believed to provide the most value added to companies. The model is presented in *Figure 12* and explains how to structure AI technology application in VC to augment human capabilities and benefit the company.

Figure 12 - Model of AI Implementation by VC Investment Stage



The first step of the VC process is identified with deal sourcing. As highlighted in the theoretical background section of this dissertation, VCs spot deals through outbound sourcing and inbound sourcing. Traditionally, outbound activities comprehend networking, events, and proactive search, and inbound activities refer to marketing and spontaneous applications intake. Aside from normal practices, data can further enable VCs' outbound sourcing. Thanks to data analytics, VCs can use web crawlers to spot deals more efficiently. By searching among incubators, accelerators, startup cohorts, social networks, public data repositories, and other online sources promoting new product developments, VCs can count on an additional data-led deal stream that feeds opportunities pipeline. In a data-driven VC, the inputs of both the traditional and the data-enabled sourcing converge into a vast list of existing deal opportunities. This list gets automatically and manually updated as new companies are spotted through data and manual search over time. The ventures' list constitutes the VCs' proprietary database of deal opportunities. Through AI and automation, VCs are able to enrich the startups' records with important company features useful for deal screening. VCs can gather information on ventures by scraping online sources such as websites, newspapers, and public and private databases like CrunchBase and Pitchbook. Moreover, when dealing with traditionally sourced companies, information can be gathered manually by investors to integrate available online resources. After data collection, the company features get structured into predefined categories. For instance, VC firms that mainly value team components for decision-making will search for specific information on the founders' experience, educational background, and connections. Instead, other VCs more focused on alternative factors might research geographical, industrial data, and patent information. Finally, the collected data gets standardized and cleaned up to reduce noise and allow for practical usage. These steps allow the VCs to build a proprietary database that includes thousands of startup companies and all their relevant information.

In the deal screening phase, previously developed and trained AI algorithms use the proprietary database information to recommend the best deals for the fund. The most common AI models currently implemented by VCs leverage ML and NLP algorithms and can be trained with supervised or unsupervised learning. The algorithms first screen the deals by setting a filter that concerns the VC fund strategy regarding geography, industry, and investment stage. Then, they use past cases to predict the deals' appeal. The details of this step are specific to each VC, as their algorithms remain secret and highly customized. The outputs of AI for deal screening vary from company to company, but generally, they provide a composite score that reflects the deal's attractiveness based on many variables, weighted differently depending on the VC strategy. Investors can either get periodic reports that include deal recommendations, or they can use online pages or companies' internal applications to review deals. Alternatively, some VCs have developed search interfaces leveraging NLP models that allow searching for opportunities according to specified attributes. The recommended deals present the highest score based on the inputs provided by the VCs. The effectiveness of the AI models employed in this stage, coupled with clear and organized database data structuring, allows VCs to discover the best deals in the market and potentially gain a competitive advantage.

Once the top deals have been identified, it is up to investors to continue screening companies. Partners and associates still need to meet the entrepreneurs to establish a connection that can lead to a profitable partnership. The next stages of the process are still highly manual but are enhanced by technology through CRM platforms that allow investors to keep track of their day-to-day operations. CRM platforms are also leveraged in the evaluation and deal structuring stages to monitor the deal's status and to record investors' opinions. VCs can develop their own CRM or acquire third-party software. However, developing it in-house is the preferred option as it allows for seamless integration between the deals recommended by the AI and the follow-up investigation by investors.

Every step of selection until the final investment decision is taken is used to retrain the AI algorithm. Retraining allows to reflect the VC investors' interest in companies and provides a better fit for future recommendations. Moreover, models can exploit cases of successful startups to update algorithms over time. However, due to the highly innovative environment in which VCs operate, deterministic interventions are also expected to adjust the algorithmic rules to reality. As technological developments and new trends in the industry are hardly foreseen by historical cases, VCs need to interpose their judgment and observations to allow for productive technology functioning. The mixture of retraining from selection decisions, analyzing successful cases in the VC market, providing deterministic interventions allows the AI algorithm to learn and develop more expertise in the long run.

Finally, AI can be leveraged in the monitoring and assisting phase once the capital has been deployed to the venture. Proprietary or third-party AI-enabled platforms are majorly employed by lead investors, as they retain higher stakes into supporting startups developing their untapped potential. These technologies allow to match investors with entrepreneurs and to find potential human resources to expand the startups' team and expertise. The usage of AI in this step allows to leverage platform economies and provides higher value to portfolio companies, which might consider it a discriminant when deciding the VC to rely on. The data-driven support continues until the exit stage is reached, eventually resulting in a successful IPO or acquisition.

AI can be used for other relevant use-cases along the whole VC process. For instance, it can be leveraged for writing automated memos. The deals are initially reviewed by individual investors, who must convince the other fund's partners to invest in a promising startup. By generating automated memos on the opportunity, investors save time and can easily rely on them to persuade other VC stakeholders. Another AI use case is market mapping. The AI algorithm, leveraging the information in the database and the scores assigned during deal screening, can provide snapshots of startup industries, suggesting the attractiveness of different markets.

The model presented provides an overview of the current practical application of AI technology in the VC industry. It aims to provide a general framework that enriches literature review on VC decision-making, by analyzing in detail data integration at each step of the investment process. Furthermore, it generates useful insights for many industry stakeholders. On one side, VCs that have already started leveraging data can use this model to reflect on their practices and improve them, implementing AI along the value chain. On the other side, it provides a structured map that shows how to build a data infrastructure for VCs that still rely on traditional methods. Moreover, the high complexity of the model hints at the necessity for VCs to hire specialized data scientists to develop and maintain technology. As evidenced in the model, AI is updated, retrained, and sometimes needs deterministic interventions. Thus, having an in-house skilled workforce allows to perform these actions continuously and to explore new possibilities for AI applications in the sector, potentially discovering new practices that can further revolutionize the industry.

5.3 AI Adoption Future Opportunities

The future role of AI in VC will be to support investors in their daily work to increase operational efficiency even further. Current AI use cases can be ameliorated, and new developments can be studied by either relying on present technology or waiting for future advancements. As the AI paradigm will shift from ANI to AGI and eventually ASI, technological progress and new frontiers in the industry will open opportunities for developing novel applications in VCs. As for now, VCs can either leverage the existing technologies by improving models already built or by developing new algorithms and including them in the technology stack. This dissertation will propose three new AI use cases developed by existing AI capabilities, which will enrich the literature on AI technologies applications in organizational contexts.

First, as NLP technology is particularly suited to the VC business, new models could be developed and used to perform accurate consumer and market research by leveraging surveys as inputs. Investing in a new product or service for venture capitalists is a leap in the dark, as the market, technology, and adoption risk are very high. By performing more market research through surveys, for instance, VCs could understand the attractiveness of the opportunity they have under scrutiny in terms of the potential of future consumers' adoption and market appeal. VCs could leverage NLP algorithms to analyze responses of consumers, as well as sentiment analysis on social networks to investigate market readiness.

Second, AI can be further used to predict human attitudes from behaviors. As many VCs are trying to discover the ingredients that make founders successful based on their backgrounds, AI could be leveraged to spot the most promising future founders. Through unsupervised models, AI could find patterns of behaviors or founders' characteristics that presumably can lead to startup success. Research on this topic is already ongoing, but there is still a lot to uncover before possibly leveraging these kinds of insights in investment decisions. As new developments will be discovered, VCs should leverage this knowledge in their deal sourcing and screening stages.

Third, VCs could try to partially automate the review of entrepreneurs' first pitches. In this case, they could leverage Computer Vision technology to analyze videos or live recordings of ideas presentations to study body language. This could be further connected to research on successful founders' behaviors and attitudes. Since this type of technology application is already used in other industries, it could be interesting to experiment with its functioning as a further screening tool in the investment process before passing to in-person meetings and manual deal evaluation and due diligence.

5.4 AI Adoption Limitations

Even though the evidence gathered in this research has proved the possibility of AI application to VC firms' operations, it also recognizes its limitations. AI can augment VCs capabilities by improving operational efficiency and possibly granting a competitive edge; however, it is still unthinkable for investors to rely solely on AI for decision-making in the VC space. Limitations of AI application in the VC industry mainly generate from two instances: the first refers to data-related elements and the second to the strong human factor of the VC job.

First, VCs face data availability issues due to the relatively scarce quantity of data, specifically in the early investment stage, as evidenced in literature. Even though future developments in AI technology could allow for models to be effectively trained based on fewer data points, as of now, data scarcity can cause biased or erroneous predictions. VCs can currently mitigate this limitation by building up a vast information database, leveraging online and proprietary resources. Moreover, they can structure it in a way that can give them a competitive advantage over time. Interviewees believe that AI will become a commodity in VC in the future. Thus, superiority in the sector will be a matter of how much data each firm possesses and how efficiently they can use it for investing purposes. The quest for information, however, embeds ethical risks. When gathering personal data on the founders and team members, privacy is a concern that VCs must take into account in their data collection activities. As more VCs start using AI, privacy will become an area of interest that investors will have to deal with.

Second, AI is not fit to accompany the nature of the VC job in all its aspects. As VCs' work consistently leverages personal relationships and impressions, AI is deemed inappropriate to fully measure certain characteristics, such as team resilience and attitude in problem-solving. Moreover, VC investors need to feel the connection with the startup to ensure a fit that can predispose a successful partnership along the way, both in terms of business opportunities and personality. Furthermore, it is mandatory to eyeball the founders before deploying capital. In a market where asymmetry of information is very high, trust within partners is crucial in the pre-investment stage where technology and execution risks are still high; thus, investing in an idea that is still at the very early stage requires human effort other than AI application.

While the first limitation can be mitigated by future technology developments and data collection strategies, the second one will be more difficult to overcome. The VC business is a people business and will always require a certain degree of human interaction. Human connection remains one of the most important factors in this industry, and gut feel still plays a role in taking final investment decisions. Moreover, capacity for negotiation and personal relationships can hardly be replicated by AI. Considering the entrepreneurs' point of view, they might also refrain from entrusting an utterly data-driven approach to VC. Entrepreneurs carefully choose the most suitable VC company that can support them with their idea by evaluating various factors, which also include the company's reputation. As VCs reputation is not only built by successful investment but also by participating to events, networking, and meeting with founders, retaining a personal component still matters. All in all, taking the human component away to rely solely on AI would degenerate the nature of the VC job and would only backfire on investors in the long run.

Chapter 6: Conclusion

6.1 Final Remarks

This research paper aimed at providing an overview of AI technology adoption in the VC industry, exploring its current application status and future opportunities. The topic has been investigated by addressing the research questions in the analysis and discussion sections. The current market status (*RQ1*) has been analyzed in *chapter 4.1* and discussed in *section 5.1*. The practical application of AI technology in the VC sectors (*RQ2*) has been addressed in *section 4.2* and further discussed in *chapter 5.2*. Finally, the benefits and opportunities (*RQ3*) and the potential limitations (*RQ4*) of AI implementation have been explored in *sections 4.3* and *5.3 and 5.4* of the dissertation.

From the analysis conducted, it is possible to conclude that AI adoption in VC is at a very early stage. In terms of market adoption, a scarce number of VCs have started to proactively implement new technologies and take on data-driven approaches to investment. The list of VCs produced by this study is considered to represent the group of market innovators or early adopters that will set up an industry standard for the adoption of AI technology and will boost the diffusion of AI innovation amongst new entrants. In terms of technology application, there is a high variance among firm practices. Even though the sector is characterized by low transparency of information, it was discovered that VCs generally apply AI in the first stages of the VC process, namely deal sourcing and deal screening. The most common AI use cases leverage ML and NLP models and comprehend startup search, recommendation provision, and memo writing. VCs have highlighted the added business value of developing AI software in-house and have shared the need to hire data scientists to maintain and update the firm's technology applications. From these insights, a model for practical AI technology adoption was created. The model overviews the technology stack setup for all the stages of the VC investment process and provides a tangible example of AI application to VC operations. All stakeholders in the industry can leverage it to improve their current data practices or find inspiration to experiment further and innovate. Finally, the study has highlighted that the most important benefit that AI can provide to VCs is efficiency in business operations. By leveraging technology, deals can be discovered in a timely fashion, possibly providing a competitive advantage in the market. By beating competitors in time, VCs adopting AI can spot the best deals and potentially generate more returns in the long run. However, due to data availability limitations in the early investment stage and the highly personal nature of venture capitalists' job, this research points to the still very relevant role that human interaction has in the industry.

This paper suggests that AI can play a relevant role for VC firms, increasing operational efficiency and providing better support to portfolio companies. Overall, it can give adopters a competitive edge, ameliorating their market position. However, this dissertation also acknowledges the impossibility of VC firms to rely solely on AI technology, disregarding traditional processes. The findings and implications presented throughout the paper hint at the optimal solution of developing a model of "*AI-augmented VC*" in which technology is leveraged along with human practices. In this ideal setup, AI is put at investors' service to improve their day-to-day operations, supporting them in most of the VC process stages. Investment decisions will rely gradually upon more data analysis, and firms will shift their strategies opting for a technological approach. However, the final decisions will always depend on human factors, which cannot be taken away from this industry. Thus, notwithstanding the innovativeness of AI technology, its adoption will hardly disrupt the VC industry, but it will create a new standard of VCs that efficiently integrate technological and human components to create more value.

6.2 Study Limitations and Further Research Opportunities

The study presented possible limitations that derive from the novel character of the research. First, it is recognized that information availability issues constrain the results' comprehensiveness. Due to the low transparency of the market and reduced companies' disclosure propensity, it is reasonable to suspect that the real-life cases of data-driven VCs might exceed the ones presented in this paper. However, all the market's key players that are currently applying AI and intensely leveraging data have been mapped, providing a moderately complete sample on which the analysis is based. Moreover, the limited accessibility of data on mapped companies reduced the analysis capabilities. Obtaining access to private databases would allow performing a more accurate analysis, including more variables. Second, this research might suffer from data generalization problems. When considering practical technology application, only four VCs have been interviewed. Even though the response rate is considered acceptable, considering the VC market's secrecy and investors' low propensity to discuss internal operations, the restricted sample might provide a limited view of the topic. It is possible that, by interviewing more VCs and analyzing more cases, a broader set of approaches might be discovered. Third, the data collection method employed for gathering information in a database might represent some shortcomings. Having applied a bottom-up methodology, accuracy in reporting might be biased by the author and the interviewers' experience. However, considering the novelty of the topic and the lack of objective metrics that can indicate data usage in VC, this method was deemed more reliable than opting for a top-down approach. Lastly, time represents a limitation to this research. The market adoption status provides a snapshot only until the end of data collection for this paper, excluding possible cases that emerged later on. Thus, due to the fast-paced change and evolution of the market, more cases might be present at the current date.

The findings reported in this paper enrich the literature on the topic and open future research possibilities. Upcoming studies might focus on market adoption, enlarging the presented database by adding more cases that could emerge in the future. Moreover, the already discovered companies could be validated by leveraging insider information, providing higher reliability in the analysis. After some years, the diffusion of AI technology adoption in VC could also be studied by using adoption models, such as Rogers' Diffusion of Innovation curve or other related economic frameworks. Furthermore, when more information will be available, it would be interesting to perform quantitative analysis to understand the potential of AI application in this industry analytically. In five or ten years, some startups backed through AI will potentially have reached an exit, and it will be possible to study their relative performance compared to normally sourced deals. Finally, future research might develop case studies on renowned VCs, exploring in depth their technology stack and the adoption process from an organizational point of view. As the topic is innovative, literature review on the topic will be gradually populated in the following years, with even more study opportunities opening up for future researchers.

References

Alexander, D. (2020). *Neural networks: History and applications*. New York: Nova. Retrieved on May 10l 2022 from https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2323785&site=ehost-live

Alexander, O., Finnerty, P., Geis, B. M., Kong, M., Libby, S., O'Callaghan, A., . . . Yazdani, D. (2020). *The power to reshape the future.* (). PwC. Retrieved on April 29, 2022 from https://www.pwc.com/gx/en/industries/financial-services/assets/wealth-management-2-0-data-tool/pwc_awm_revolution_2020.pdf

Allen, J. F. (1998). AI growing up: The changes and opportunities. *AI Magazine, 19*(4) Retrieved on -May 15,2022 from https://doi.org/10.1609/aimag.v19i4.1422

Amit, R., Brander, J., & Zott, C. (1998). *Why do venture capital firms exist? theory and canadian evidence* Elsevier BV. doi:10.1016/s0883-9026(97)00061-x. Retrieved on May 8, 2022.

Amornsiripanitch, N. (2019). *More than money* SSRN. doi:10.2139/ssrn.2586592. Retrieved on May 12, 2022.

Antretter, T., Blohm, I., Grichnik, D., & Wincent, J. (2019). Predicting new venture survival: A twitter-based machine learning approach to measuring online legitimacy. *Journal of Business Venturing Insights, 11* doi:10.1016/j.jbvi.2018.e00109. Retrieved on June 10, 2022

Arroyo, J., Corea, F., Jimenez-Diaz, G., & Recio-Garcia, J. A. (2019). Assessment of machine learning performance for decision support in venture capital investments. *IEEE Access, 7*, 124233-124243. doi:10.1109/ACCESS.2019.2938659 Retrieved on May 25, 2022

Asaftei, G. M., Alex, E., Bernow, S., Bielenberg, A., Boland, B., Christy, J., . . . Aditya, S. (2022). *Private markets rally to new heights - McKinsey global private markets review 2022.* (). Retrieved on April 19, 2022 from https://www.mckinsey.com/~/media/mckinsey/industries/private%20equity%20and%20principal%20investors/our%20insights/mckinseys%20private%20markets%20annual%20review/2022/mckinseys-private-markets-annual-review-private-markets-rally-to-new-heights-vf.pdf

Ayodele, T. O. (2010). Types of machine learning algorithms. *New Advances in Machine Learning, 3*, 19-48. doi:10.5772/9385. Retrieved on May 8 2022.

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide* Sage.

Böhm, M. W. (2017). *The business model DNA: Towards an approach for predicting business model success* Retrieved on June 10, 2022 from https://explore.openaire.eu/search/publication?articleId=od________518::35d22b119d6f810e8c34ac965fdce337

Bonaccorso, G. (2017). *Machine learning algorithms* Packt Publishing.

Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies, 2.* Retrieved on May 15, 2022.

Bughin, J., Hazan, E., Manyika, J., & Woetzel, J. (2017). *Artificial intelligence, the next digital frontier.* ().McKinsey Global Institute. Retrieved on May 10, 2022 from https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx

Caruso, C., Enriquez, F., Oshotse, A., & Pradeep, G. (2015). *Algorithmic venture capital predicting valuation step-up multiple in venture backed companies through deep learning techniques. Retrieved on June 10, 2022 from http://cs230.stanford.edu/projects_fall_2020/reports/55791766.pdf*

Cavana, R., Delahaye, B., & Sekeran, U. (2001). *Applied business research: Qualitative and quantitative methods*. Australia: John Wiley & Sons.

CB Insights. (2022). *State of venture Q2'22 report.* Report. Retrieved on May 2, 2022 from https://www.cbinsights.com/research/report/venture-trends-q2-2022/

Chalmers, D., MacKenzie, N. G., & Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice, 45*(5), 1028-1053. doi:10.1177/1042258720934581. Retrieved on May 15, 2022. Retrieved on May 10, 2022.

Charafeddine, M. (2016). The VC value chain (part 1). *LinkedIn.* Retrieved on May 10, 2022 from https://www.linkedin.com/pulse/vc-value-chain-part-1-mohamad-charafeddine/

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* Sage.

Chui, M., Hall, B., Signla, A., & Sukharevsky, A. (2021). *Global survey: The state of AI in 2021.* ().McKinsey. Retrieved on April 30, 2022 from https://www.mckinsey.com/business-functions/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine, 7*, 2050312118822927. doi:10.1177/2050312118822927. Retrieved on June1 17, 2022.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology, 13*(1), 3-21. https://doi.org/10.1007/BF00988593. Retrieved on June 17, 2022.

Corea, F. (2018). AI and venture capital. *An introduction to data* (pp. 101-110). Cham: Springer International Publishing. doi:10.1007/978-3-030-04468-8_15 Retrieved on May 20, 2022 from http://link.springer.com/10.1007/978-3-030-04468-8_15

Cote, C. (2021, July, 13). 3 key types of private equity strategies. *Harvard Business School Online,* Retrieved on April 29, 2022 from https://online.hbs.edu/blog/post/types-of-private-equity#:~:text=There%20are%20three%20key%20types,%2C%20growth%20equity%2C%20and%20buyouts.

Da Rin, M., & Hellman, T. (2020). *Fundamentals of entrepreneurial finance* Oxford University Press.

Da Rin, M., Hellmann, T., & Puri, M. (2013). A survey of venture capital research. *Handbook of the economics of finance* (pp. 573-648) Elsevier B.V. doi:10.1016/B978-0-44-453594-8.00008-2 Retrieved on May 8, 2022 from https://dx.doi.org/10.1016/B978-0-44-453594-8.00008-2

Dataiku. (2022). *3 enterprise trends driving AI into everyday use: 2022 and beyond* Dataiku. Report. Retrieved on May 16, 2022 from https://content.dataiku.com/2022-ai-trends

David P. Stowell. (2018). Asset management, wealth management, and research. *Investment banks, hedge funds, and private equity*(Third Edition ed., pp. 145-156) Elsevier Inc. doi:10.1016/B978-0-12-804723-1.00006-2 Retrieved on April 29, 2022 from https://dx.doi.org/10.1016/B978-0-12-804723-1.00006-2

Dayan, P. (n.d.). *Unsupervised learning*

Fairview Capital. (2018). *The future of data driven venture firms. Retrieved on May 15, 2022 from https://fairviewcapital.com/ourview/the-future-of-data-driven-venture-firms-d67/*

Foy, P. (2021). Applications of AI and machine learning in venture capital. Retrieved on May 20, 2022 from https://www.mlq.ai/ai-machine-learning-venture-capital/

Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (2006a). What you are is what you like—similarity biases in venture capitalists' evaluations of start-up teams. *Journal of Business Venturing, 21*(6), 802-826. doi:https://doi.org/10.1016/j.jbusvent.2005.07.001. Retrieved on May 12, 2022.

French, R. M. (2000). *The turing test: The first 50 years* Elsevier BV. doi:10.1016/s1364-6613(00)01453-4. Retrieved on May 15, 2022.

Fried, V. H. (2003). Venture capital and the university: The endowment's role. *The Journal of Private Equity, 6*(2), 79-85. doi:10.3905/jpe.2003.320042. Retrieved on May 12, 2022.

Geronimo, A. (2022). AI start-ups attracted the largest number of VC investors in 2021: Report. Retrieved from https://www.itp.net/business/ai-start-ups-attracted-the-largest-number-of-vc-investors-in-2021-report. Date retrieved May 21 2022

Gobble, M. M. (2019). The road to artificial general intelligence. *Research Technology Management, 62*(3), 55-59. doi:10.1080/08956308.2019.1587336. Retrieved on May 15, 2022.

Goertzel, B. (2014). Artificial general intelligence: Concept, state of the art, and future prospects. *Journal of Artificial General Intelligence, 5*(1), 1-48. doi:10.2478/jagi-2014-0001. Retrieved on May 15, 2022.

Goertzel, B., & Pennachin, C. (2005). *Artificial general intelligence* Springer.

Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics, 135*(1), 169-190. doi:10.1016/j.jfineco.2019.06.011. Retrieved on May 8, 2022.

Gompers, P., Kaplan, S. N., & Mukharlyamov, V. (2016). What do private equity firms say they do? *Journal of Financial Economics, 121*(3), 449-476. doi:10.1016/j.jfineco.2016.06.003. Retrieved on May 8, 2022.

Gompers, P., & Lerner, J. (1994). The rise and fall of venture capital. *Business and Economic History, 23*(2), 1-26.

Gompers, P., & Lerner, J. (1999). An analysis of compensation in the U.S. venture capital partnership. *Journal of Financial Economics, 52*, 3-44. Retrieved on May 8, 2022.

Gompers, P., & Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives, 15*(2), 145-168. doi:10.1257/jep.15.2.145. Retrieved on May 8, 2022

Gorman, M., & Sahlman, W. A. (1989). *What do venture capitalists do?* Elsevier BV. doi:10.1016/0883-9026(89)90014-1. Retrieved on May 12, 2022.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of Qualitative Research, 2*(163-194), 105.

Gupta, A. K., & Sapienza, H. J. (1992). Determinants of venture capital firms' preferences regarding the industry diversity and geographic scope of their investments. *Journal of Business Venturing, 7*(5), 347-362. doi:https://doi.org/10.1016/0883-9026(92)90012-G. Retrieved on May 8, 2022.

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review, 61*(4), 5-14. doi:10.1177/0008125619864925

Hall, J., & Hofer, C. W. (1993). *Venture capitalists' decision criteria in new venture evaluation* Elsevier BV. doi:10.1016/0883-9026(93)90009-t. Retrieved on May 12, 2022.

Heredia, L., Bartletta, S., Carrubba, J., Frankle, D., Mcintyre, C., Palmisani, E., . . . Sheridan. (2021). *The $100 trillion machine. Retreived on April 29, 2022 from https://www.bcg.com/it-it/publications/2021/global-asset-management-industry-report*

Hellman, T. and Puri, M. (2002) Venture Capital and the Professionalisation of Start-Up Firms: The Empirical Evidence. Journal of Finance, 57, 169-197. https://doi.org/10.1111/1540-6261.00419. Retrieved on May 8, 2022.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science (American Association for the Advancement of Science), 349*(6245), 261-266. doi:10.1126/science.aaa8685. Retrieved on May 15, 2022.

Hisrich, R. D., & Jankowicz, A. D. (1990). Intuition in venture capital decisions: An exploratory study using a new technique. *Journal of Business Venturing, 5*(1), 49-62. doi:https://doi.org/10.1016/0883-9026(90)90026-P. Retrieved on May 12, 2022.

Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly, 60*(4), 634-670. doi:10.1177/0001839215597270. Retrieved on May 8, 2022.

IBM. (2022). IBM global AI adoption index 2022. Retrieved on May 30, 2022 from https://www.ibm.com/downloads/cas/GVAGA3JP

Isaacs, A. (2022). Revolut CEO Nik Storonsky to launch new VC fund that will "disrupt" the venture world. Retrieved on May 22, 2022 from https://www.altfi.com/article/9260_revolut-ceo-nik-storonsky-to-launch-new-vc-fund-that-will-disrupt-the-venture-world

Jiang, X., Chen, Y., & Wang, L. (2019). *Can china's agricultural FDI in developing countries achieve a win-win goal?—Enlightenment from the literature* doi:10.3390/su11010041. Retrieved on June 17, 2022.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. . *Science (American Association for the Advancement of Science), 349*(6245), 255-260. doi:10.1126/science.aac4520. Retrieved on May 28, 2022

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *The Journal of Artificial Intelligence Research, 4*, 237-285. doi:10.1613/jair.301. Retrieved on May 15, 2022.

Kaplan, S. N., Lerner, J., School, H. B., , N., Loderer, C., Runde, J., . . . Advisors, H. (2010). *Spring conference overview: MCMI-III and MBMD* American Psychological Association (APA). doi:10.1037/e459592008-011. Retrieved on May 8,2022.

Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2009). Should investors bet on the jockey or the horse? evidence from the evolution of firms from early business plans to public companies. *The Journal of Finance (New York), 64*(1), 75-115. doi:10.1111/j.1540-6261.2008.01429.x. Retrieved on May 8,2022.

Kaplan, S., & Lerner, J. (2016). Venture capital data: Opportunities and challenges. *National Bureau of Economic Research Working Paper Series, 22500* Retrieved from http://www.nber.org/papers/w22500. Retrieved on May 8,2022.

Khan, S. N. (2014a). Qualitative research method: Grounded theory. *International Journal of Business and Management, 9*(11), 224-233. https://doi.org/ 10.5539/ijbm.v9n11p224. Retrieved on June 17, 2022.

Khan, S. N. (2014b). Qualitative research method-phenomenology. *Asian Social Science, 10*(21), 298. https://doi.org/10.5539/ass.v10n21p298. Retrieved on June 17, 2022.

Khanin, D., & Turel, O. (2012). Short-termism, long-termism, and regulatory focus in venture capitalists' investment decisions. *Venture Capital (London), 14*(1), 61-76. doi:10.1080/13691066.2012.666072. Retrieved on May 12, 2022.

Khurana, P., Vishav, M., & Yadav, R. (2012). Computation of user- based query using natural language processing. *International Journal of Computer Applications, 55*(11), 20-23. doi:10.5120/8800-3024. Retrieved on *May 10, 2022.*

Kraus, M., Feuerriegel, S., & Oztekin, A. (2019). *Deep learning in business analytics and operations research: Models, applications and managerial implications. https://doi.org/10.1016/j.ejor.2019.09.018. Retrieved on May 10, 2022.*

Kurzweil, R. (2000). *The age of spiritual machines: When computers exceed human intelligence* Penguin.

Lai, Y. (2019). A comparison of traditional machine learning and deep learning in image recognition. *Journal of Physics. Conference Series, 1314*(1), 12148. doi:10.1088/1742-6596/1314/1/012148. Retrieved on May 10, 2022.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature (London), 521*(7553), 436-444. doi:10.1038/nature14539. Retrieved on May 12, 2022.

Lehot, L. (2022). Q2 2022 venture capital funding data shows a steep drop, but still exceeds pre-pandemic levels. *The National Law Review* Retrieved on May 2, 2022 from https://www.natlawreview.com/article/q2-2022-venture-capital-funding-data-shows-steep-drop-still-exceeds-pre-pandemic

Liddy, E. D. (1998). Enhanced text retrieval using natural language processing. 24(4), 14-16. https://doi.org/10.1002/bult.91. Retrieved on May 12, 2022.

Litvak, K. (2009). Venture capital limited partnership agreements: Understanding compensation arrangements. *SSRN Electronic Journal, 76*(1), 161-218. doi:10.2139/ssrn.555626. Retrieved on May 8, 2022.

Macmillan, I. C., Siegel, R., & Narasimha, P. N. S. (1985). *Criteria used by venture capitalists to evaluate new venture proposals* Elsevier BV. doi:10.1016/0883-9026(85)90011-4. Retrieved on May 12, 2022.

Manyika, J., Miremadi, M., Bughin, J., George, K., Willmott, P., Dewhurst, M., & Chui, M. (2017). *A future that works: Automation, employment, and productivity* McKinsey Global Institute. Retrieved on May 20, 2022 from https://www.mckinsey.com/~/media/mckinsey/featured%20insights/Digital%20Disruption/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works-Full-report.ashx

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* Sage Publications. Retrieved on June 17, 2022 from https://books.google.it/books?id=Wt3Sn_w0JC0C

Marsland, S. (2014). *Machine learning: An algorithmic perspective, second edition* CRC Press LLC.

Martin, I. (2022,). Revolut's billionaire founder Nik Storonsky to launch AI-led venture capital fund. *Forbes* Retrieved on May 22, 2022 from https://www.forbes.com/sites/iainmartin/2022/05/17/revolut-ceo-nik-storonsky-to-launch-ai-led-venture-capital-fund/?sh=31f79ed86224

McCarthy, J., Minsky, M. L., Rochester, N., & & Shannon, C. E. (1955). A proposal for the dartmouth summer research project on artificial intelligence.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics, 5*(4), 115-133. https://doi.org/10.1007/BF02478259. Retrieved on May 6, 2022.

McIntyre, C., Zhu, K., Bartletta, S., Bhattacharya, I., Carrubba, J., Frankle, D., . . . Slack, B. (2022). *From tailwinds to turbulence global asset management 2022.* ()Boston Consulting Group. Retrieved on May 20, 2022 from https://web-assets.bcg.com/ba/c8/5b65e9d643abac4fa8e6820e86f4/bcg-global-asset-management-2022-from-tailwinds-to-turbulence-may-2022-r.pdf

Metrick, A., & Yasuda, A. (2010). The economics of private equity funds. *The Review of Financial Studies, 23*(6), 2303-2341. https://doi.org/10.1093/rfs/hhq020. Retrieved on May 8, 2022.

Minsky, M. (1986). Semantic information processing.

Mishra, S., Bag, D., & Misra, S. (2017). Venture capital investment choice: Multicriteria decision matrix. *The Journal of Private Equity, 20*(2), 52-68. doi:10.3905/jpe.2017.20.2.052. Retrieved on May 25, 2022

Norvig, P. R., & Intelligence, S. A. (2002). A modern approach. *Prentice Hall Upper Saddle River, NJ, USA: Rani, M., Nayak, R., & Vyas, OP (2015).an Ontology-Based Adaptive Personalized E-Learning System, Assisted by Software Agents on Cloud Storage. Knowledge-Based Systems, 90*, 33-48. Retrieved on May 15, 2022.

Pan, Y. (2016). Heading toward artificial intelligence 2.0. *Engineering, 2*(4), 409-413. doi:10.1016/J.ENG.2016.04.018. Retrieved on May 15, 2022.

Patel, N. (2015). 90% of startups fail: here's what you need to know about the 10%. *Forbes*. Retrieved on May 20, 2022 from https://www.forbes.com/sites/neilpatel/2015/01/16/90-of-startups-will-fail-heres-what-you-need-to-know-about-the-10/?sh=1112d5bb6679

Preqin. (2022). *2022 preqin global venture capital report.* Prequin. Retrieved on May 2, 2022 from https://www.preqin.com/insights/global-reports/2022-preqin-global-venture-capital-report

Puri, M., & Zarutskie, R. (2012). On the life cycle dynamics of venture-capital- and non-venture-capital-financed firms. *The Journal of Finance, 67*(6), 2247-2293. doi:https://doi.org/10.1111/j.1540-6261.2012.01786.x. Retrieved on May 12, 2022.

QuantumLight (2022). Retrieved on May 22, 2022 from https://quantumlightcapital.com/

Ragothaman, S., Naik, B., & Ramakrishnan, K. (2003). Predicting corporate acquisitions: An application of uncertain reasoning using rule induction. *Information Systems Frontiers, 5*(4), 401-412. doi:10.1023/B:ISFI.0000005653.53641.b3. Retrieved on June 10, 2022

Ransbotham, S., LaFountain, B., Khodabandeh, S., Kiron, D., Candelon, F., & Chu, M. (2020). Expanding AI's impact with organizational learning. *Journal of Information Technology, 27*(4), 319-320. doi:10.1057/jit.2012.27. Retrieved on May 10, 2022.

Rimol, M., & Costello, K. (2021). Gartner says tech investors will prioritize data science and artificial intelligence above "Gut feel" for investment decisions by 2025. Retrieved on May 10, 2022 from https://www.gartner.com/en/newsroom/press-releases/2021-03-10-gartner-says-tech-investors-will-prioritize-data-science-and-artificial-intelligence-above-gut-feel-for-investment-decisions-by-20250

Rimol, M. (2021). Gartner identifies four trends driving near-term artificial intelligence innovation. *Gartner* Retrieved on May 22, 2022 from https://www.gartner.com/en/newsroom/press-releases/2021-09-07-gartner-identifies-four-trends-driving-near-term-artificial-intelligence-innovation

Rogers, E. M. (1962). *Diffusion of innovations* New York: Free Press of Glencoe.

Ross, G., Das, S., Sciro, D., & Raza, H. (2021). CapitalVX: A machine learning model for startup selection and exit prediction. *The Journal of Finance and Data Science, 7*, 94-114. doi:10.1016/j.jfds.2021.04.001. Retrieved on: May 25, 2022

Sahlman, W. A. (1990). *The structure and governance of venture-capital organizations* Elsevier BV. doi:10.1016/0304-405x(90)90065-8. Retrieved on May 8, 2022

Schmidt, C. M. (2018). The Impact of Artificial Intelligence on Decision- Making in Venture Capital Firms. Universidade Católica Prtuguesa. Thesis.

Shepherd, D. A., & Zacharakis, A. (2003). A new venture's cognitive legitimacy: An assessment by customers. *Null, 41*(2), 148-167. doi:10.1111/1540-627X.00073. Retrieved on May 12, 2022.

Shepherd, D. A., Zacharakis, A., & Baron, R. A. (2003). VCs' decision processes: Evidence suggesting more experience may not always be better. *Journal of Business Venturing, 18*(3), 381-401. doi:https://doi.org/10.1016/S0883-9026(02)00099-X. Retrieved on May 12, 2022.

Sheth, A., Akhtar, U., & Halloran, B. (2022). *Chasing disruption: The brave new world of growth investing.* (). Retrieved on April 29, 2022 from https://www.bain.com/globalassets/noindex/2022/bain_report_global-private-equity-report-2022.pdf

Silver, D. A. (1985). *Venture capital: The complete guide for investors* John Wiley & Sons, New York.

Singh, A., Thakur, N., & Sharma, A. (2016). *A review of supervised machine learning algorithms* IEEE. Retrieved on May 18, 2022 from https://ieeexplore.ieee.org/abstract/document/7724478

Sofaer, S. (1999). Qualitative methods: What are they and why use them? *Health Services Research, 34*(5 Pt 2), 1101-1118. Retrieved on June 17, 2022.

Sørensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *The Journal of Finance (New York), 62*(6), 2725-2762. doi:10.1111/j.1540-6261.2007.01291.x. Retrieved on May 12, 2022.

Statista Research Department. (2022a). Global total corporate artificial intelligence (AI) investment from 2015 to 2021. Retrieved on August 7, 2022 from https://www.statista.com/statistics/941137/ai-investment-and-funding-worldwide/

Statista Research Department. (2022b). *Value of VC funding worldwide 2020, by region. Statista. Retrieved on May 2, 2022 from https://www.statista.com/statistics/1095957/global-venture-capita-funding-value-by-region/*

Stebbins, R. A. (2001). *Exploratory research in the social sciences* Sage.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques. 2nd Ed.

Strömberg, P., & Kaplan, S. N. (2001). *Venture capitalists as principals: Contracting, screening, and monitoring.* (). Cambridge, Mass: National Bureau of Economic Research. doi:10.3386/w8202 Retrieved on May 22, 2022 from http://www.nber.org/papers/w8202. Retrieved on May 8,2022.

Taddy, M. (2018). The technological elements of artificial intelligence. *The economics of artificial intelligence: An agenda* (pp. 61-87) University of Chicago Press. Retrieved on May 15, 2022.

Teare, G. (2022). Global venture funding and unicorn creation in 2021 shattered all records. *Crunchbase News,* Retrieved on June 17, 2022 from https://news.crunchbase.com/business/global-vc-funding-unicorns-2021-monthly-recap/.

Torres, J. P. (2020). The venture capital feedback cycle: A critical review and future directions. *Entrepreneurship Research Journal, 12*(1), 1-34. doi:10.1515/erj-2017-0205. Retrieved on May 8, 2022.

Tricot, R. (2021). *Venture capital investments in artificial intelligence. OECD Digital Economy Papers, 319, 1-46-* doi:https://doi.org/https://doi.org/10.1787/f97beae7-en. Retrieved on May 8, 2022.

Trocha, B. (2019). Data-Driven VCs: How 83 venture capital firms use data, AI & proprietary software to drive alpha returns. *Medium. Retrieved on April 3, 2022 from https://medium.com/hackernoon/winning-by-eating-their-own-dogs-food-83-venture-capital-firms-using-data-ai-proprietary-da92b81b85ef*

Turing, A. M. (2009). Computing machinery and intelligence. *Parsing the turing test* (pp. 23-65) Springer.

Tyebjee, T. T., & Bruno, A. V. (1984). *A model of venture capitalist investment activity** Routledge. doi:10.4324/9781315235110-6. Retrieved on May 12, 2022.

Vorholes, W. (2016, February 23,). Artificial general intelligence – the holy grail of AI. Retrieved on May 12, 2022 from https://www.datasciencecentral.com/artificial-general-intelligence-the-holy-grail-of-ai/

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience, 2018*, 7068349-13. doi:10.1155/2018/7068349. Retrieved on May 12, 2022.

Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence, 10*(2), 1-37. doi:10.2478/jagi-2019-0002. Retrieved on May 15, 2022.

Wei, C., Jiang, Y., & Yang, C. (2009). Patent analysis for supporting merger and acquisition (M&A) prediction: A data mining approach. (pp. 187) doi:10.1007/978-3-642-01256-3_16 Retrieved on June 10, 2022 from https://ui.adsabs.harvard.edu/abs/2009debs.book..187W

Wells, W. A. (1974). *Venture capital decision making.*

Wright, M., & Robbie, K. (1998). Venture capital and private equity: A review and synthesis. *Journal of Business Finance & Accounting, 25*(5-6), 521-570. doi:10.1111/1468-5957.00201. Retrieved on May 12, 2022.

Xiang, G., Zheng, Z., Wen, M., Hong, J., Rose, C., & Liu, C. (2012). *A supervised approach to predict company acquisition with factual and topic features using profiles and news articles on TechCrunch. Retrieved on June 10, 2022 from http://www.cs.cmu.edu/~guangx/papers/icwsm12-long.pdf*

Yankov, K. (2018). Phase-plane models in korelia software. Paper presented at the 1-4. doi:10.1109/InfoTech.2018.8510747 Retrieved on June 10, 2022 from https://ieeexplore.ieee.org/document/8510747

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018, Recent trends in deep learning based natural language processing [review article].*IEEE Computational Intelligence Magazine, 13*, 55-75. doi:10.1109/MCI.2018.2840738 Retrieved from https://ieeexplore.ieee.org/document/8416973

Zacharakis, A. L., & Meyer, G. D. (2000). *The potential of actuarial decision models* Elsevier BV. doi:10.1016/s0883-9026(98)00016-0. Retrieved on May 12, 2022.

Zhang, D., Masle, N.., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika J., . . ., Perrault, R. (2022). *Artificial intelligence*
index report 2022. ().AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University. Retrieved on May 13, 2022 from https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf.

Appendices

Appendix 1- Interviews Guide

Thematic Bucket		Question
General Questions	1	What is your role in your VC?
	2	Could you provide general information on your VC firm? For instance, what is the fund strategy? What are your investment stage, geography, and industry focus?
	3	Would you define your VC firm as a data-driven VC? In which sense?
AI-related questions	4	Do you specifically use AI technology in any of your processes?
	5	In which venture stage do you use AI? (Deal sourcing, deal screening, due diligence, negotiating, closing, post-deal)
	6	Which type of AI technology are you using? Is it supervised or unsupervised? Could you provide specific information about the development of AI models?
	7	Is it a proprietary technology? Do you license it to other firms in the industry?
	8	How long did it take to develop the AI?
	9	How does it work? Which output does it produce? Who uses it? Who updates it and how often?
	10	Have you hired any data scientists (or similar roles) in your company?
Firm-level adoption decision questions	11	Did your VC start as a data-driven firm, or did you recently adopt new technologies?
	12	Have you or any of your colleagues experienced resistance to implementing new practices? In which sense?
Personal Opinion questions	13	Do you think that AI brings a valuable contribution to your VC? In which sense?
	14	Do you see the possibility of AI disrupting the industry and becoming a common practice in VC? If so, with which timeframe?
	15	Do you think that AI usage can lead to having a competitive advantage over competitors in the industry?
	16	Which are the limitations of AI application in VC, in your opinion?

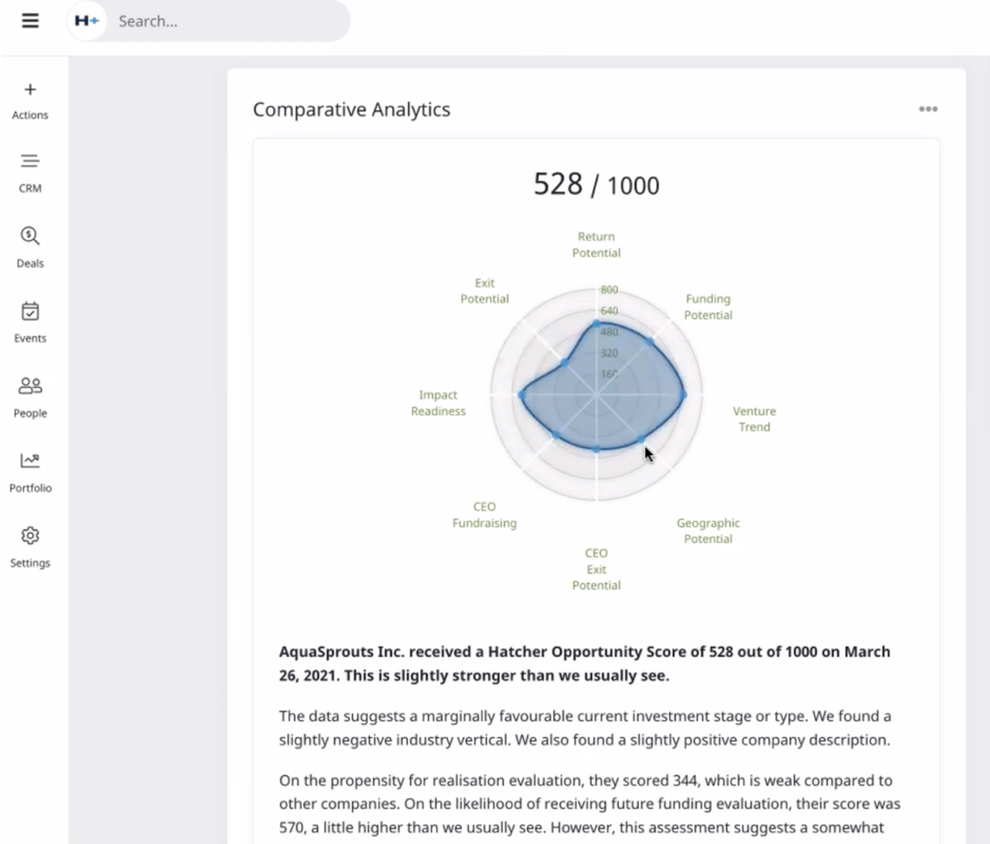
Appendix 2 – Comprehensive VC Database

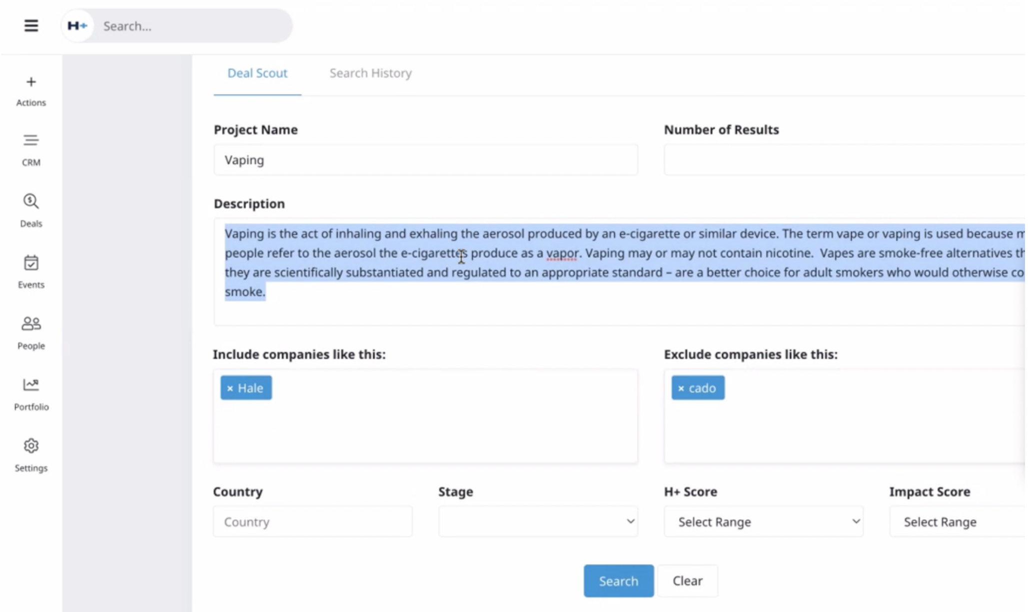
An online interactive version of the database presented In *Table 3* is available through this QR code[10]:

The online resource allows future researchers to leverage the database created, facilitating eventual data manipulation and re-elaboration.

[10] Link of the online resource: https://airtable.com/shrZgGuVb2WaH93VE/tblSsosY8ps5raSlZ

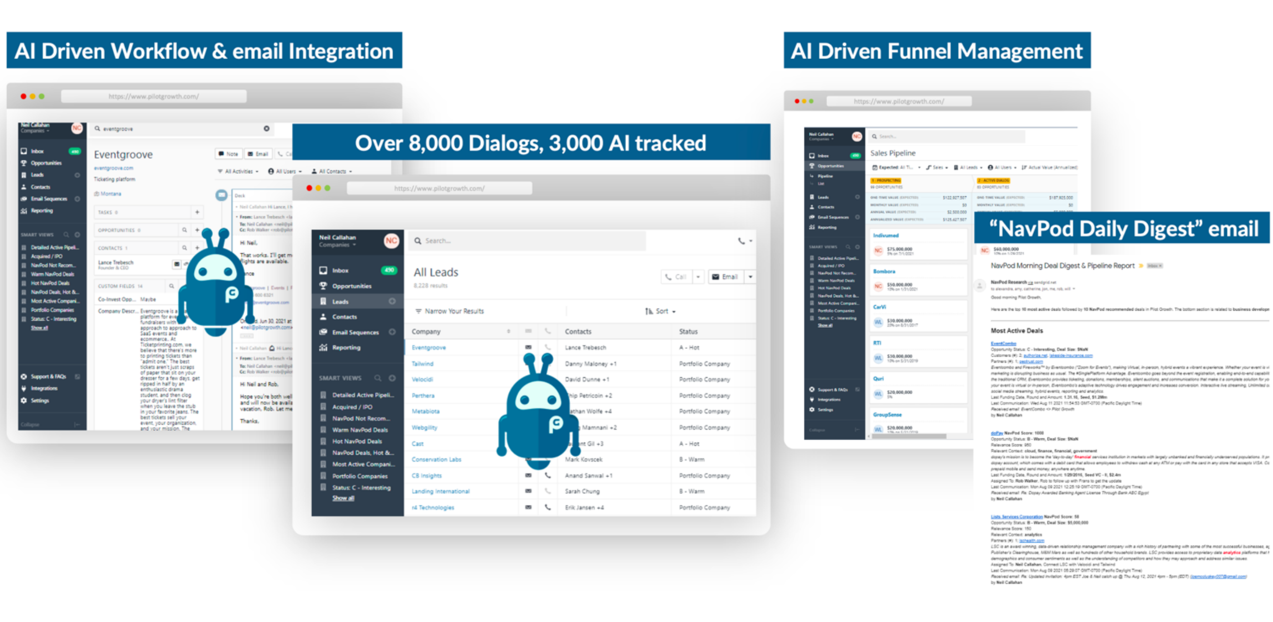
Appendix 3 – Hatcher+ VAAST Platform Interface - Opportunity Score View and Deal Scout Feature





Source: Hatcher+

Appendix 4 – Pilot Growth Equity NavPod Interface and Recommendation Email



Source: Pilot Growth Equity

Appendix 5 – Interviews Transcriptions

Interview 1 – Redstone

Interviewer 0:05: I found Redstone VC online and I decided to interview you because you define yourself and your vision as to combine your extensive experience in VC with data analytics and you also claim to have a proprietary analysis platform that enables you to predict industry trends, benchmark the segments and identify appropriate investment targets. To what extent do you think data is important in your VC and do you specifically use AI as a technology in any of your internal operations?
Interviewee 1 0:40: So first of all, let me thank you again for having me here, it's my pleasure to answer these questions. Let's start with a very brief introduction on what Redstone VC is, because it's very much related to how and why we use data. Redstone was born already 8 years ago. We were born as VC as a service company, which basically means if you want to simplify, a sort of consulting firm specialized in the VC industry, which means that our main clients were large corporations that were approaching the VC ecosystem, as a whole. And what we did was market intelligence services, and due diligence services. We helped corporates to establish their investment thesis and to find targets, to support them in their investments to establish their own funds, and ultimately, manage those funds. And then, over time, we've completely kind of twisted our business model to fund this fund manager. We have set up several funds under management with different configurations. It's kind of interesting, we can also dig deeper but let's say on the surface for now. And so, we have a strong heritage of relationship with corporates. Market intelligence, analysis of the space is still part of our DNA. And therefore, we try to establish a very data driven approach, which means that in the first place, we started just collecting data from several different sources. And we build up our proprietary database, which is called Phoenix. And whit this proprietary database what we do is we basically collect data from several data sources, structured and unstructured databases. We partially dig some data from companies' websites, we standardize everything and then we generate the database. We have a huge database of data of startups with more than 2.2 million companies as of today. We have also all the information related to transactions, to people, to funding rounds, to whatever you may think would be interesting related to this stuff. On top of that, what we have been doing over the last year now, we start to develop a series of tools that basically help us in making our investments, operations, more efficient. And this collection of tools that help us in answering critical questions, we call this whole system Sofia. You may ask, why Sofia? It's like a fun story. I don't know if you are aware about Greek mythology, but it's the name of the Goddess of knowledge and stuff like this, so that's the reason why. And among these tools, we also partially use artificial intelligence algorithms, specifically, we mostly use as of today natural language processing. And then we leverage these algorithms for competitive search, for startup searching. Just to make an example: we have a sort of Google for startups, let's say that. You can type a sentence, and based on the words that you write, these algorithms look for companies, and similar. So, you can imagine one line of text as a sort of multi-dimensional vector, and then we check all the vectors that are close to this one and we get the company with that are similar. And we do something similar for competitors so you can type a couple of companies that are similar into the search bar, and you get all the competitors. So, we use actually artificial intelligence. We also have another tool when we have a generate lists which is a recurring task for our company. And it's probably very specific. We generate a list of companies to map a specific space, to understand what's going on in this space. If we have to do any project with corporates they are often interested in the market, which are the main players, how many companies and there are, how much funding there is. So, we have to come up with a list of companies that are kind of consistent with the mandate that they gave us and this process was super manual at the beginning, we kind of partially automated it and now we have tool that is generating an automated list based on the inputs provided. So, you provide a sort of input file. You can imagine it as an Excel file for the sake of simplicity. Where you type the definition of the different things you will explore. Then you put it in this tool, and it will ultimately automatically generate the list of companies. Then we have other tools where we don't use AI per se, but we have a very driven approach. In the sense that based on the data we have, for instance, we generate a lot of analytics, a lot of dashboards, data visualization tools that we use to map the market where we can see for instance, which are the trends of new companies founded in a specific sector, which are the trends of funding. Just to give you an example, one of these dashboards that we generated, we tried to map the transactions worldwide to see if there was any difference and something very interesting that from 2020 to 2021, you see that the average transaction amount across all funding rounds, has doubled compared to the previous year. This is a signal of more capital being deployed in the market on one side but also that valuations are crazy which you may have heard about it. This is certainly something relevant. Then all this data we have allows us to put companies in rankings. We rank companies in terms of funding, in terms team, so we have a ranking to make it simple, of course, where you can see how strong the team is. And we give these kinds of scores to the team based on the previous funding experience, exit experience, we look at if founders have worked for top companies, if they have studied in top universities. And we look at the industry experience, so it's quite complex, but it is something that allows us to be super data driven. Of course, is not a perfect system but it's still a very good way in which we can narrow down the scope and focusing on the top

companies. Another thing we did very recently, we developed a new type of ranking for growth companies. We tried to see: can we assess how our fast are growing these companies? So, what we did, we developed a new indicator that tells us how fast growing these companies are. And we were looking at several metrics such as web traffic, employee growth, funding days, age of the company, number of job vacancies, so we kind of aggregate the data and we come up with an indicator. We were able to compare them and say okay, these 15 - 20 companies are the ones that are showing stellar growth. Of course, there were stupid names like gorillas, okay, thank you! Everybody is aware about that. But also, you can spot other companies. I can keep going so stop me if you have any question.
Interviewer 9:00: On what you said in the first part of your answer, it's clear that you use a lot of data in a lot of different ways. But, if I understood well, you use AI just with natural language processing. So, with your sort of Google platform to search for startups, and to create lists of companies. Can you elaborate on this?
Interviewee 1 9:28: We use AI for competitor search and startup search. These are the two main use cases. So, you have a company, and you want to see who the competitors are besides the well-know. You put the company in the system, and you get the list of competitors. This is one use case. The second one is "I want to see the space, let's say, of cucumbers delivery" as an example. You type grocery delivery and whatever and you get the list of companies. So, these are the two main use cases. Then a third use case, which is related to the list generation is also a partially semi-automated taxonomy. You have millions of companies as I said, you cannot type them manually. We actually leverage all the information we collect from our data providers, on top of that, through the use of AI we partially automate the taxonomy values, which means I want to see okay, which are the main tags assigned to a given company, and we do it in a partially automated way. When I say partially. It's because, of course, AI as you may know, needs a sort of training. What we do is we have the market we think how we could structure it, and based on the taxonomy values we have the system learn it and assign these taxonomy values automatically to the company. So, this is another use case.
Interviewer 11:20: So, you are doing the learning and training phase.
Interviewee 1 11:24: Exactly, exactly. In this case, yes. I mean, now we have just completed the technical project to restructure the system. If we speak again, at the end of the year just to give you an understanding of what's in our current roadmap. We have this project that is called Hermes, so we are biased towards Greek mythology. This project automatically tracks companies and automatically gives me updates on the company in terms of rounds raised, new technology developed. This is still under development but is probably priority 1 in our roadmap at the moment.
Interviewer 12:46: Basically, you're currently using AI just for deal sourcing, correct?
Interviewee 1 12:49: Deal sourcing, deal screening and partially competitive analysis. Internally people want to pass to automation of investment memos and skeptic about it. What we can do is to partially automate thesis of companies since we are super structured in writing one pagers, and we also did some experiments, partially automating some parts of the memos. For instance, for an investment memo we scraped websites of competitors to actually see which are the features of the competitor, so we get competitive benchmarking, and we did it in an automated way. So rather than me going on the website of each company, we wrote a code that just went to the website and generated as an output a table with features as column cells. Identifying if a company had that feature. But at the moment those are the main use cases.
Interviewer 14:03: Is this more like sort of coding exercise rather than AI?
Interviewee 1 14:10: Absolutely.
Interviewer 14:12: And so basically, you are a VC-as-a service right? It's not super clear from your website, and I want you to understand if you raise funds yourself and you use all of these platforms and tools yourself, or if you put them at work for other corporates or other companies?
Interviewee 1 14:33: That's a very good question. Our website is very bad, and we are in the process of changing it. You have two things. First of all, our tools are only accessible to the investment, so only to us. As external clients do not have any access unless we grant the access. But usually, the access that they have refers to the data visualization part, which is basically the output of our work. So, it's not the intention tool per se. This the first thing. And then in terms of what we are. We were born as VCaaS. Now, we define ourselves as multi-VC fund asset manager. In this aspect we manage at the moment 10 sector funds. We call them sector funds, because they have a very specific focus in terms of verticals. We are not the usual one fund of 400 million that invest in European seed stage companies. We have 1 fund of X million that invest in European FinTech early-stage companies, another one that invest in retail companies, other that invest in other companies, and so on. We are now closing two more funds, on in ESG and one in impact VC. As you can see, they're all focused on one specific industry. And this gives us an edge on our competitors because we have a very vertical expertise. We have our expertise in terms of the sector, so we know what's happening in the sector and we can proactively help them. On top of that, we still partially work with corporates, I say partially because it's not our core business. These projects create insights that we can use in pour own research. At the same time all these relationships that we have with corporates are beneficial for our companies because our portfolio companies have very large access to professionals from these companies, they can access to the network, they could be like commercial partners, or potential exit channels. It's a very good thing for our portfolio companies. Nowadays our core business is fund business.
Interviewer 17:31: You are an early-stage VC, correct?
Interviewee 1 17:35: Yes.

Interviewer 17:36: Because you know that AI works on data. And usually with the early-stage VCs, companies are basically new, and most of the time, there's not a lot of data available online or through other databases. So how do you make AI work with the limited information available? Or do you not think that this is not a problem? What's your take on this?
Interviewee 1 18:05: I mean, it depends on what you are focusing on. In the sense, if you're talking about stealth startups, of course it's kind of impossible to spot them. You can do through automating sales navigator on LinkedIn. but this will not be AI, it will be more like a coding exercise. We do that actually, but it's not AI. For sure, first of all, one comment about data, it is a scarce resource. That's why we have so much interest in it.. Because we have a kind of data heritage and asset that few have in such an extensive way, maybe later I can also show you something. But you will see that the data points we have very kind of specific. So, this is one first thing. But data availability is an issue because we are not talking about public companies where everything is disclosed, and everything is available online. This is one thing, but on the other side, yes, definitely, you touched many important points. I think it depends on the use cases because what you can do is you can for instance see which are the main traits, not like physical or ethical traits, but which kind of experience you have as a founder that makes more likely that you're going to be successful. You can use some like predictive algorithms, so AI, and then based on that, you can go back to your data and look for people with similar characteristics, or background or experience. This is something that could be already put in practice. But besides that, I don't have in mind a specific case. Especially at this stage, if you join this industry you will understand, data is a very powerful tool to spot early on opportunities but then you need to speak with people. You have to understand what is actually your vision, you have to be convince, talk about execution. But for sure data availability is an issue.
Interviewer 20:20: That's what I'm getting for most of the interviews that I'm conducing. AI and data is useful to start the process and get in touch with people that might be of value and they be a good investment. And then you still need to value the personal side and the gut feeling in a later state.
Interviewee 1 20:45: For sure. The useful thing in the sourcing phase, is that instead of 1000 companies you look to only a few because you can narrow down the scope and become more operationally efficient. Imagine going though 1000 companies instead of 50. You will save a lot of time, you can be much faster, you can be the first one, you can be there before the others. Talking about operational efficiency, like what I was saying before, you can automate how you do the commendations, how you do search, how you do competitor review, and you can actually use AI or coding.
Interviewer 21:42: Considering that using AI in VC is becoming the sort of trend for the future, ideally, and more and more VCs might start implementing these practices, do you think that it could become an advantage? As of now, if a VC uses AI for your sourcing, for example, it might be a sort of competitive advantage because you track companies faster you get there first, you might become the lead investor or something like this. But what if every VC starts using AI for deal sourcing? Where do you think is the differentiation at that point? Is it in the database that you have in any other aspect?
Interviewee 1 22:19: This is a very interesting question. First of all, I think that there is still a lot of skepticism in the system. People think that this is cool, but that VC is still a people business, that it is about network, which is of course true, but to some extent. In the sense that you have to merge the two things. So, if you are able to do that and at the moment, I think we are one of the few companies that are doing that, you could really have an advantage. For instance, if you look at EQT Ventures, I'm sure you saw them. My CTO of our company says, they kind of shut down Motherbrain. Not really shut down, but it's more like now an advertising thing.
Interviewer 23:15: I'm also trying to understand if VCs are actually applying AI to the extent that they are advertising or if it's just marketing.
Interviewee 1 23:44: That's exactly the point. I think there is still a lot room of improvement also from our side. We use it, but it's not perfect, it's not on scale in terms of operations. On sourcing, it's on scale, I can also show you the platform later, it works. But within the company you still have a lot of room for growth, right? Other companies that use data, just to give you an idea. You know Balderton Capital? You should speak with Francesco.
Interviewer 24:25: I already spoke with Francesco.
Interviewee 1 24:24: You know that he's alone, and just buying tools and aggregating everything. They haven't developed any proprietary AI even though they are data driven. I love the research that they are doing, it is super interesting. But if you look at the technology, he is collecting a lot of tools and data sources. He created a tech stack that just allows them to be data driven and tp collect inputs and signals from several parties and see if a company has a spike on GitHub and this means that they are a very good company, but it's not AI, right? It's just a very refined way of tracking signals, which is still, in my opinion, very useful. Because let's say that you have patents, GitHub, LinkedIn and whatever, you see okay, a simplistic example, if I start writing stealth mode founder on my LinkedIn profile, and at the same time my GitHub profile explodes, it means that you should talk to me. It's not AI. If you have this technology now, you have a competitive advantage. In 10 years everyone has it is more a matter of how much data you have, how much in depth you can go. You don't only need the names of the companies but all the other information like who working there, what they did before joining the company, how much funding they raised, who invested, how strong is the investor. So, you have to be the best

in terms of how you collect and structure your data. Given that AI is not a commodity yet, we don't have to think about this second step.
Interviewer 26:52: And since you are one of the first to have AI and this technology on learning overtimes and getting better and better and better over the years, you will have a better performing AI than your competitors in the future, I guess. Yeah. I found a lot of companies that kind be defined as data driven. But then it's difficult to understand if they actually use AI or if they don't use AI. And also, the ones that claim on their website, such as EQT ventures that they use AI, they don't disclose how and where they use AI? So overall might be a little bit of a marketing stunt. And it's really difficult because maybe they use AI just for one operation and claim to be fully AI driven if it's just one part of a bigger picture.
Interviewee 1 27:44: That's definitely true but you have to start from somewhere, right? I think it's a good start. And so far, it worked. In the sense that we already invested in three companies over the last 12 months through our platform Sofia. We found three companies and we actually invested in them; one these is actually Italian so good news for us. But still, I agree with you. I think it is still a lot of marketing. But for example, we are not really marketing it even though we have it. I'm personally pushing a lot on this, in the sense that we should advertise much more what we are currently doing because it's something that other companies are doing. If you go to EQT ventures website, they have a part dedicated to Motherbrain and it is super cool. The website is futuristic.
Interviewer 29:00: It's difficult because also if it's just marketing and you get an interview with someone and ask them about their practices, they can still lie to support their AI case. So, there's always this problem.
Interviewer 29:18: Just a few last questions. How long did it take to develop this AI platform? If you know.
Interviewee 1 29:35: So, when we developed the sourcing thing, without counting me because I brought more the business knowledge to try to make sense of everything, they were like 3 full time people, and it took something like 6 months to develop the first version. So there was a beta after six months and then there's been incrementally improvement. The struggle was more before. So, on the data collection, data standardization part. That took a lot of time.
Interviewer 30:32: How many data points or cases did you use to train the AI?
Interviewee 1 30:40: Good question, I'm not sure. In this case, with NLP it's different. You take a company, and you transform the company based on all its taxonomy values into a multi-dimensional vector. Then, the algorithm does the same for all the companies and the closest ones or the ones that are taken as the competitors. I don't know if you got what I mean.
Interviewer 31:21: Ok let's say, you take the company, separate its features it into a multi-dimensional vector and it does this with all the companies. And then it takes the little pieces of the vectors and then it compares them within the companies.
Interviewee 1 31:34: Exactly. So imagine it like a sort of galaxy where we have all the stars. And all the stars closer to each other and similar, right? And then I don't want to go too in depth because it's not even necessary but just to give you an idea. We use KNN and average nearest neighbors were the two models we were using. So, this was this was on natural language part. While on the semi-automated list, since I think we did two weeks ago I don't know to be honest. I mean I can ask.
Interviewer 32:35: I'm also trying to understand which type of AI are VCs using because to understand if they are actually using AI and it is not just a marketing stunt. Is it neural networks?
Interviewee 1 32:58: Exactly. I mean, you can use predictive, but the thing of predictive and this is also maybe it's skeptic as well, in a sense. Predictive is very difficult. I mean, what can really use to predict the success of the company? Okay, you can do this exercise, but maybe you find that a company funded in Silicon Valley has a high likelihood of being a unicorn than a company in Italy. But this can be because in Italy we have only one unicorn, so data is already biased. This is very difficult to do, predict like this the successfulness of the companies. You can also find some papers, but they are not good quality.
Interviewer 34:06: One last question. It's been really, really useful. Do you have a data scientist, like a full-time data scientist in your VC?
Interviewee 1 34:27: We have three. Let's say that what we have a full stack developer, and data scientist. Then we have one which is software engineer and data scientist. And then we have another one which is an NLP expert. Some just focused on AI and machine learning. He just joined the team to be honest. We have a lot of expectations because we hired him expensively for this reason for kind of augmenting our AI capabilities.
Interviewer 35:11: Okay, yeah, I think it's been really useful. Thanks a lot.
Interviewee 1 35:21: Thanks.

Interview 2 - InReachVentures

Interviewer 0:03: I'm going to start with the first question. Basically, as I told you, I found InReach Ventures through a bottom up research online trying to find out who are the VCs that are using AI right now. And on your website, you say that you are at the AI powered venture capital firm, and that you use a proprietary software to discover, evaluate and

support investments in the most promising European startups. Can you elaborate a bit on this and in particular, on the role that data and AI have in InReach Ventures?
Interviewee 2 0:41: Sure. First of all, let me say, the current focus of InReach Ventures, we focus on early stage, European companies. And this determines the data we collect, the data we use and how we what we use it for. Because we're focused on early stage, the main role of data and AI is discovery. It's finding potential investment opportunities. We collect the data from multiple sources, more than 300 sources. We look at the website we look at what information is provided by each company on their website and then there is a layer of intelligence that make sense of this data. And that's a first screening of what could be interesting for our investment team. At the moment, is about 5-6 people. And there are as many software engineers in the firm, which is a very unusual balance. The whole point is that by using technology developed by this engineering team, finding investors can be much more productive than a whole team of 10 investors. So, it's removing the kind of human inefficiency in going through hundreds and hundreds of companies every single day. It also makes it for more pleasant for the investor themselves, they just work with something that is prequalified, already checked by AI.
Interviewer 2:47: So, from what you said, the AI collects data from more than 300 sources, mainly four sourcing purposes. But how does the AI actually work? How is it structured? How did you develop it? Is it supervised or unsupervised learning? Or does it use deep neural networks?
Interviewee 2 3:12: There are a lot of components because the evaluation of the company has different dimensions. I won't reveal all the secrets of course. Our approach is mainly supervised in the sense that we try to give to our investor what they consider interesting based on our investment strategies, as well as individuals preferences. Different people are focused on different areas. Or are experts in slightly different areas. Yeah, so mainly supervised. And then what components there are. We look at the people dimension. We use a lot the text data. There are a lot of NLP components for different purposes. I can tell that it's two layers. The first layer of using AI to make sense of data. And then there is another layer which is to make a final judgment on the company. There are components that classify the company into different business styles. We are interested in SAS consumer internet and marketplaces; this is our investment focus. The first the layer does that classification. And then we combine this type of information with many other information about the people about the location of company about how many downloads they on their app if they have it, the funding configuration, if they got funding recently, how much funding did they get. All this information goes into a ranking system. So, the layers are kind of a chain.
Interviewer 4:56: And the final output is a score that ranks the companies correct?
Interviewee 2 5:37: Essentially yes.
Interviewer 5:37: And how many companies do you consider investing in after the score is given?
Interviewee 2 5:53: From the point of view of capacity in our team, we look at least, let me do some quick maths... at least 200 companies a week. And then he's a bit of a funnel that shrinks, some of them we try to reach out, some of them after the first call are not considered as interesting, some of them go into later stages with us of engaging to the point of offering term sheets.
Interviewer 6:30: So, you use AI just for deal sourcing and deal screening and no further in the VC process, correct? So, no AI in negotiation, due diligence or anything else?
Interviewee 2 6:44: No, at this stage, for early-stage companies, those kinds of analysis it's okay to do them with the investment team. Also, because the volume of deals has shrunk by that point, it has shrunk so much that it is manageable for the number investments.
Interviewer 7:04: In relation to deal sourcing, how do you search for companies? Do you have a platform on which you do your search?
Interviewee 2 7:12: We developed our own internal platform. It is a web and mobile app that our investors use to explore, look at the ranking and manage the workflow as well, with the stages of first call, second call and so on.
Interviewer 7:38: Okay, and how long did it take to build and develop this platform, if you know?
Interviewee 2 7:44: Well, I've been in the company for a year and a half, but they started in 2016 and that's been a continuous improvement. There is never the point where all of this is done.
Interviewer 7:59: I understand. In this platform, do you separate deal sourcing and screening with AI from normal processes? Like if I'm a startup and I want to apply for funding can I still send my application?
Interviewee 2 8:19: We also have an external facing component to submit funding application. This goes into the same deal funnel.
Interviewer 8:30: And what are the AI application specifically that you would say you use more for deal sourcing and deal screening? I mean, you mentioned that you use NLP. How does it work? Do you have a database with a lot of companies, and you type something and NLP searches on your database? Or how does this work?
Interviewee 2 9:00: We do have a component of searching. I wouldn't say it's really powered by NLP. It's some fairly simple text search, which nowadays you don't need a google-like system or a very complicated language model to achieve some form of search. It's more like keywords basically. So, NLP, we use it more to understand something deeper about

the company, for instance is it a marketplace or is it a SAS company or also classifying it into industries and technologies. What are the technologies what that are the industries of these companies.

Interviewer 10:12: How is this different from simple coding? What's the value added that your AI brings in rather than just using coding in deal screen and deal sourcing?

Interviewee 2 10:24: Well, it is tailored to our investment team, to our investors and to the things we are interested on. Maybe your question is: why AI and not a simple heuristic? Like an if else logic? A simple if else logic becomes pretty complex. And also in real life, you can ask the investor: what is interesting to you? Describe it to me in an unambiguous way so that I can code and that kind of exercise will be very long, probably not get to results that are as good as let the investor say: this is interesting or not interesting. And having a machine that catches the aspect that actually important to them.

Interviewer 11:44: Before you said that it was mostly supervised, but wouldn't this be more on the unsupervised side?

Interviewee 2 11:53: No. I stand by my description. It is supervised. We show the investors a bunch of companies every week. And they say: “I want to talk with these, I don't want to talk with these, this is interesting this is not interesting”. That training data we use to understand which are the characteristics or the features of a company that our investors are interested in. And this is changing over time. People change, their judgment and their interest changes. Also by looking at the company and making a decision in the company, you kind of get the real judgment rather than having a long abstract conversation, you say: “what do you think is interesting? What do you think you consider interesting?” And then you actually look at the data and what they may tell you in words doesn't reflect the way they judge. Sometimes the judgment is also imperfect, because it's human. And the machine also helps them to look at themselves in the mirror and say: “well I used this judgment for this company, why am I not using this judgment for these other companies?”. It helps to reflect on all the decision you've done in the past and make even better decisions.

Interviewer 13:29: Before you talked about running an early-stage VC. In early-stage VC, there's not a lot of data available on the companies because they are basically new companies. There's not a lot of data points that you can actually gather. So, how do you deal with this? Which are the main data points that feed into your AI to give out a score in the end? Like, which proxies do you use to understand the success of companies?

Interviewee 2 13:59: We try to make it easier for us. The goal of our current AI is of course in changing the future. The goal is not really to say the company will be successful or the company is a good investment for us, but it's more to identify good leads, potentially good leads. So, something that an investor say: “I want to know more”. And then they decide if it's a good company or a good investment. Because as you say, lots of information is not on the web. The simplest one you can think of is the cap table. If we get to a company that is very interesting, maybe solving an interesting problem, and the 20 founders, each has a little share, they give up a lot of shares the company to the first seed investment, so the company structure is a bit messy. This automatically becomes not interesting for us. But that information is not available, we discovered it later. The goal of the AI is to find the potential leads and then there is a deep dive that can be easily done by the investing team. And if you are talking about later stage, where there are there are models of evaluation, you can even estimate the valuation of the companies. It can be done by some rule-based models. But we are not at the investment stage that requires this kind of processes.

Interviewer 15:49: So which kind of benchmarks do you use to give you the lead? I mean, do you look at market kind of metrics? Do you look at competitors?

Interviewee 2 16:02: Yes, we look at competitors, we look at the market. But this is already that level of deep dive and investor judgment that is done by the investment team. That big advantage of AI for an early-stage firm like us, is going from one million companies to a short list of several thousands. And then our investment team can judge, can apply judgment, and also engage. Because investing is not just throwing money, but also creating relationships helping them understand if we can add the value. So, there's a human component. I think if you remove it, you lose.

Interviewer 16:54: I agree with you. It's a people business. I understand the usage of AI that you're doing, but I'm trying to understand how you go from having a very large list of companies to have less companies, and which are the kind of rules that leads you this reduction? Like: you go to the investor team, and you ask them: tell me which are the companies that you're interested in within this bucket. And then they tell you the companies. And this feeds the AI algorithm, but which are the metrics that you track in the companies to get to the list in the first place?

Interviewee 2 17:44: So, every week our investors make several hundreds of decisions. Those decisions, it's their job. When they say: “I want to talk to this company”, then that company goes into our process. The investor just operates and then at the machine, the AI, every week basically, looks at all the decisions that have been made in the past, and we’re talking of several thousands of decision, and say: given what we liked, what our investment liked in the past and even the characteristics of those companies, which are the best companies I should recommend them next week? And I mentioned all the features we look at are: people, information about the company, the space. We don't invest in robotics, we don't invest in hardware, we don't invest in medical things. So that goes automatically at the bottom of the ranking. And then in certain periods of time, there are some spikes of interest in some subjects and topics. So, the following we might find an increase in recommendations in that same topic.

Interviewer 19:57: I understand. But if the investment team just focuses on an investment topic or a sector for one week, then wouldn't it feed in the AI the same sector for the next week? And so on, like in a virtuous circle? Is the only input the decisions from the investment team or do you get inputs from the market, from trends, from new technologies? How do you break this cycle?
Interviewee 2 20:30: We try not to break this cycle. That was just an example of how things change over time. I'm not saying that if in a week someone does a couple of more calls about sustainability, next week, everything is going to be sustainability. Also, our investment do find things. They do their own searches on the platform. So, they don't just evaluate what is recommended to them. They do some creative research about some market, some spaces. So that removes the bias as well.
Interviewer 21:17: I understand. So, AI is basically just a help that you provide to the investment team. And did you find a lot of startups through this AI system? Did you invest in a lot of them successfully at the end of the whole negotiation process?
Interviewee 2 21:40: There are quite a few, I don't know exactly the percentage also because I haven't been with the company for the entire life, so I don't know much about the past investment. Everything goes through the system and is managed through the system. Sometimes there are cases where someone heard of a company through some other means. And then actually that company was already in the system and there was already high in the ranking. So, in that case, was it the investor that found out about the company or the system that recommended it?
Interviewer3 22:30: So, you don't have a way to distinguish companies, because it goes all on the same funnel, correct? Can we call it like a CRM system, in the end?
Interviewee 2 22:45: Yes, it's a CRM under many aspects but sometimes the merit is kind of both at the same time.
Interviewer 22:59: A couple of last questions. This is proprietary software; do you just use it yourself? Or do you use license it to others?
Interviewee 2 23:13: Just ourselves, it's just an internal software.
Interviewer 23:15: Okay, and how often do you update it? You told me that there are six software engineers, before.
Interviewee 2 23:23: Everyday we do several releases. Sometimes the changes are very small. But it's a continuous improvement. Like a proper software company.
Interviewer 23:38: Did InReach Venturea start as an AI-led VC, or was it more of a traditional VC in the beginning and then changed?
Interviewee 2 23:46: It started as an AI VC. The initial idea, and the initial found was driven by AI.
Interviewer 23:58: I'm also trying to understand if there's some cases of VCs that are traditional and then changed to this type of sourcing and screening, but it seems that this is not really the case. VCs usually start already with this vision in mind.
Interviewee 2 24:13: Yes. It throws back to Roberto Bonazinga. He has been a traditional venture capitalist for many years. He has worked at Balderton, has done very traditional work, I mean, meeting people, having lots of calls, participating to events, flying around the world and working with a lot of people that were sourcing deals for him or sourcing deals himself. And the human is the bottleneck in this process. His vision was, I can solve lots of the problem in the first step of the value chain, with sourcing technology.
Interviewer 25:08: Ok that's all**.** Thanks a lot for answering my questions, it has been very useful.
Interviewee 2 25:11: Thanks to you, I hope it was helpful. Bye.

Interview 3 – Hatcher+

Interviewer 0:04: Starting with a few general questions, I read that you offer VC-as-a-service. I wanted to understand if you are also a VC yourself, so if you also have your portfolios and you invest yourself? And in the VC-as-a-service part, who are your main clients? Is it corporates that are trying to do CVC or just other ventures to which you're offering just the AI deal flow service?
Interviewee 3 1:45: How we came to build our AI strategy is as a result of one my founders initial experiences in 2013 of leading an angel group to create a venture strategy. This group of angels, which later on expanded to 49, basically put together $20 billion and spread it across 15 companies that they thought made sense. What they realized pretty early on was it five went out of business very soon. The remaining 10 had to be worked hard in terms of getting the management team to change or in some cases to pivot the business. And there were no answers that they could give to the other investors and whether they were going to see their money back. And that's typically what happens. When you come in at the earliest stage in the venture world, statistically 60% of VCs that take the money never return the money. 75%, including the 60% will give you returns which are equal to NASDAQ ETF returns. This means that you're better off investing in NASDAQ ETF than investing in venture capital if you want your money back. Only 25% of VCs give you anything in excess of 2x. The majority don't. And that became our quest or John's quest initially to find if there is a way to kind of deliver predictable index returns. What we figured out was by aggregating venture data spanning 22 years that forms the foundation of our AI as well as the data science model. We constructed 4.3 billion virtual portfolios and we

thought we'd get 10-15 strategies out of those, we got one strategy. And that one strategy was that you need to construct a very large portfolio in excess of 500-2000 companies. In our case it was 1350 companies. That will throw a return of 4.2x on average after taking into account mortality. The other thing that we found is constructing a portfolio 1300 companies sitting in this small island called Singapore is going to be very difficult. So how do we get that unicorn out of LA or New York or whichever part of the world? So, what we did was we went and interviewed 160 VCs and family offices across the world. We took copious notes, everyone had their own strategy, but one thing that became apparent was that professional investors or seasoned VCs were investing in top 1% of the deals that came across their business on their desks, right? We've seen 100 business cards; they were picking up one and investing. So, we said "okay, now to construct 1350 companies that means we need to look at 135,000 companies". That's a lot and manually, it's difficult. So, it is possible, but it will not testify in terms of the cost benefit. So that's kind of what prompted us to come up with a fund strategy. We're not a VC ourselves. What we do is we empower other venture capital firms. In in this case, we partner with first degree asset management, and we've created a fund strategy that they manage, and we basically aggregate the deals, and we use the technology and invest the money that they have across those deals. So that became our strategy. What we realized was that we needed to do a pilot. We raised a small amount of money in 2018. By 2020, we had made 100 investments, the last one and a half years unfortunately thanks to COVID have not got as much ahead. But we are back and what we are doing now is we already have 114 companies. And what we typically do is we come in at the earliest stage now data suggests that this is the best place where you make returns. Yes, you will make money in terms of multiples in series B, C and D. But as you can see both the dispersion of returns as well as the returns is reducing. So, the risk is very high, but returns are high. This is Angel round. What we believe happens in angel round is angels typically have a halo on their head. And typically, it's friends and family who come and ask. There are no questions asked for valuations. You will cut out a check and give it. You always overpaid and you probably don't get money back. This is the worst place to invest. What we decided to do was to invest into any company at this stage. And we take 5% stake for $50,00 to $100,000 we exercise our pro rata at pre-seed, seed, or series A and then we stop. What we want to do is basically push up the pro rata rights to our investors or LP so they can continue the journey on successful companies. because the other thing that is happening in the venture world is this is a trend shift. If you've been following the news, all the big boys whether it is Sequoia, Andreessen Horowitz, Softbank, they have now gone on to say that they are investing in more than 500 companies. We probably are the first ones to come out with research in 2018 saying that you need to construct a large portfolio and this is summarized here, which I mentioned. On average, this green line is a 100-company portfolio under the portfolio, 40 initial investment 60 follower. On average, it can give you 2x returns. 75% of VCs are here. 25% are here. The yellow line is representative of 500 startups, which is an accelerator based in the US again I guess you're familiar with them. The first one is returning a net 8x to their investors, which is simply phenomenon. What they are trying to do is push the envelope a little bit further. If we do anything without biasing towards good quality on average, we'll return 4.2x worst cases 2x. And if you do anything towards quality will end up somewhere here. Logically, if you look at it, the way it works is if you study portfolio theory in public markets, diversification beyond 12 or 20 companies does not make any difference. Fund managers do it to avoid concentration risks. There is no incremented value in the size of the portfolio. But in the private venture market, it's like casting a fishing net in an ocean and not in a pond. If you were to create a portfolio of 10 companies, and if none of them are unicorns or multi baggers, basically under small fishes, they will die, and you will get nothing out of it. But the moment you cast a big fishing net in the ocean, you will catch some small fishes you'll get some tunas, and you will catch the occasional whale which will return. So, that's what happens in venture. The moment you go from 100 to 500 companies you are doubling the returns in venture capital. After that it becomes a marginal increase. This is what is happening in the venture world right now. One is that everyone's creating mega portfolios. Two, everyone's trying to get in early. And the reason everyone's trying to get in early is because late-stage valuations are high and if you're sitting on an Airbnb early on, Sequoia is not going to allow you to sit on Airbnb data. So, everyone's trying to get in early and aggregate the pro rata rights. And for all of this you need some kind of screening process, which we believe AI can provide. And the funny thing is that, on average, none of the venture capital investors or leading VC funds invest in data science and AI.

Interviewer 10:54: I decided to write my thesis on this topic because I think that the industry is moving in that direction and I think that AI can provide efficiency, so I wanted to understand the status and the practices of the early adopters of this technology in VC.

Interviewee 3 11:57: If I look at it, holistically, deals can be everywhere, and everyone seems to be getting into the startup business. The question is whether it is Federica that will succeed or Interviewee 3. We don't know right? The question is: are we able to use AI to take some salient features from the pitch deck, on the on the background, on the idea on the geographic potential, on the ability of investors to either exit or fundraise? And then say, the first analysis is good, you should go and interview the founders and do a manual due diligence. So, when you get 1000 applications instead of trying to use your gut instinct, which is what is happening in the VC world, by using a little bit of AI. Yes, AI has its constraints, because AI cannot at least as of now, gauge what the new trends or new technologies are going to be. You need data, right? I mean, if you don't know what the new technology is, there is no way you can predict what is going to be new technology. If you have enough data, then yes, you can train the model to give you returns that you want to see. Just to

give you an idea this is my platform, as you see. So let me go to the deals page. This will basically show the 9000 companies that we have passed on in the last few years and we've invested 140 companies. What the AI is able to do is for us is to pull in the application form data. We have more than 60 API's. Let's say, if I take this company which came from a partner accelerator in New York. By the way, we have somebody from Milan, one of our partners is in Milan Fashion Technology Accelerator. So, what we do is we aggregate these who are qualified accelerators, and we invest in all the companies that they invite to the cohort, what typically happens is when an application comes in to the engine, the engine first and foremost will check for data quality, if the data quality is robust, then it is pushed into the AI engine. So that the AI engine does not get corrupted. What AI does is it basically takes in the details from the application form. I am Jack is my website. This is what the idea is all about. We are based in Austin, Texas. This is my pitch deck, short biography of the CEO. This is the money that I want. This is what will come to you in the very early stages for a company there is going to be no quantifiable data. So, the AI picks this up. We'll also go and check the 60 API exchanges that we have and get everything that we can find on Jack from LinkedIn, Facebook, CrunchBase, Twitter and also the company. AI basically comes out with a scoring on this.
Interviewer 15:33: Okay.
Interviewee 3 15:33: We call this the first pass. As an example, I know this company is in the business of aquaponics hydroponics, the venture trend is pretty strong at this point in time. Geography potential of only having this company doing something in Austin, Texas is pretty average. The AI cannot take anything about it All the data points we have are similar in other similar companies and stage in our database. The CEO is 24-25 year old, he has very limited experience or startup experience. The ability of his own network to exit or fundraise is below average compared to other founders that the platform has seen right? The company does score on the Zero Hunger SDG. At this point in time, there is no company in the aquaponic space which has had an exit so the AI is not able to score it so it gives an average score. On a composite level, we use text dimensionality, which is an important feature of AI to come out with a memo like this. This uses NLP text dimensionality, and we use it for relevance matching. Everything that you see out here is auto generated. There is no manual intervention. AI comes out with this memo, we push this out to the analyst or the partner, they can read through is they can figure out what algo models were used. And if the score is decent enough, they will invite Jack to meet them in person. Then you're going to get into more specifics.
Interviewer 17:13: So, the input from the AI is the application of the startup. So basically, the AI just does deal screening and not deal sourcing, it does not come up with companies itself?
Interviewee 3 17:31: We have that, but we don't use it for our own portfolio. We don't use it right now. But what typically happens is once the data quality is robust, this is the score in the last 500 days in terms of data quality and predictive EIB scores, but this is the predictive score. If you've got 1000 applications, you can basically decide I don't want to look at applications below 550. I only want to look at the top scoring application. So, it becomes a screening tool and to answer your question, there is something called Deal Scout feature. Just before this call, I was presented to someone this seems to be a bug and I'm not sure if it is still working or not. But what it does, if you have a mandate, you are looking for a company in vaping, for example, today what you will do, you will go to Google and type vaping or you will go to Crunchbase and do the same. Or you will put AI technology in vaping I don't know I'm just making it up right now. That is a string-based search that Google will use or Crunchbase will give you everything that matches AI everything that matches vaping and everything that matches AI vaping. You want to know what is vaping all about is there something which is relevant to this and another feature that comes into play is text dimensionality, we are trying to find out which is the most relevant thing that can come out and close this to the market described. So, if you go and read the descriptions of all the companies in out platform and come up with the top 250 countries which would be the nearest. So right now, it's broken, I think it will not work.
Interviewer 19:41: It's not a problem. I just want you to understand if you also use AI for deal scouting.
Interviewee 3 19:46: We do. It goes into 219,000 companies. As you can imagine, only 10,000 have come through put platform at Hatcher. Yes, maybe some things have come through other people who are using a platform, but they don't want to disclose. And we also pull in deals from the web and CrunchBase and pitch book.
Interviewer 20:10: Okay, and do you also use AI for any other venture stage later in the process? For example, due diligence negotiating, closing?
Interviewee 3 20:37: No, we don't use anything after this stage. At this point in time we are contemplating, we are gathering data to see if there is a way we can use data points to suggest if there is a better way to kind of identify founders who are going to be successful. But we are not embarked on that journey. we're collecting data at this point.
Interviewer 21:13: So you just use this platform yourself and provide the output to VCs you are working with, you don't license the platform?
Interviewee 3 21:27: Anyone can use it. We can create a community for Mackenzie, we can create a community for angel groups. There are more than 200 groups are using it.
Interviewer 21:43: Okay. And how long did it take to develop?

Interviewee 3 21:51: We've been at it since 2016. From 2016 to 2018 we were largely in the process of gathering data and running the various simulations as I say that we've built 4.3 billion data science models, and after that, we got into coding and it's taken about three and a half years and $10 million + to be here.
Interviewer 22:19: And which type of AI do you use? Is it supervised, unsupervised?
Interviewee 3 22:32: I will qualify that answer saying that I like to refer you to Dan. He is the guy whom you should be speaking to for more technical aspects of the thing. I don't know if you had a look at our white paper.
Interviewer 22:45: No, not yet.
Interviewee 3 22:47: Okay, so if you go to our website, you can download our white paper. We have a lot of other research in terms of blocks. There'll be written on AI and human. Because a lot of people have questioned as saying that humans are good at judging. We don't dispute that. What we do believe that AI helps in removing a lot of noise and that's what we would like to do.
Interviewer 23:14: Do you think that a lot of VCs are going against this technology innovation because maybe they don't trust the AI or they don't trust the output that it gives, or they prefer traditional model? Do you think that this could be a blocker in in the adoption of AI in in this stage for VCs?
Interviewee 3 23:36: A research that Gartner did last said that by 2025 75% of VCs will be using data science. And the question he asked is and it mentioned to you earlier that an average VC does not spend more than $100,000 on data science. They invest in companies that use data science and promote AI, but they themselves are not using anything to that effect and the reason for that is that everyone is busy deploying capital and investing capital using the traditional process. I guess we will be the ones who power them or at least get to that qualifying stage. At this point in time, when we speak to VCs, or more established VCs. Having said that, some of them are doing it. We do know that 500 startups is using kind of data screening model and I'm not sure if it is an evolved AI. We do know that EQT Ventures is using something called Mother Brain. They are basically trying to gauge data points on founders to see whom to back and not the back. That's something they're out there. We know 500 startups, Sequoia has built something internally, they use that. But again, these are all closed group software. We probably are the only ones on a global scale that's doing what we have done.
Interviewer 25:23: There's not a lot of information that's available online on who is actually using AI and who is not using AI and instead just doing some automation processes. So, it would be interesting to get in touch with as many of them to understand who's actually using AI. I think it's really interesting that you mentioned that because I think it's a really different thing to build up arbitrary software for your own VC, rather than just have a platform that you have, that can be used by many. What do you think is the best way for a VC to implement AI?
Interviewee 3 26:03: At this point, depending purely on AI to make an investment decision in venture is far way out. AI cannot do some things. Number one, it cannot anticipate future trends because there is not enough data available. Number two, AI cannot cannot gauge the passion, sincerity, deep domain knowledge of a founder. You need to eyeball the founder, you need to ask those valuable questions, you need to feel the connection. So, at this point in time, it has not evolved to that level yet. What we do believe is that as we gather more and more data, we can get more refined and we'll heading in that direction. But at this point in time, it will largely assist in decision making, it will not make the investment decisions. There are a few groups who use our platform, who use our software and they are purely depending on AI. We do not recommend it. They seem to have had a good experience so far. But as I said, we don't recommend that we would censor that. Do the manual due diligence. What AI helps do is screen the top companies that come to you and then you interview them and identify the top one.
Interviewer 27:30: You think you might lose a little bit of human connection that you can have in VC that is very important. Do you think it's also a matter of investment? In the sense that building and developing your own software takes a lot of capital, so maybe the big VCs are the ones that can actually build their own proprietary software, while the smaller VCs might have to rely on third party providing the software. Do you think that in the future that might be the case that the smaller VCs will have to rely on someone, like as a third party, because they might not be able to develop it themselves? And if so, wouldn't it be difficult for them to get a competitive advantage in the industry having standardized software?
Interviewee 3 28:39: Two questions you ask, one is capital and the second is who has the data, right?
Interviewer 28:57: Yes.
Interviewee 3 28:57: It took us one and a half years to collect data, 600,000 transactions, 22 years of data in a doc. In this collection exercise you need to basically clean and make assumptions and provide for all of that training. So, it took us a good amount of one and a half years to do that, get that data and clean that data and then run the models. Even if even if you look at the big boys out there, on average, I don't think there is even more than 2500 applications in a year. Getting data will take time. The second question that you asked was, will it give a competitive edge? I mean, again, everyone has their own strategy. Some people invest in 10 companies some investment 20 Some are invested in 500. So that is one is the scale. Some people believe they're good at picking founders. So be it. 500 startups and everyone else will say no, we rely on numbers. We don't know which one is going to succeed. So, while some of it might seem that it is going away, but people will find different ways to differentiate themselves. It could be based on themes it could be based on geography

is it could be based on stages. It could be based on sectors and everything and mix and match. So again, everyone may not use the same cutoff of 500 companies. The biggest challenge in VC is deal flow. Even if you have the best AI and no deal flow, it's not helping. You need a quality deal flow. Of course, the big boys have access to what they believe is quality deal flow. And the question is whether the AI can tell them that out of this, you don't need to spend so much bandwidth on 75% of the companies, just spend it on 25% of the companies. That's what the AI is going to do. It's going to allow them to give more latitude to spend more time on top 25 companies rather than 100 companies. They can dig deeper; they can get more efficient in deploying capital across those 25 or less.
Interviewer 28:59: So, you are saying that the AI is not going to provide directly the competitive advantage but the strategy and the way of using AI is going to improve deal flow?
Interviewee 3 29:10: Correct.
Interviewer 29:12: That was my last question, thanks a lot.

Interview 4 – Pilot Growth Equity

Interviewer 0:00: Starting with the first question, I read that you define yourself as an AI-driven deal sourcing and portfolio company value creation VC. Can you elaborate a little bit on this in general terms?
Interviewee 4 0:10: We're pilot growth, so we help pilot the growth of other b2b software companies, and our founders have been software developers, created software businesses, and grew them and sold them. And so not only are we investing in companies that use AI, but we also use AI ourselves. The value added, the value creation side of it, once we make an investment in a company, we use our operational experience and our network and our advisors to really help the companies that we invest in helping them grow by helping them first and foremost get new investors, but also kind of institutionalize different practice areas in their betas, whether it's going to market, or product management, or corporate governance. So that's kind of how we, we add value. But on the deal sourcing side, we built our own deal sourcing engine called NAV Pod, which leverages AI.
Interviewer 1:31: Yeah, I read about that. And I was wondering if you use NAV Pod and AI just for deal sourcing, or if you use the AI in any other stage of the VC process? So, for example, deal screening, due diligence and negotiating, closing and post-deal, as you said, helping the companies you invest in grow? Or is it just deal sourcing?
Interviewee 4 1:54: It's just the deals sourcing and deal screaming. At the end of the day, the people are still cold-calling the companies and building a direct relationships. It helps us, we start with about 35,000 companies. And we want to narrow that down to about 3000 companies which NAV Pod then follows really on a daily basis. It says hey, here are the 100 hot deals. It sends us an email every day and says: here are the companies you could call today. Because every single day it's trying to find the most relevant deals we should be calling. So that's really what NAV Pod does.
Interviewer 2:42: Okay, that's super clear. And which type of AI does it use? Is it supervised learning or unsupervised learning?
Interviewee 4 2:50: It's unsupervised learning.
Interviewer 2:59: Which kind of indicators does it track to understand which are the best deals day to day?
Interviewee 4 3:11: So, we track 15 leading indicators, and then there are dozens of other data points. But we're looking for the signal around a point of inflection of growth. We want to be the first institutional capital. We're looking for companies signaling that they're going through a growth phase. And it could be a good opportunity for them to take on an investment round. So, the first things we're looking for are the space they're in. It's going to be in b2b software. It's going to be in some distributions they put on their websites and their marketing materials. We use natural language processing to sort of breed everything that's from the company. We also look at any sort of financial filings, state filings, we look at their form D filings, about what they've done in terms of capitalization or registration as the company, whatever LLC or C Corp. We look at patent data, patent files that are pending. We look at grant data, any grants that they've been awarded any grants that they are highlighted for, that's sort of on the patent, a government filing side. We also look at product and product releases, so we're looking at them launching new products, promoting new products, training. We're looking at them speaking at events, about their products, about their offerings, so really trying to understand the evolution of the product and as it grows, with new features and benefits. We are able to pull down customer data and partner data. So that helps you understand the growth, what kind of customers they have, we understand the names of the customers, the customer count, as well as channel partners, you know who they are working with in technology, in a go-to market capacity or in a technology partnership capacity. And we also look at job postings and hiring data. We're looking at the types of jobs that they are hiring for the numbers of jobs are hiring for, the locations in which they're hiring, at what level they're hiring, and what functions they are hiring. That's a big piece of it. And we're also looking at their competitors, right so we're looking at how are they? What's their signal and context, you know, their competitors, whether they're in FinTech or whether it is supply chain, or cybersecurity, understanding what are the other companies in the space and how do they look comparatively across kind of all those factors?
Interviewer 6:19: And you have any data point related to founders to understand their expertise and value? Do you track their past experience? Or is it a data point that does not fit in NAV Pod?

Interviewee 4 6:33: No, we do track the founders, we track the key or key contacts on each profile. We focus on key contacts and those are usually the founder, the CEO, perhaps the CFO or a corporate person, and we do profile their backgrounds, and we look at the companies they've been involved with. We look at the different work experiences. We look at the university background educational background. Recently there's some controversy. People criticize it that you know, you continue to propagate profiles of people who have had privilege and that sort of thing. The whole kind of equity movement out there around like using profiling. But at the end of the day, our job is to buy companies that will perform. We're not a government agency. We don't have a role in the world. But you know, people who don't have a good track record and don't have a good educational background, you know, you're not necessarily going to highlight them.
Interviewer 8:00: They are less likely to perform a statistically. And how long did it take to develop NAV Pod and the whole process?
Interviewee 4 8:11: To get it to where it was working the way we want it to work took about two years. I think part of it was to have to build the model, and then to have it learn from how all the data which we were gathering manually around these companies and all the data feeds that we were connecting into the platform. So, it took about two to three years to get all of the inputs into one place. And to use the ones that we thought were giving off signal and then for the perceptron to really grow and start to be able to replicate what we would do manually if we were making the decisions of highlighting who's a hot deal and who's not a hot deal.
Interviewer 8:59: And how many data points did you start with to start the AI?
Interviewee 4 9:05: We started with probably between 50 and 100 data points, we started with that we didn't use anymore but I would say from a profile perspective, like 50 to 100. That's about it. It wasn't much more than that.
Interviewer 9:25: And do you just use NAV Pod yourself, or do you license the AI to any other VC?
Interviewee 4 9:33: We just use it ourselves and we've had a question, truly every time we talk about NAV Pod. It's very focused on our strategy. And you think about private equity. So public equity, there's lots of public information available. But I think all the different strategies in buyout, which is your patrol investing, they're run by bankers. Those processes are run by bankers, so everyone has the same information pretty much. And so you don't really have an information advantage and buyout except, Are you willing to pay a higher price? And then in early-stage venture, there really isn't much signal because the companies don't have anything. They don't have a product. They're hiring people and they might raise money, but they're not really engaged with the market, right? In growth, these companies are growing quickly. And they're spending a lot of their energy creating signal, hiring people going to conferences, writing marketing materials, talking to customers, helping customers. So, there's a lot of signal in growth. But in venture there's not a lot of good signals, because there's not there's not a lot of activity. And then in buyout there's a lot of signal, but everyone gets the same signal, and it's managed by third parties, right? So, you can get some advantage. To take NAV Pod and kind of change it for other strategies would be quite a bit of work. And we wouldn't obviously license it to people in our own strategy because that's what we're competing against.
Interviewer 11:22: What do you think about new software companies that are developing AI software's that are useful for deal sourcing or VC as a service in general, that might be applied to different VCs in different geographies in interested in different industries in different companies? Do you think it's a viable option? Do you think it's customizable enough in that way?
Interviewee 4 11:48: I do. Absolutely. I think that all software innovation is about creating perfect information and leveling the playing field. And the idea of finding companies is being rapidly commoditized. It used to be an advantage and for Pilot, we're not finding companies that no one else can find. We're finding companies much more efficiently than other people can find. I can find a company to have a conversation and spend $1 do that. Where my competitors in other funds like me have to spend $100 to find that thing. The idea of having software that helps you find the right people to talk to, I think is a good thing. Because it's going to commoditize that. So everyone can find the same companies, right? Where do you compete? So I think finding deals won't necessarily be a competitive advantage. But finding really, really good deals that you want to invest in. No, software is going to help you do that. So if you want to have hundreds of companies to look at, you really want to do that really quickly and really cheaply versus building your own team. But you also have to realize that all of the other competitors that you are against, will now be looking at all of that company.
Interviewer 13:25: I guess it's still really, really important to have a proprietary deal flow that you get through your network and not through the AI software, correct?
Interviewee 4 13:35: Yes. The software can help you if you have a list of names. You could use that software to help you manage it better. Similar to what CRM is. If you've got a list of all of your customers, the CRM software has helped you manage your customer relationships better. I think if you're saying oh, I'm going to subscribe to this service to find deals. I think as an investor, you're in big trouble. If you are relying on a third party to find the deals you're going to invest in. I think you can subscribe to lists to get named then. But then you've got to take that list. And we do that today, right? We subscribe to data platforms to get to our 35,000 names. And then we whittle it down.
Interviewer 14:24: Yeah, so the value added of developing your own platform is to include all of this and have something that can actually create value and be a competitive advantage compared to a competitor that just uses third party software.
Interviewee 4 14:40 Yes, yes.

Interviewer 14:42: I think it really makes sense. And in regard to NAV Pod itself, have you hired a data scientist or someone that takes care of NAV Pod if it has any problems or anything? Or do you just do it yourself like anyone in the company?

Interviewee 4 15:04: We were very fortunate to have one of our founding general partners, William Lee. He was a computer scientist at Carnegie Mellon and built three companies prior to pilot. And when we decided, hey, we want to use machine learning, to help us manage our deal sourcing process, he was able to personally build the product himself, and not a lot of firms have that. But I would say having people in house is important. I think that when I talk to lots of firms, and they'll hire a consultant, or they'll hire an independent contractor and say, "Hey, build this for me, I'll pay you some money", and then they go away. This is a continuous process. We don't build it once and then it's finished. It's never finished. And so you look at companies like Blackstone. Their private equity team recently hired the head data science person from Steve Cohen's, hedge fund called Point 72 and they built 50 person data and analytics team within Blackstone, to help them with their process of looking at deals and analyzing data. I think that you're going to see firms have these data teams, just like public investing firms have data teams, you'll see private equity teams having more and more internal and that's helped them find deals, but deploy it might help them also.

Interviewer 16:43: Sorry the line cut off. Where are you adding something else?

Interviewee 4 17:12: I think AI can be applied to class the other functions of a private equity firm, not just the way we're using it, but it could be applied to other areas of management customers to client reporting, things like that.

Interviewer 17:26: Did you notice any difference in terms of profitability of your portfolios and of your funds after you started using NAV POD compared to before? And if so, which kind of KPI or metric do you use to track that?

Interviewee 4 17:49: The main difference was that our deals were completely on strategy. When you're looking at deals when you have your strategy of what you want to do, but a lot of times you'll meet a company you'll meet an entrepreneur and you'll really kind of fall in love with the deal, you want to do it and you do it. And it might have some aspects of the deal which are not exactly on strategy and you kind of make exceptions for them. And you make a little bit of a well, it has some things we are looking for but it's missing a few. And previously, a discipline of staying on strategy was harder without NAV Pod, but with NAV Pod when you look at all the deals we've done, it keeps us on strategy of looking at the deals that that meet our criteria, and then investing in the deals that meet your criteria. Where if you don't have that discipline, you can tend to stray a little bit and then you end up with companies you've invested in, and you say this isn't really working out the way that we thought it would. And the reason is because it's not on strategy. So that's been the biggest impact. I think it keeps us focused on what we want to do, because it's very easy to get distracted. And you see that today with a lot of these big companies that are investing in seed stage companies. Everyone wants to be a VC now, right? And the reality is that it's very hard to be a VC. But you see people because they think it's popular string into venture strategies because they think it's easy, where they should stick to what they really do well, which might be buy out or other types of strategies. So, we'll see how those worked out for them.

Interviewer 19:39: Are you afraid that maybe with the evolution of these software and AI and machine learning will help other companies to become a VC and compete against already existing VCs? Or do you think that experience in the field will be still valued a lot in the future to succeed in this industry?

Interviewee 4 20:00: I think that if you are a follower, and you don't lead deals, and you're just trying to be part of a syndicate. I think this will help find help you find deals you might want to invest in. Now whether they let you in or not, is a different story. I think it's really helpful for people who don't want to lead deals and say there's a $50 million deal and they hope they get like 2 million and they don't negotiate terms. They just take what they can get. This is very good for that. But if you're going to lead deals and help an entrepreneur develop the company, this doesn't help you do that at all. Having a venture partner with an entrepreneur and it helped them guide really, most of success and venture investing. Because when you invest in in technology, there's 30 or 40 companies that are doing the exact same thing with pretty much the exact same technology. So, you're not investing the idea of like, "Oh, we're going to find this one technology company that has this one thing, and it's going to change the world". When you look at a space, there's 50 companies doing the exact same thing with the exact same technology. And so it's the teams that navigate and get success who win and then they consolidate all those other companies where they go out of business. So, helping the company as a venture capitalist and helping guide them is the business. Finding deals, it's either going to find deals cheap, or following the expensive way to find deals or cheap way to find deals. like that's, that doesn't matter as much. It's going to help the person who wants to follow who's basically the dumb money. We're like, "Hey, if I can put $2 million in this company at $100 million valuation, I'll put it in. I'm not going to set the valuation. I'm not going to negotiate terms. I'm not going to be on the board. I'm not going to help those people, but still I have to let you in the deal". If I find this company and call them up, and they say, "well, Neil's a lead investor, so Neil decides who gets it". When they talk to you, they are like, "who are you? why would I let you in?" So, if you're trying to use technology to find leads, and you don't know how to add value and work with entrepreneurs. You're going to lose all your money.

Interviewer 22:36: How have you welcomed PAV Pod in the organization? Has there been any resistance on your side or on your colleague side, or any difficulty in starting using it or anything that could cause any problems?

Interviewee 4 23:01: I think human nature is that we have recency bias and confirmation bias or too big name biases. So, we make decisions based upon a lot of bias. A friend of yours will call and say "I have this deal I should look at". The recency bias, and it came up today, and it's a friend of yours. There's confirmation bias, you will tend to say "oh, we should look at this deal". And spend four hours looking at this deal rather than spending those four hours looking at the hot deals in PAV Pod. We're always networking, you're always looking at NAV pod. But we're also listening and hearing and meeting other people, right? It's always a challenge to stick to your process like with anything in life. Like if you follow the process and are disciplined about it, you will get good results, whether it's exercise or playing an instrument or playing sports or cooking. If you stick to the process, you'll get results. But human nature, it's very hard to stick to the process. So the key is constantly managing and when we look at our pipeline and making sure that we're working Nav Pod by deals, that we're spending our time on them and we know we get the best results we do that but when we meet other entrepreneurs or we get we have an introduction or we hear about a company that we're interested in so there's a constant adherence to using technology and I think that's consistent across all technology within companies. You have great technology, there are people using it. Now, it's easier for us because we were born this way. We started this way. If you were a company that had an old process and you're trying a new technology process, you probably have a more difficult time getting adoption. Especially people who were trained in a different way. They were trained to get deals from networking, or they were trained to get deals from bankers. There'll be a lot. It's hard for them to switch and change their behavior.
Interviewer 23:00: When Nav Pod send you the everyday email with the most relevant deals, do you go through all these deals? Do you dive deep more and more and more to find like to understand if it's worth it or not?
Interviewee 4 24:41: We divide up the list and we make sure someone is calling them and starting. Because our goal with Nav Pod is to get to have a direct conversation with the founder. And that's the goal. Once we have a conversation with the founder, Nav Pod is over. And so we have to call and get in touch with the founder. And then ask that person more questions about the company and introduce Pilot and say: "hey, we think we could be a good partner." And then many times we will call and get in touch and we look for companies with $5 million or more of recurring revenue. So many times, we'll call a founder and have great conversation. They'll say, "Oh, I'm doing 1 million a year of ARR". And we'll say "okay, that's great, we'll keep tracking you" and we have track companies who've gone from 1 million to 5 million and not invested. So, we have to do all that screening. So, a lot of the companies fall out because they don't meet that 5 million. Sometimes we'll call and they'll say: "Oh, I just closed the series A last week". Our signal is that good. We're saying you should call this company and like literally a week earlier, they close their series A. So our goal is to get them on the phone and then it's also a personality match. If there's someone who we don't think we could partner with, then we will continue the conversation. Its goal is to have us have a direct dialogue with the founder. And then from there, we take over and start our due diligence process of gathering information about the customers, we try to get a product demo we try to understand who the customers are and talk to them. We talked to their management team, we talk, we get the financials. And then we have a better picture. Do we want to issue a term sheet or not? And then once you have a term sheet, complete due diligence to close the deal? So it's kind of the beginning of that process goes from deal sourcing to kind of deal negotiation.
Interviewer 28:18: And who reviews this everyday email? Is there someone specifically or does everyone do it?
Interviewee 4 28:24: Everyone gets it. It's a bit of a we call it jump ball. It's everyone's responsibility. A part of Nav Pod is our own CRM. And if you're going to call a name, the first thing you do is go into Nav Pod and see that profile. And you'll see if someone already called it, so you don't have to coordinate. I'll see like, "oh Federica called him this morning. Okay, great, and then we'll go to the next one". And they'll say: "oh, we have called them already". So, it's a very much a self-managed process. There's no centrally manage, there's no one assigning you to do this. You have to do that. And then we look at all the deals that are being worked. It is anyone that can do the calling and then sometimes I will look at a company and I'll say: "I think by looking at the profile of his company, it's in cybersecurity and Federica has a better background and maybe matches this entrepreneur better. I think you should call them because I was going to call them, but I think you'd probably be a better match". Because you have a better understanding, it's all about matchmaking within this level. You want to build a relationship between the entrepreneur and the investor.
Interviewer 29:32: I think that was my last question. Thanks a lot.