

Jia, Peiyi; Stan, Ciprian

Article

Artificial intelligence factory, data risk, and VCS' mediation: The case of bytedance, an AI-powered startup

Journal of Risk and Financial Management

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Jia, Peiyi; Stan, Ciprian (2021) : Artificial intelligence factory, data risk, and VCS' mediation: The case of bytedance, an AI-powered startup, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 14, Iss. 5, pp. 1-19, <https://doi.org/10.3390/jrfm14050203>

This Version is available at:

<https://hdl.handle.net/10419/239619>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Article

Artificial Intelligence Factory, Data Risk, and VCs' Mediation: The Case of ByteDance, an AI-Powered Startup

Peiyi Jia ¹ and Ciprian Stan ^{2,*}¹ Robert J. Manning School of Business, University of Massachusetts Lowell, One University Avenue, Pulichino-Tong Business Center Suite 228, Lowell, MA 01854, USA; Peiyi_Jia@student.uml.edu² Rinker School of Business, Palm Beach Atlantic University, 901 S Flagler Drive, West Palm Beach, FL 33401, USA

* Correspondence: Ciprian_Stn@pba.edu

Abstract: The AI factory is an effective way of managing artificial intelligence (AI) processes, enabling broad AI deployment in a firm. The purpose of this study is to explore the role of the AI factory in an entrepreneurship context. How do AI-powered startups leverage AI to grow, and manage data risks? What is the role of venture capitalists in this process? We answer these research questions by conducting an in-depth study of an AI-powered startup: ByteDance. Our study extends both AI and entrepreneurship literature by showing that AI-powered startups adopt the AI factory approach to optimize scale, scope, and learning. Our discussion also emphasizes the critical role played by venture capitalists in assisting AI-powered startups in building AI factories and in reducing data risk.

Keywords: artificial intelligence; scaling growth; AI factory; AI-powered startup; data risk



Citation: Jia, Peiyi, and Ciprian Stan. 2021. Artificial Intelligence Factory, Data Risk, and VCs' Mediation: The Case of ByteDance, an AI-Powered Startup. *Journal of Risk and Financial Management* 14: 203. <https://doi.org/10.3390/jrfm14050203>

Academic Editor: Michael McAleer

Received: 23 March 2021

Accepted: 23 April 2021

Published: 2 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence can process, analyze, and interpret data much faster and at a scale unachievable by human capabilities, potentially leading to increased prosperity, knowledge, and comfort (Fountain et al. 2019; Mannes 2020). Despite the increasing attention paid to AI and its capability for business enhancement, little research focused specifically on AI in an entrepreneurship context (Obschonka and Audretsch 2020). It is far from obvious how AI and machine learning technology are applied at scale within a startup, and how it can help accelerate a company's scaling growth. This study focuses on AI in the startup context and aims to answer three research questions: (1) What is an AI-powered startup and how does it manage AI processes? (2) How does an AI-powered startup manage its data risks? (3) What is the role of venture capitalists (VCs) in managing AI and its related risks?

We define AI-powered startups as firms who establish and design their organization around AI from their inception, and leverage AI to achieve fast growth. ByteDance is a typical example of an AI-powered startup. As one of the most valuable startups, ByteDance applies AI to every aspect of its business (Chen 2020; Byford 2018). Its application of AI algorithms and machine learning techniques power a large variety of apps, including the popular video-sharing app TikTok and its flagship news aggregator, Toutiao (Li 2013; ByteDance 2020). The key feature behind these apps is the AI-backed recommendation algorithm (Li 2020). The data of active users is collected and processed to build a workflow of recommendations through advanced AI algorithms.

Through an in-depth examination of ByteDance, we show that AI-powered startups optimize their scale, scope, and learning by adopting the AI factory approach. The AI factory is an effective way of managing AI capability and generating AI solutions at scale (Iansiti and Lakhani 2020). It allows firms to industrialize the collection of data, its analysis, as well as the decision-making processes in an organized and systematic way. An AI factory

integrates data, algorithms, experimentation, and software infrastructure and creates two virtuous cycles to drive the company's scale and scope.

Through examining the organizational characteristics of ByteDance, our study shows that effective organizational architecture and agile culture are critical conditions for building and running an AI factory. However, an AI factory approach significantly raises the level of data risk, not only due to cybersecurity challenges that may result in data breaches, but also due to the potential amplification of user biases on the one hand, or due to algorithmic bias, where input attributes directly impact output quality, on the other hand. These data risk challenges are mainly due to the very large amounts of data collected regarding their customers. Addressing these risk issues requires additional resources. We specifically look into the role of venture capitalists (VCs) in managing the risks associated with AI factories. VCs can be solid partners for startups, as their business and financial skills can play a crucial role in the success of new ventures (Rosenbusch et al. 2013; Zider 1998). Recently there has been a large increase in the number of AI-powered startups that have been backed by VCs, with numbers more than quadrupling over a decade (Brynjolfsson et al. 2018). AI-powered startups often represent high risk/high return opportunities for VCs, given their unproven but highly promising technologies, but they must be able to transform their innovations into commercial products in order to attract investment (Santos and Qin 2019).

Our study shows that VCs have influence on AI-powered startups beyond their traditional roles of monitoring, information intermediation, strategy analysis, and development. We provide further evidence that VCs' mediation could reduce AI-related risks through multiple activities such as enhancing data governance, lobbying with governments, and structuring strategic deals.

While AI-powered startups leverage VC's resources to improve data governance and reduce political risk, they also engage in corporate VC (CVC) activities themselves by investing their own funds in external startups to further enhance their AI capabilities. Such investments exclude internal venturing, which remains a part of the company, and can have financial or strategic goals (Chesbrough 2002). Corporate VC investments may be made for financial reasons, when AI-powered firms consider that they are likely to generate high returns due to their superior knowledge regarding technology and markets. Conversely, CVC investments may be made for strategic reasons, ranging from learning about new technologies, potential expansions into new markets, acquisition of firms, to the leveraging of the firm's own AI platform to other firms (Maula 2007).

Our aim is to make four contributions to the literature on entrepreneurship and AI: (1) Entrepreneurial firms which place AI at their core from inception can derive faster growth of scale, scope, and increase their learning; (2) A flat organizational structure and an agile culture are crucial in the establishment of the AI factory for entrepreneurial firms; (3) VCs play a critical role in assisting AI powered startups deal with ethical challenges and data risks; (4) AI powered firms engage in CVC investments themselves to gather knowledge and to improve their own AI.

2. AI-Powered Companies

2.1. The Pains of Digital Transformation

AI deployment in businesses is mainly discussed within the context of large technology companies since the well-resourced large companies enjoy advantages in acquiring AI talents and making AI-related investments. Montes and Goertzel (2019) noticed that big tech firms are the dominant players in AI and that they shape its development trajectory as they control most AI resources like data, hardware infrastructure, and intellectual property. However, realizing AI's full benefits when applying it at scale within a company is much more complicated than just inserting AI into existing processes (Davenport and Ronanki 2018).

Digitalized companies that are able to realize the full benefit of AI utilize a different operating model. Thus, they achieve wider scope and higher levels of scalability by

utilizing AI to quickly learn and adapt. However, the transformation to an AI-powered company is difficult for large companies (Vial et al. 2021). Iansiti and Lakhani (2020) suggest that traditional large firms transitioning to AI-powered companies need to rewire the organization, internalize the need for constant organizational change, and develop a data-centric organizational architecture. Marginal improvements, such as the creation of an AI department, are not sufficient, but radical change to the core of the firm is required. Some organizational leaders make investments in AI pilot programs by developing the data infrastructure, software tools, and model development, expecting a plug-and-play technology that will result in immediate results, but despite early positive results, they become disappointed with the lack of long-term major company-wide wins (Fountaine et al. 2019). Realignment of the organizational culture, structure, and operations are required when transitioning to an AI-powered firm, not only the more obvious technological and human resources improvements.

This transition process is challenging and painful for large and resourceful companies as they are encumbered by existing structures and capabilities. The transition can take a long time due to high levels of organizational inertia, since employees might resist the shift due to fear of job losses, as the technology makes some of them redundant (Fountaine et al. 2019).

2.2. The Rise of AI-Powered Startups

The transformation challenges faced by large companies offer agile startups an opportunity to compete, by designing and forming their organization around AI. Thus, AI-powered startups can better leverage AI at scale by building an AI factory without going through a painful transformation. Once the well-organized system for managing AI capability is established, companies will be explicitly formed around digitalization and AI technology. In this case, AI becomes the core engine rather than a complementary enabler that powers the company's growth.

AI-powered startups require significant resources to develop AI capabilities to build their AI factory, and manage AI-related risks. This is where VCs come into play, as they serve multiple roles in the companies in which they invest, including monitoring, information intermediation, strategy analysis, and development (Hsu et al. 2014). They monitor their investments by spending time on site, acting as board members, and replacing executive team members, including the CEOs, in extreme circumstances (Gorman and Sahlman 1989). VCs can provide information intermediation, especially among the firms in which they invest, potentially helping with R&D partner selection (Reuer and Devarakonda 2017). They also benefit firms by engaging in management recruitment, raising extra funds, and consulting on strategy analysis and development (Gorman and Sahlman 1989).

In the context of AI-powered startups, VC funding offers important financial resources that are required for building the AI factory and managing AI processes. Additionally, political risk is among the most important factors facing AI-powered startups which are expanding globally. VCs expertise in managing political risk is thus especially important. VCs are helpful in assisting AI-powered startups in developing robust governance systems to deal with multiple legal and ethical issues related to AI and big data. In certain circumstances, VCs can help reduce political risk as a matchmaker for strategic deals.

Building on the concept of the AI factory approach (Iansiti and Lakhani 2020), we argue that AI-powered startups achieve a superior level of scale, scope, and learning, leveraging a well-developed AI factory that effectively manages AI deployment at scale. They minimize the risk associated with AI deployment with help from VC investors.

3. Methodology

We adopt a case study approach to explore the role of the emergent AI factory in driving the scaling growth of AI-powered startups. Much of recent management literature recognizes the importance of case study design in research (Hyett et al. 2014). An inductive case study is especially helpful for generating insights from the phenomenon under study

when there is a limited theory in the focused area (Ozcan and Eisenhardt 2009). Given the limited theories related to AI factories and specifically regarding AI-powered startups, a case study approach was adopted to analyze the application of the AI factory approach in a startup context.

This study aims to provide a deeper understanding of how startups leverage their AI capabilities to drive scaling growth and how they deal with challenges. A single case study is most appropriate when the research focuses on a single group (Yin 2009). It allows researchers to make a more focused study and get a deeper understanding of the subject (Dyer and Wilkins 1991). Furthermore, Siggelkow (2007) argues that single case studies can better describe the existence of a phenomenon. Therefore, we conduct a single case study focusing on the startup context, choosing ByteDance as it represents a successful startup operating in a highly competitive and turbulent market, whose success is driven by AI and machine learning.

Data triangulation, the use of multiple sources of evidence, is one of the most important principles for a case study (Corbin and Strauss 2008). In order to improve the validity of our study findings, we relied on search engines to collect data from different sources such as extant research, company websites, news reports, media articles, video footage of interviews in both English and Chinese. Our data covers an eight years period, from 2013 to 2021.

We generate emerging insights based on the information we gathered and triangulated among various data sources. First, we established a comprehensive understanding of ByteDance's business development and generated emerging insights by analyzing multiple reports (Chen 2020; Geng 2020). For instance, we found that ByteDance has a different structure and operation model compared to other large IT companies like Tencent (Chen 2019). Its diverse products share data, basic AI algorithms, and technological support under its flat structure, indicating an effective approach for managing its AI processes. Next, the insights we generated are triangulated with data from other sources such as online video footage (ByteDance 2020, 2021), magazine articles (Economist 2018, 2021), media review (Cornerstone 2020; Qin 2020) and news report (Li 2013; Li 2020) to ensure validity. Finally, we examined all the information collected with a special focus on how AI is implemented in an integrated and systematic way within the company.

4. Case Overview

4.1. Company Inception and Early Development

ByteDance is a Chinese technology company founded by a former Microsoft employee, Zhang Yiming, in 2012, known globally for its short video app TikTok. By the end of 2020, the startup was valued at \$180 billion (McKinnon and Leary 2021). The company has been utilizing AI starting with its first product, Toutiao, a newsfeed application, throughout all of its 21 apps, including TikTok and its Chinese sister app, Douyin. Both apps offer short, highly addictive homemade videos for a wide range of interests, including dancing, lip-syncing, cooking, and many others. Users can easily create and edit videos on their smartphones, having many special effects available. Utilizing the same AI structure for both apps allows ByteDance to address the requirements of censors in China with the Douyin app, while TikTok can focus on conquering global markets by delivering localized content. By developing and fine-tuning machine learning in its AI engine initially in the Chinese market, which is less restrictive in terms of customer data protection, and where it did not face competition from American social media sites, the firm was able to quickly expand globally and gain market share. TikTok successfully entered and dominated the global market for short video apps outside of the Chinese market, while Douyin competed successfully in its home market.

When comparing TikTok's AI to some of its competitors', like Facebook, YouTube, and Spotify, we note that the former decides directly what videos viewers should see, while the latter only offers recommendations. As the algorithms learn more about user preferences, by analyzing their interactions while using the app, they continuously improve

their recommendations, leading to increased user satisfaction and engagement. The AI analyses location data and individual viewing habits to customize and localize video selection.

ByteDance is backed by a group of renowned venture capitalist firms, from Japan, U.S.A., and China. Given the firm's full private ownership and lack of Chinese state ownership, the company faces lower interference from the state at home, but still encounters scrutiny over data security concerns and lack of trust overseas. While ByteDance has the support of various venture capitalists, the firm also seeks to acquire an interest in trendy apps, such as U.S.-based Flipagram, a photo and video creation app, as well as the French-based News Republic, an aggregator of global mobile news. The competition between ByteDance and its Chinese rival Tencent has been reflected in their acquisition patterns, as the latter swept away Reddit, an American social news aggregator, previously targeted by the former.

As a result of geopolitical tensions between India and China, in June 2020, India temporarily banned TikTok and 200 other Chinese apps, followed by a permanent ban in early 2021. As about a third of TikTok's 2 billion global downloads were from India, making it ByteDance's largest foreign market by downloads, this decision had an immediate direct impact on firm strategy and performance. Authorities in multiple of its other markets, ranging from France, Italy, U.S.A. to South Korea, scrutinized firm practices regarding the protection of user data, especially the mishandling of young users' information (Lin and Xiao 2020). The vast amount of user data processed by AI to drive deep learning is a major concern of regulators.

4.2. Global Growth

The firm's founder Yiming acknowledged that China only hosts about 20% of the world's Internet population, so in order to reach the other 80%, the company sought to expand on a global scale. As Douyin's main short-video Chinese competitor Kuaishou, which is backed by Tencent, had already expanded internationally, the firm started its own aggressive international expansion. ByteDance's Douyin earns most of its money from online ads and sales of virtual goods, including stickers and emojis, in contrast to Kuaishou, which makes more than two thirds by collecting a portion of the tips that viewers give to live streamers, so-called "live-stream gifting" (Economist 2021). To quickly enter foreign markets, ByteDance acquired in 2017 the fast-growing competing short-video sharing app Musically, with its 200 million global subscribers for \$1 billion. Thus, the company engaged in a global growth strategy utilizing its AI technology by developing global products that utilized localized content. Since the Musically acquisition brought customers mainly from the U.S.A. and Europe, the firm rebranded it TikTok to compete in global markets, thus differentiating it from its Chinese sister app, which retained the Douyin brand. TikTok used the advanced AI technology developed by Douyin to build global creation and interaction platforms but emphasized a localization strategy, utilizing customized filters, effects, and stickers specific to the needs of various markets. TikTok was able to achieve more than 2 billion downloads worldwide by mid-2020 (Economist 2021).

5. Proposition Development

5.1. AI Factory

The speed and scale of AI applications in a startup's businesses require a robust system to organize the AI-powered process. Like a traditional factory that manufactures tangible goods at scale, an AI factory consistently creates AI solutions for the company on a large scale. An AI factory integrates data, algorithms, experimentation, and software infrastructure to help the company better leverage its AI capabilities and empower its fast growth. A typical AI factory consists of 4 components: data pipeline, algorithm development, experimentation platform, and software infrastructure (Iansiti and Lakhani 2020).

5.1.1. Data Platform

The first component in the AI factory is the data pipeline, where information from the user is collected, cleaned, processed, and preserved in systematic, scalable processes. The exponential increase in the volume and diversity of data in recent years has led to fundamental advances in AI systems. Firms are able to extract high volumes of data from regular business activities in a process labeled datafication ([Leonardi and Treem 2020](#); [Schafheitle et al. 2020](#)).

Data is the fuel for AI and machine learning algorithms ([Agrawal et al. 2019](#)). An AI-powered startup can tap into the power of AI relatively easily when its business is simple, and the scale is relatively small. It does not need to process and analyze data at an extremely large scale or deal with problems from siloed and complex organizational structures. However, as the AI-powered start-up continues growing, it needs to develop a systematic and scalable way to clean and normalize data in order to maximize its value. In addition to the multiple benefits of the data platform, it also presents potential risks. The continued development of AI and its requirement for large datasets raise cybersecurity concerns. Due to digital amplification, the potential damage of cybersecurity, such as data breaches, could be highly significant. Hence, the data platform should be designed not only to manage the access and processing of data, but also to tackle data security issues.

Back in December 2014 ByteDance's first product, Toutiao had only 1 million daily active users ([Han 2019](#)). The main focus at that time was on product management and operations. There were only a few data engineers focusing on data processing, analysis, and recommendation algorithms. As the business grew, the small team of data engineers was no longer sufficient to support a more extensive business scale. Driven by the data-centric mindset, a large variety of data demands has emerged from different business units such as business operation, product management, and customer relationship management ([Wang 2017](#)).

As ByteDance expanded its business scope, data started to flow from multiple sources in an enormous volume, making its collection and processing even more challenging. The data integration and further analysis become difficult if different functions or teams use their own methodology for data collection and processing. In order to better support data analysis in various teams, ByteDance developed a data portal platform aiming to offer integral data solutions to different business units by supporting different stages of the data cycle. The declarative tools provided by the data platform are user-friendly. It enables data engineers to achieve their analysis objectives without learning intricate details. The data platform significantly improves efficiency by allowing the analysts, and product managers to focus on their own analysis tasks ([Wang 2017](#)).

The data platform has multiple essential functions. The user preference and behavioral data are transmitted, stored, processed, and provided to various business units on demand. It offers a data toolset for analysts, including reporting, query, and metadata. The product management and operation team can rely on the general user behavior analysis platform to generate business insights.

5.1.2. Algorithm

The second component of the AI factory is algorithm development, when rules are established for machines to follow to generate decisions, predictions, or answer certain problems. Without any human input, algorithms can thus autonomously create individual solutions for a specific user, and when learning capabilities are embedded, by repeatedly going through the problem-solving loops, solutions improve over time ([Verganti et al. 2020](#)). Humans are required only to establish the product offering and to develop the problem-solving loop, but then technology takes over. An advantage is that algorithms can easily be scaled without redesign and can deliver multiple solutions without the need for further sizable investments. Algorithm development is critical for generating business insights and for facilitating decision-making. However, algorithmic biases such as selection bias and labeling bias can lead to flawed decision-making and raise the level of risk for

the company. Although it is impossible to remove these biases, an AI-powered company should seek to reduce data risk.

ByteDance's deployment of AI and machine learning was ahead of the industry peers. When the IT industry was still focusing on recruiting technicians, the firm had already recruited numerous machine learning and AI algorithm talents from its inception (Tan 2019). When ByteDance launched its first product Toutiao, there were only 50 employees and half of them are algorithm technicians (Li 2013). They have laid the foundation early on for the later success of the recommendation algorithm system. The algorithm platform provides the most basic algorithm recommendation technology for each of ByteDance's product lines.

5.1.3. Experimentation

Forecasts generated by AI algorithms are validated in the experimentation platform with AI-powered firms, like LinkedIn running forty thousand experiments each year, while Google runs more than a hundred thousand per year. Randomized control trials are undertaken to determine if algorithm predictions indeed have a causal effect on the outcome. The platform is fully automated thus allowing firms to run experiments at scale (Thomke 2020).

An AI-powered startup does not just collect and analyze data, it runs experiments to generate actionable decisions. Well-executed and consistent A/B testing is one of the secrets behind ByteDance's remarkable growth. The insights generated are always tested to find the optimized solution. Tools such as A/B testing are utilized to determine the best solutions for user experience improvement, and new product development. From the selection of the name "Toutiao" to the user interface of Douyin, ByteDance continuously runs numerous experiments relying on its experiment platform.

ByteDance also uses experimentation to improve its core algorithms. Recommendation algorithms drive the exponential user growth of ByteDance's flagship apps Douyin and Toutiao. There is no universal algorithm model that can be applied to all recommendation scenarios. Therefore, ByteDance also developed a flexible algorithm experimentation platform to do rapid experiments on various algorithms and different complex combinations. ByteDance apps like Xigua, Huoshan, and Douyin utilize the same set of basic algorithm systems as Toutiao, but each differs in detail.

5.1.4. Software Infrastructure

Data platform, algorithm development, and experimentation all require the support of robust and effective software infrastructure. To align with their technical ideas and better support data demand from the frontend, ByteDance initiated the creation of its Infrastructure 2.0 in 2018 (Geng 2020). This new infrastructure has a deeper connection with its business and offers architecture support for rapid business development. To build infrastructure 2.0, ByteDance made an adjustment in its organizational structure, thus integrating the online and offline infrastructure into a team. The integration offers three major infrastructure components spanning offline and online storage, computing, and R&D systems. It serves as the common base supporting all ByteDance product lines such as Toutiao, Douyin, and Feishu.

The deployment of pooled data storage helped the company unify the underlying implementation and benefits the upper-level systems. ByteDance also integrated online and offline computing systems at the computation level and created an integrated resource system to offer on-demand resources. The improvement of the R&D system helped realize the benefits of unified resource deployment. More importantly, the new infrastructure optimized the collaboration process, enhancing the synchronization of information and facilitating collaboration within the whole company (Geng 2020).

Efficiently applying AI into every aspect of the business at scale requires a robust and scalable approach. The AI factory is such an approach that organizes firm digital processes

and facilitates the continued development of new products leveraging its AI capability. We propose:

Proposition 1. *AI-powered startups tend to adopt the AI factory approach to achieve rapid growth.*

5.2. Scaling Growth of AI-Powered Startups

An AI factory is characterized by two virtuous cycles which are created to drive both the scale and scope of the firm. The more data the company collects to train the model, the better the algorithms are for improving the customer experience, which in turn leads to more usage and more data. This virtuous cycle accelerates user growth and scale expansion. As more data is accumulated, the algorithms are better at generating useful business insights, which are particularly helpful for the company to identify new opportunities and thus expand its business scope to enter new markets. As the business scope expands, a wide variety of businesses result in high data volume as well as variety. An effective AI factory integrates, cleans, and processes data from different sources via its data platform to initiate another turn of the cycle. See our model in Figure 1.

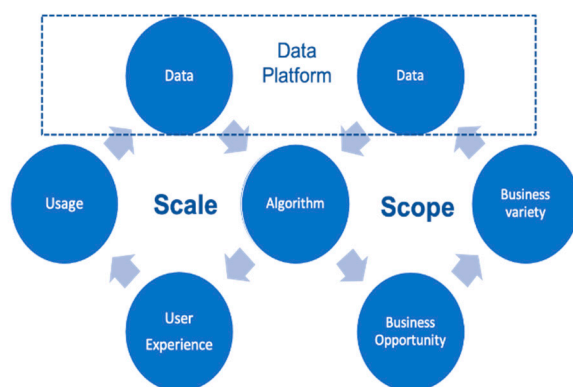


Figure 1. Virtuous growth cycle of AI.

5.2.1. Unleashing Growth Potential: AI-Powered Scale Expansion

ByteDance is known for its fast product iteration. Since its inception in 2012, ByteDance has launched over 21 popular products (Li 2020). In China, it received the nickname: “app factory”, as the rapid production of apps has created a large customer base and enabled the accumulation of a considerable volume of data. Among all of ByteDance’s products, the most well-known apps in China are Toutiao and Douyin, while globally it is TikTok.

Toutiao was launched in 2012 as the firm’s first product. Leveraging its AI recommendation algorithms, Toutiao integrates news from online resources and delivers daily personalized content to each of its users based on their interests. Toutiao had attracted approximately 10 million users after only 90 days of launch (CGTN 2014). It is worth noting that Toutiao did not have a reporter or editor. As the founder Zhang Yiming emphasized to investors, Toutiao was an AI company in the search business instead of a news company. In this way, Toutiao was founded on AI algorithms instead of human labor, making it highly scalable, as there is nearly zero marginal cost associated with an additional user.

Most importantly, the AI factory creates a virtuous cycle that reinforces user growth. Toutiao generates revenue by incorporating advertisements in the news delivery. As the number of users increases, more data is accumulated, and algorithms better predict user preferences. As a result, the recommendation algorithm is enhanced, and customer experience is improved, which leads to more users and more data, thus attracting more advertisers. This reinforcing loop allowed Toutiao to achieve amazing growth in scale. In August 2019, Toutiao had 115 million daily active users (DAU) in China (Song 2019).

Eight years after its establishment, ByteDance has achieved massive scale at an incredible speed. It has grown into one of the most globalized internet companies in China. Leveraging on its success in China, ByteDance launched corresponding versions of the

same products overseas. Its global expansion in news and information platforms started from emerging markets like India and Indonesia and gradually developed into the European and American markets. In 2015, TopBuzz, Toutiao's overseas version, was born, while in 2016, ByteDance invested in India's largest news aggregation platform-Dailyhunt, and Indonesia news recommendation reading platform BABE. In 2017 it acquired News Republic, a global mobile news aggregator. Similarly, ByteDance started its internationalization in the short video field in 2016 following the massive success of Douyin in China. A number of overseas versions of short video apps such as TopBuzz Video, TikTok, Vigo video, and Helo were born after 2016 (Yang and Wang 2020). In the short video field, ByteDance took advantage of TikTok's global impact to add users on a massive scale. In contrast, TopBuzz Video, Helo, and other short video products are designed for meeting local, personalized needs in regional markets.

ByteDance's speedy internationalization journey is powered by its AI factory. On the one hand, the core AI algorithms for recommendation and personalization are standard and can be easily applied to overseas products. At ByteDance, the existing modules can be re-used to reduce the development cost when developing new apps. On the other hand, the AI factory also enhances the effectiveness of decision-making in the internationalization process. Digitalization enables the company to learn about foreign markets faster and more effectively (Clark et al. 2018). ByteDance gained access to information and data on local markets to power its AI engine, thus significantly enhancing its internationalization speed through investment and acquisition. Thus, we propose:

Proposition 2a. *The AI factory enhances the scale expansion of an AI-powered company.*

5.2.2. Breaking Business Boundaries: AI-Powered Scope Expansion

An AI factory offers a common set of building blocks that can be deployed in different businesses and drive the company's scope. After developing a huge user base from Toutiao and Douyin, ByteDance rapidly expanded its business beyond news distribution and short video to education, healthcare, enterprise services, etc. Following its outstanding performance in the B2C market, ByteDance began to enter the B2B market. In 2019, ByteDance officially launched its enterprise messaging and productivity app, Feishu, which was intended as an internal collaboration tool (Feng 2020). In 2020, its overseas version Lark was born. Moreover, in 2020 ByteDance launched a technological service platform, "Volcano Engine," to offer technical products and solutions for business customers. Volcano Engine aims to empower external business partners with technical and AI capabilities such as big data, AI, recommendation algorithms, smart videos, and growth concepts and operating tools accumulated by ByteDance in the past. Based on the effective development of the AI factory for the B2C arena, the firm was able to successfully extend its product offering to the B2B market. Therefore, we propose:

Proposition 2b. *The AI factory enhances the scope expansion of an AI-powered startup.*

5.2.3. Quick Decision and Fast-Moving: Learning Enhancement

AI-powered startups are continuously learning from customers through experimentation and through collaboration. AI factory enhances quick and effective learning by enabling quick decisions and fast-moving adaptation. An AI Factory enhances the company's effectiveness and efficiency of learning from customers. Traditionally, companies make extensive efforts to learn from their customers. Human labor is widely involved in customer research, by sending out surveys, using focus groups, reading social media comments, etc. AI and machine learning significantly improve the efficiency of learning from customers, and AI also creates a virtuous learning cycle. When there is more data available, the AI and machine learning algorithms are better trained to discover helpful insights for predicting user preference and improving customer experience. As a result,

more users are attracted, thus more data is available for learning and improvement of the AI algorithms.

Organizations that have a high potential to innovate appear to make a habit of learning from experimentation. This activity helps organizations gather valuable information and respond to evolving contexts and new challenges. ByteDance has a robust experimentation platform that allows the company to run hundreds of experiments every day to learn about new product opportunities and risks. Learning from experimentation can also be viewed as a virtuous cycle rather than a linear process. Generally speaking, a learning cycle of experimentation might involve generating a hypothesis, gathering data, and conducting tests, distilling results into critical insights, and then adapting for the next iteration—a new cycle of experimentation. This learning cycle of experimentation fosters the culture that recognizes the value of failure and encourages firms to “fail fast and fail forward.”¹ Even when experiments do not support a proposed solution, the insights generated from the failure are valuable to organizational dynamics.

AI algorithms are better trained when a large volume and variety of data is available. However, the problem of data islands has widely existed. Due to user privacy, commercial secrets, laws, and regulations, and for various other reasons, some organizations cannot integrate data together to train a larger and better model. ByteDance Federated Learning Technology Team open-sourced its self-developed federated learning platform Fedlearner in 2020 (Cai 2020). Federated learning allows joint modeling of data distributed between institutions hence enhances AI algorithms without sharing data. It enables the improvements of AI and of machine learning algorithms and maximizes the value of data on the basis of ensuring data privacy as well as legal compliance. We propose:

Proposition 2c. *The AI factory enhances the learning of an AI-powered startup.*

5.3. Building the AI Factory

Prior literature suggests that realizing the full benefit of IT investments requires substantial investments in organizational capabilities and the redesign of the organizational structure (Bresnahan et al. 2002). Bresnahan and colleagues argue that new organizational assets are needed to complement the adoption of IT capital. The requirements for building and running an AI factory go far beyond making certain complementary organizational investments. The AI factory represents a whole new way of working. It requires the AI-powered company to be built around an integrated and modular foundation rather than a variety of siloed processes (Iansiti and Lakhani 2020). As such, it requires the organization to be both stable and agile, effective, and efficient.

5.3.1. Effective Organizational Architecture

A data-centric technological infrastructure and a strong capability foundation are required to build an effective organizational architecture. A strong emphasis on data and artificial intelligence demands centralization and consistency. A large volume of information becomes data assets only when it is organized and interpreted through a data-centric architecture. Such a data-centric infrastructure puts data at the core of the company and avoids the problem of siloed information. With a data-centric design, the functionalities that were historically handled by individual applications are instead integrated and managed at the company level.

Data-centric infrastructure and standard processes are critical when the organization advances in its AI implementation efforts. ByteDance developed a data-centric infrastructure by designing its data platform in a modular way that avoids siloed structure. A data asset is therefore integrated through the spectrum of apps to realize the benefits of the AI factory.

In order to further consolidate the capability foundation for AI algorithm development, ByteDance established its artificial intelligence laboratory (AI Lab) in March 2016. The AI Lab’s research areas include machine learning, natural language processing, computer

vision, human–computer interaction, and other cutting-edge technologies (Li 2020). The algorithms are applied to ByteDance products to promote continuous optimization of user experience in terms of communication between humans and information. The AI lab plays a critical role in the virtuous cycle of the AI factory through the improvement of algorithms and ensuring their implementation. Unlike traditional research teams focusing on basic research, ByteDance’s AI Lab establishes a team that integrates technology development and implementation. It has both the research capability, as well as the ability to ensure successful technology implementation. As such, we propose:

Proposition 3a. *Effective AI-powered startup organizational architecture builds a foundation for the AI factory.*

5.3.2. Creating an Agile Organizational Culture

Generating and delivering AI solutions at scale is inherently uncertain, thus requiring rapid experimentation and the involvement of almost the entire organization. As a result, an agile culture is essential to an AI-powered startup, improves flexibility, drives effective communication, and helps cultivate an agile, product-focused mentality. An agile organizational culture helps build a roadmap which ensures that the AI factory will frequently release successful products.

A flat organizational structure promotes effective communication and information transparency. It allows more efficient decision-making and gives the team great freedom and flexibility to innovate. For instance, the insights generated through data analysis need to be verified by extensive experimentations to form a business decision. If each decision requires going through multiple layers of management approval, a project cannot move forward quickly, and mistakes are not easily identified and corrected.

ByteDance does not have a CFO, CMO, or CTO (Cornerstone 2020). Fourteen executives report directly to CEO Zhang Yiming, covering all the functions. Project tasks are driven and completed by the feature team. It is rare at ByteDance for one person to be in a fixed position for a long time. The firm’s resources follow product development. There are only three or four levels of reporting relationships under the flat organizational management, and decisions are made fast. The whole company shares technology, including the internationalization, commercialization, audit, and sales teams. A middle platform and backend infrastructure support frontend app businesses. To further ensure information transparency and the flow of information, the company undermines hierarchies. Employees at ByteDance are asked to address each other by name instead of their job title. By weakening the management level and position, information can flow smoothly within the organization.

An agile organizational culture aligns with ByteDance’s values. “Always day 1” is one of the five values of the company (ByteDance 2021). That is to say, no matter how fast the company grows and how high its market value is, it must be as fast and flexible as they were on their first day of business, and continuously iterate. This culture emphasizes agile processes that go hand in hand with an AI-powered operating model. At ByteDance, communication is highly efficient, with teams collaborating and working in a rapid and agile fashion. For instance, context is required in project collaboration, so that the groups involved can know enough about the project backgrounds to be able to move fast.

ByteDance sets bi-monthly objectives and key targets, which are highly compatible with an agile culture that encourages trial and error. Its promotion of “fail forward and fail fast,” allows employees to keep trying in order to achieve their goals, allowing them to obtain sufficient resources and support in exploring new directions. For instance, any new idea or insights can be proposed and tested without a lengthy approval process. Therefore, we propose:

Proposition 3b. *AI-powered startups tend to create agile organizations to support their AI factory.*

5.4. VCs Role in AI-Powered Startup Data Governance

5.4.1. Ethical and Legal Issues Faced by AI-Powered Startups

The challenges created by the impact of AI on society become increasingly critical, potentially leading to harmful impacts both physically, when autonomous robots or cars operate in the real world, as well as non-physically when privacy violations or financial loss ensue (Mannes 2020). AI's legal and ethical issues may become the bottleneck of an AI-driven company and make the company susceptible to sudden failure. AI-powered firms such as Facebook, Google, and Amazon have been accused by U.S. government regulators of spreading misinformation and misleading information, especially in regards to anti-vaccination propaganda targeting measles as well as COVID-19 vaccines. Chinese officials made similar accusations against Baidu, claiming that AI algorithms do not distinguish between factual information and false statements.

AI-powered firms, where data, analytics, and AI are coupled with a high level of network connectivity, present specific ethical and legal issues that raise the firm's level of data risk. Various organizational stakeholders face increased risk from: (1) Digital amplification—algorithms can provide content that unintentionally reinforces biases, providing echo chambers, where dissenting information is ignored, potentially targeting vast populations with political, social, and health misinformation (Iansiti and Lakhani 2020). (2) Algorithmic bias—the quality of the data input into algorithms determines the quality of output (Bolander 2019). In selection bias, the input data is not representative for the population and may lead to flawed decisions, while in labeling bias, gender or race discrimination is at times pervasive, with the algorithms amplifying such biases (Caliskan et al. 2017; Kellogg et al. 2020). (3) Cybersecurity—AI-powered firms accrue massive amounts of data to power their processes, becoming targets for actors that engage in attacks resulting in data breaches, such as that faced by Equifax in 2017, where sensitive information for almost half of the U.S. population was exposed (Economist 2017). Russian-backed digital hijacking attempts sought to influence and interfere with political campaigns in both the UK and U.S.A. (4) Platform control—data-rich AI-powered firms must control access to their users' information, seeking to prevent issues like the Cambridge Analytica's unauthorized access to Facebook users' psychological information, later used in political campaigns, such as UK's Brexit and the U.S.' 2016 presidential election (Economist 2018). (5) Access to customers on Apple and Android devices are controlled by Apple and Google, respectively, with the former charging firms 30% of all in-app purchases, a strategy challenged by many developers such as Spotify and Epic Games, who tried to bypass the firm's App Store. The two companies act as gatekeepers, thus raising ethical issues due to the control they have over access to their customers' devices and information.

To answer the multiple legal and ethical issues faced by AI-powered startups, firms need to focus on developing strong in-house data governance processes, as well as accessing help from outside stakeholders to strengthen their digital performance. Governments are powerful stakeholders, who set the rules of the game for AI-powered firms. In order to minimize political risk, strong engagement with various levels of government is beneficial, taking into account the varying demands placed on data collection, storage, and sharing in different parts of the world (Jia and Ruan 2020). Typically, the EU and U.S. place much stricter regulations for AI-powered firms when compared to the rest of the world including China. In addition, VCs represent a powerful stakeholder for AI-powered startups, as they can provide not only the financial resources required to invest in growing the firm, but also expertise in managing political risk, upgrading managerial capabilities, and access to technological know-how (Gorman and Sahlman 1989).

5.4.2. ByteDance Data Factory Development with VC Backing

ByteDance has multiple VC co-investors that financed the development of the firm's AI factory. Almost 40% of the firm is held by US-based VCs including General Atlantic, Sequoia Capital, GGV Capital, KKR, in addition to the non-US VCs including Japan's Soft

Bank, and China's SIG China and Source Code Capital. The founder Zhang Yiming controls 25% and employees hold around 20% of the firm (Rolfe et al. 2020a).

Sequoia Capital, one of Silicon Valley's largest VCs, provides not only financial resources to ByteDance, but a rich experience having made early investments in some of today's tech titans, including AI-powered firms: Google, Apple, Uber, WhatsApp, and Cisco (Rolfe 2015). In addition to the typical benefits that VCs bring to the table, such as strategy analysis and development, management recruitment, and investment, the association with such a prestigious VC represents a powerful signal and vote of confidence to other potential co-investors. General Atlantic, another U.S.-based VC brings its own AI factory expertise having made prior investments in some of the world's largest AI-powered firms: Alibaba, Facebook, and Slack (Gottfried 2019). Japanese VC investor, Soft Bank brought their experience from their portfolio which includes successful investments in its own AI-powered firms: DoorDash, Uber, and Alibaba. Having VC backers with extensive experience investing in AI-powered firms, represents an advantage for ByteDance. Such VCs can contribute much more than financial resources, as they can finetune their firm strategy consulting, based on their extensive expertise with highly successful AI-powered firms from their portfolio, especially in the area of data governance.

ByteDance faced a major threat in the U.S.A. in 2020 due to concerns raised by the Trump administration regarding data collection and sharing policies, being temporarily forced to divest of its U.S. operations. Sequoia Capital and General Atlantic, two of the major U.S.-based VC investors in ByteDance, sought to intermediate a deal with Oracle, in order to prevent a ban in the U.S.A., rather than be excluded from an alternative deal proposed by Microsoft (Rolfe et al. 2020b). This intermediation was an attempt to reduce the political backlash against the firm, as well as an action intended to strengthen the position of the two VCs, which had a high stake in ByteDance. However, as of March 2021, the Biden administration had postponed this prior ruling indefinitely, making the VC-backed sale unnecessary (McKinnon and Leary 2021). Therefore, we propose:

Proposition 4a. *Venture capital investments help establish and strengthen data governance, and aid in minimizing data risk.*

5.5. ByteDance as a CVC Investor

AI technology is constantly evolving, requiring continuous learning both internally as well as externally. AI-powered firms can engage in internal corporate venturing, by developing knowledge within the boundaries of the firm, as well as by engaging in corporate venture capital (CVC), which represents boundary spanning operations, when the firm participates in external corporate venturing, by making minority investments in high tech small ventures (Maula 2007). CVC programs resemble those of traditional venture capital firms, with the focal firm investing in independent external companies. AI-powered firms engage in CVC either for financial gains, but especially for strategic objectives, such as (1) Learning about the market, new technologies and to access external R&D; (2) Options to acquire companies or to expand into new markets; and (3) Leverage their own AI-platform, and to add new products to the platform, as well as to leverage their human resources (Maula 2007). AI-powered startup CVC investors benefit new ventures by bringing financial but especially nonfinancial resources, such as AI and machine learning expertise, as well as other complementary assets and industry experience (Park and Steensma 2012).

As ByteDance engages in CVC activities, by investing in other firms, they can either follow the lead of other VCs, or they can rely on their own expertise in machine learning to analyze potential opportunities and gauge the risk of investments (Arroyo et al. 2019). Corporate investors utilize CVC investments as a window to new technologies and to learn from the new ventures in which they invested; hence improving their own innovation outcomes and performance (Dushnitsky and Lenox 2005; Wadhwa and Kotha 2006). Exposure to new technologies can increase firm absorptive capacity, its ability to appreciate the value

of acquiring, assimilating, and applying novel external knowledge, thus improving the firm's own internal R&D productivity (Cohen and Levinthal 1990; Benson and Ziedonis 2009). Firms can expand their learning by acquiring promising firms, as ByteDance did in 2020 when it bought LevelupAI, an AI-powered startup in the video game industry; however, this may prove a riskier and costly endeavor. Alternatively, CVC investments seek to reduce risk, while decreasing the amount of investment and increasing the potential learning effects. ByteDance has made CVC investments in IReadyIT, an AI-powered startup that seeks to provide a green industrial internet platform, as well as Lingxi, an AI-powered fintech firm which uses machine intelligence to augment human capabilities (Liao 2020). Insights into AI in the fintech industry are valuable for ByteDance as the firm has launched in 2021 Douyin Pay, a mobile payment system offered to the platform's almost 700 million Chinese users, who had previously used Tencent's WeChat Pay or Ant Group's Alipay when making payments on Douyin (Marketline 2021).

CVC investments present two important options to firms: either to acquire the portfolio company later, if the firm proves valuable, or to potentially enter new markets, in which case the CVC investment is used as a probe. The identification and assessment of targets to be potentially acquired are one of the goals of AI-powered firms' investments. Alternatively, when seeking to enter different industry sectors, unfamiliar to the AI-powered firm, a CVC probing investment may be pursued in order to develop the necessary skills and ascertain proper market timing. Thus, the investment will allow the firm to hedge its bets and develop stakes in emerging AI technologies before a dominant design develops (Keil 2000). ByteDance joined in 2020 two of its VC investors, Sequoia Capital and Hillhouse Capital as co-investors into Narwal Robotics, in order to develop knowledge about the industry. ByteDance again followed Sequoia Capital into a CVC investment in JuLive, a real-estate e-commerce platform. Furthermore, to gain information regarding the electric vehicle market, ByteDance pursued a CVC investment into the auto manufacturer Lixiang in 2019. The previously discussed investment in Lingxi, an AI-powered fintech firm specializing in insurance sales and debt collection, offers an insight into machine learning in the financial industry, which may be further expanded as ByteDance holds an insurance broker license (Liao 2020).

AI-powered startups may engage in CVC to encourage demand for their existing services, by investing in firms that use their platform. A goal may be to proactively shape markets, seeking to establish standards around their platform, by engaging into venture investment into favorable companies (Maula 2007).

ByteDance can use its CVC investments as windows into technologies that may allow it to enter adjacent industries, developing into a network hub, as WeChat and Baidu have done. Network hubs link consumers with companies and whole industries (Iansiti and Lakhani 2020). Being able to incorporate the almost 700 million daily users from the short video app Douyin and the news app Toutiao into a network hub, using AI to link to other industries where ByteDance has CVC investments, would allow the firm to compete with its main Chinese and international rivals. Therefore, we propose:

Proposition 4b. *AI-powered startups tend to engage in CVC to amplify their AI capabilities.*

6. Discussion

6.1. Contributions

Our study makes four contributions to extant literature. First, this study complements current digital entrepreneurship literature (Nambisan 2017; Nambisan and Baron 2019) by investigating how AI application can empower the high growth rate of startups. Applying AI into business operations is a foreseeable and likely unavoidable trend. However, AI has not yet been widely adopted at scale in business processes (Ransbotham et al. 2018), as most companies apply AI only to solve discrete problems (Fountain et al. 2019). However, realizing the full benefits of AI to power business growth and expansion is not only about applying new technology but also involves the adoption of a systematic approach. By

closely examining ByteDance's success, we show that AI-powered startups place AI at their core and apply it at scaling growth in their operations. Such firms tend to adopt the AI factory approach to speed up expansion in scale and scope and to enhance learning. These findings have significant strategic implications for startups seeking to apply AI in scaling their operations and models.

Secondly, our study brings attention to the limited literature that focuses on the relationship between AI and organizational structure (Brock and Wangenheim 2019). We emphasize the importance of organizational structure and company culture in creating and running an AI factory. Existing research on the effects of AI on company organizational structure is primarily constrained to well-established companies (Davenport and Ronanki 2018). Existing studies focus mainly on structural reform alongside the digital transformation within large firms. In contrast, our study uses the case of ByteDance to provide recommendations for startups who do not face organizational resistance challenges and offers novel insights to describe competition in the age of AI. Our analysis provides implications for agile startups to develop and design their structure to fully take advantage of AI and to challenge incumbent companies.

Third, in addition to the benefits of the AI factory, our study takes one step forward in identifying the ethical challenges and risks faced by AI-powered startups. The case of ByteDance sheds light on the role of venture capitalists in helping AI-powered startups not only by financing their growth, but by providing expertise on how to deal with data governance issues, as well as helping them manage political risk. VC experience with prior AI-powered firms can provide a valuable benefit to their AI-powered startups.

Fourth, we emphasize the role of AI-powered firms as CVC investors, and exemplify with ByteDance's multiple investments in firms involved in AI, chips, robotics, fintech, and electric car manufacturing. Such investments may be made for financial but have predominantly strategic reasons, seeking to gather knowledge that may improve the firm's own AI. As ByteDance grows into adjacent industries, such as their launch of Douyin Pay, their online payment system, CVC investments allow a window into the operations of relevant startups, which may lead to later acquisitions or potential expansion to new markets. As AI-powered startups seek to develop into network hubs, with ByteDance trying to emulate its Chinese competitor Baidu, the role of knowledge gathering regarding markets and new technologies through CVC investments, becomes critical.

6.2. Implications

This study offers significant strategic implications for entrepreneurial ventures seeking to scale their operations and models. First, using the case of ByteDance, we show how AI-powered startups leverage the AI factory approach to achieve rapid growth in scale and scope. Second, our close examination of company culture and organizational structure offers startups recommendations on how to fully take advantage of AI and challenge incumbent companies. Startups are advised to design a flat organizational structure and cultivate an agile culture to improve flexibility and drive effective communication. Third, we identify challenges and risks faced by AI-powered startups. Although data risk is impossible to remove, it is important to understand its pervasiveness and strive to reduce it.

One of the main implications for VCs is the stringent need to develop expertise in regard to the various types of data risk faced by AI powered startups in their portfolio. In addition to the usual type of resources that VCs provide their firms, such as financing, and strategy consulting, developing data risk expertise takes an increasing role for their AI powered startup investments. In the case of ByteDance, its VCs had developed prior expertise regarding data risk mitigation from their previous investments in AI powered firms.

6.3. Limitations and Future Research

Our study presents certain limitations which may lead to valuable directions for future research. First, the case study method has inherent limitations in terms of generalizability. It is challenging to elucidate the characteristics of a broad population using a case study method, especially from a single case. Therefore, future studies could expand the scope of analysis to more cases to explore other factors relevant in building the AI factory. For instance, the current research focuses on ByteDance, which evolved in the institutional environment of China. It would be interesting to examine how AI-powered startups manage AI capabilities in different institutional environments by conducting multiple case comparison studies.

Second, although there are five types of data risk in theory, the ByteDance case alone does not address the impact of VCs on all of them individually. Using data triangulation to ensure validity, we are only able to identify the role of VCs in mitigating data risk as a composite construct in the case of ByteDance. In a future study, we plan to utilize additional cases to develop a model focusing on how VCs detect and reduce specific types of data risk. This more detailed examination would further extend our understanding of the role of VCs in helping startups manage individual data risks and offer practical implications for both startups and VCs in the age of AI.

Third, due to the limitations of the case study method, focusing on the in-depth analysis of a single case, we are not able to compare and contrast the evolution of the AI factory between startups and established companies. Future research may compare AI-powered startups and firms that transitioned from traditional to AI-powered to further describe how AI-powered startups compete with incumbents by taking advantage of the AI factory from the firm establishment.

7. Conclusions

Through an in-depth analysis of ByteDance, using the case study method, our study extends both AI and entrepreneurship literature by showing how AI-powered startups leverage AI at scale to optimize scale, scope, and learning. Our discussion also emphasizes the critical role played by VCs in assisting AI-powered startups to build AI factories and in reducing data and political risk. As an early attempt to investigate AI deployment within an entrepreneurship context, this study advances our understanding of AI application in businesses. Our findings offer important practical suggestions for startups when adopting AI and seeking to become AI-powered companies.

Author Contributions: Conceptualization, P.J. and C.S.; Methodology, P.J.; Validation, P.J. and C.S.; Formal analysis, P.J. and C.S.; Investigation, P.J. and C.S.; Resources, P.J. and C.S.; Writing—original draft preparation, P.J. and C.S.; Writing—review and editing, P.J. and C.S.; Visualization, P.J.; Supervision, P.J. and C.S.; Project Administration, P.J. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

¹ From John C. Maxwell Quotes: Fail early, fail often, but always fail forward.

References

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2019. Economic policy for artificial intelligence. *Innovation Policy and the Economy* 19: 139–59. [[CrossRef](#)]

- Arroyo, Javier, Francesco Corea, Guillermo Jimenez-Diaz, and Juan A. Recio-Garcia. 2019. Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments. *IEEE Access* 7: 124233–43. [CrossRef]
- Benson, David, and Rosemarie H. Ziedonis. 2009. Corporate venture capital as a window on new technologies: Implications for the performance of corporate investors when acquiring startups. *Organization Science* 20: 329–51. [CrossRef]
- Bolander, Thomas. 2019. What do we lose when machines take the decisions. *Journal of Management and Governance* 23: 849–67. [CrossRef]
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* 117: 339–76. [CrossRef]
- Brock, Jürgen Kai-Uwe, and Florian Von Wangenheim. 2019. Demystifying AI: What digital transformation leaders can teach you about realistic artificial intelligence. *California Management Review* 61: 110–34. [CrossRef]
- Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2018. Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In *The economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press, pp. 23–57.
- Byford, Sam. 2018. How China's ByteDance Became the World's Most Valuable Startup. Available online: <https://www.theverge.com/2018/11/30/18107732/ByteDance-valuation-tiktok-china-startup> (accessed on 15 March 2021).
- ByteDance. 2020. Hear from Our Engineers at ByteDance Singapore [Video]. Available online: <https://www.youtube.com/watch?v=pweTBZnNqlk> (accessed on 15 March 2021).
- ByteDance. 2021 Explaining ByteStyle: Always Day 1 [Video]. Available online: https://www.youtube.com/watch?v=G20bn_sdhbA (accessed on 15 March 2021).
- Cai, Fangfang. 2020. ByteDance Breaks Federal Learning: Open Source Fedlearner Framework, 209% Increase in Advertising Efficiency. Available online: <https://www.jiqizhixin.com/articles/2020-11-03-9> (accessed on 15 March 2021).
- Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* 356: 183–86. [CrossRef] [PubMed]
- CGTN. 2014. Colors of China: The Story of App Toutiao. Available online: <https://america.cgtn.com/2014/10/06/colors-of-china-the-story-of-app-toutiao> (accessed on 15 March 2021).
- Chen, Junfeng. 2020. ByteDance, an Algorithm-Driven Approach to Media Advancement. Available online: <https://www.jingpt.com/article/41663b6b-3dbb-4efe-bfd0-86a5c947b5c8> (accessed on 15 March 2021).
- Chen, Meng. 2019. Lonely Tencent, Dancing Byte. Available online: <http://www.invest-data.com/eWebEditor/uploadfile/2019071709044959153424.pdf> (accessed on 15 March 2021).
- Chesbrough, Henry W. 2002. Making sense of corporate venture capital. *Harvard Business Review* 80: 90–99. [PubMed]
- Clark, Daniel, Dan Li, and Dean Shepherd. 2018. Country familiarity in the initial stage of foreign market selection. *Journal of International Business Studies* 49: 442–72. [CrossRef]
- Cohen, Wesley M., and Daniel A. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–52. [CrossRef]
- Corbin, Juliet, and Anselm Strauss. 2008. Strategies for qualitative data analysis. *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory* 3. [CrossRef]
- Cornerstone. 2020. Why Is ByteDance “No Management at All”? Available online: <https://segmentfault.com/a/1190000023056271> (accessed on 18 March 2021).
- Davenport, Thomas H., and Rajeev Ronanki. 2018. Artificial intelligence for the real world. *Harvard Business Review* 96: 108–16.
- Dushnitsky, Gary, and Michael J. Lenox. 2005. When do incumbents learn from entrepreneurial ventures. *Research Policy* 34: 615–39. [CrossRef]
- Dyer, W. Gibb, Jr., and Alan L. Wilkins. 1991. Better stories, not better constructs, to generate better theory: A rejoinder to Eisenhardt. *The Academy of Management Review* 16: 613–19. [CrossRef]
- Economist. 2017. Once More into the Breach. The Big Data Breach Suffered by Equifax Has Alarming Implications. Available online: <https://www.economist.com/finance-and-economics/2017/09/16/the-big-data-breach-suffered-by-equifax-has-alarming-implications> (accessed on 18 March 2021).
- Economist. 2018. Is Democracy Safe in the Age of Big Data? Available online: <https://www.economist.com/open-future/2018/08/25/is-democracy-safe-in-the-age-of-big-data> (accessed on 18 March 2021).
- Economist. 2021. Kuaishou's Shares Surge on the Video App's Stockmarket Debut. Available online: <https://www.economist.com/business/2021/02/05/kuaishous-shares-surge-on-the-video-apps-stockmarket-debut> (accessed on 18 March 2021).
- Feng, Coco. 2020. ByteDance Launches Dedicated Cloud Documents App for Feishu as TikTok Continues to Face Headwinds Abroad. Available online: <https://www.scmp.com/tech/big-tech/article/3110389/bytedance-launches-dedicated-cloud-documents-app-feishu-tiktok> (accessed on 18 March 2021).
- Fountaine, Tim, Brian McCarthy, and Tamim Saleh. 2019. Building the AI-powered organization. *Harvard Business Review* 97: 62–73.
- Geng, Chen. 2020. App Factory in the Age of Algorithm: The In-Depth Research Report of ByteDance. Available online: http://pdf.dcfw.com/pdf/H3_AP202001131373897373_1.pdf (accessed on 18 March 2021).
- Gorman, Michael, and William A. Sahlman. 1989. What do venture capitalists do? *Journal of business venturing* 4: 231–48. [CrossRef]
- Gottfried, Miriam. 2019. Banking and Finance: General Atlantic taps 3 presidents. *Wall Street Journal*, November 6.
- Han, Jasper. 2019. ByteDance Story 2: Get 1M DAU in 4 Months. Available online: <https://medium.com/chinese-unicorns-through-the-looking-glass/bytedance-story-2-get-1m-dau-in-4-months-157c63d7853d> (accessed on 15 March 2021).

- Hsu, Dan, Michael Haynie, Sharon Simmons, and Alexander McKelvie. 2014. What matters, matters differently: A conjoint analysis of the decision policies of angel and venture capital investors. *Venture Capital* 16: 1–25. [CrossRef]
- Hyett, Nerida, Amanda Kenny, and Virginia Dickson-Swift. 2014. Methodology or method? A critical review of qualitative case study reports. *International Journal of Qualitative Studies on Health and Well-Being* 9: 1–12. [CrossRef]
- Iansiti, Marco, and Karim R. Lakhani. 2020. *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Boston: Harvard Business Press.
- Jia, Lianrui, and Lotus Ruan. 2020. Going global: Comparing Chinese mobile applications' data and user privacy governance at home and abroad. *Internet Policy Review* 9: 1–22. [CrossRef]
- Keil, Thomas. 2000. *External Corporate Venturing. Cognition, Speed and Capability Development*. Helsinki: Helsinki University of Technology.
- Kellogg, Katherine C., Melissa A. Valentine, and Angèle Christin. 2020. Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals* 14: 366–410. [CrossRef]
- Leonardi, Paul M., and Jeffery W. Treem. 2020. Behavioral visibility: A new paradigm for organization studies in the age of digitization, digitalization, and datafication. *Organization Studies* 41: 1601–25. [CrossRef]
- Li, Jane. 2020. The Chinese “App Factory” behind the Global Rise of TikTok. Available online: <https://qz.com/1803609> (accessed on 15 March 2021).
- Li, Jinwen. 2013. Toutiao Has More Than 50 Million Users, and It Is Trying to Import Traffic to the News Source Website. Available online: <https://www.36kr.com/p/1641762979841> (accessed on 18 March 2021).
- Liao, Rita. 2020. TikTok Parent ByteDance Leads \$6M Round in Financial Startup Lingxi. Available online: <https://techcrunch.com/2020/05/19/bytedance-invests-in-debt-collection-ai-company-lingxi> (accessed on 15 March 2021).
- Lin, Liza, and Eva Xiao. 2020. TikTok founder wonders what hit him. *Wall Street Journal*, August 27.
- Mannes, Aaron. 2020. Governance, Risk, and Artificial Intelligence. *AI Magazine* 41: 61–69. [CrossRef]
- Marketline. 2021. ByteDance Debuts “Douyin Pay” Mobile Payment Service in China. Available online: https://advantage-marketline.com/News/bytedance-debuts-douyin-pay-mobile-payment-service-in-china_32974 (accessed on 20 January 2021).
- Maula, M. 2007. Corporate venture capital as a strategic tool. *Handbook of Research on Venture Capital* 1: 371–92.
- McKinnon, John D., and Alex Leary. 2021. TikTok sale to Oracle, Walmart is shelved as Biden reviews security. *Wall Street Journal*, February 10.
- Montes, Gabriel Axel, and Ben Goertzel. 2019. Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change* 141: 354–58. [CrossRef]
- Nambisan, Satish, and Robert A. Baron. 2019. On the costs of digital entrepreneurship: Role conflict, stress, and venture performance in digital platform-based ecosystems. *Journal of Business Research* 125: 520–32. [CrossRef]
- Nambisan, Satish. 2017. Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. *Entrepreneurship Theory and Practice* 41: 1029–55. [CrossRef]
- Obschonka, Martin, and David B. Audretsch. 2020. Artificial intelligence and big data in entrepreneurship: A new era has begun. *Small Business Economics* 55: 529–39. [CrossRef]
- Ozcan, Pinar, and Kathleen M. Eisenhardt. 2009. Origin of Alliance Portfolios: Entrepreneurs, Network Strategies, and Firm Performance. *The Academy of Management Journal* 52: 246–79. [CrossRef]
- Park, Haemin Dennis, and H. Kevin Steensma. 2012. When does corporate venture capital add value for new ventures. *Strategic Management Journal* 33: 1–22. [CrossRef]
- Qin, Chloe. 2020. Talent VS Mechanism? Enhance Organizational Competitiveness and Help Enterprise Develop. Available online: <https://runwise.co/innovation/applied-leadership/23958.html> (accessed on 18 March 2021).
- Ransbotham, Sam, Philipp Gerbert, Martin Reeves, David Kiron, and Michael Spira. 2018. Artificial intelligence in business gets real. *MIT Sloan Management Review*. Available online: <https://sloanreview.mit.edu/projects/artificial-intelligence-in-business-gets-real/> (accessed on 15 March 2021).
- Reuer, Jeffrey J., and Ramakrishna Devarakonda. 2017. Partner selection in R&D collaborations: Effects of affiliations with venture capitalists. *Organization Science* 28: 574–95.
- Rolfe, Winkler, Jing Yang, and Alexander Osipovich. 2020a. Trader's bet on TikTok pays off. *Wall Street Journal*, October 2.
- Rolfe, Winkler, Miriam Gottfried, and Cara Lombardo. 2020b. TikTok investors press for a sale to Oracle. *Wall Street Journal*, October 2.
- Rolfe, Winkler. 2015. The secretive club funding Silicon Valley. Scouts help venture capital giant Sequoia Capital find the next big thing. *Wall Street Journal*, November 13.
- Rosenbusch, Nina, Jan Brinckmann, and Verena Müller. 2013. Does acquiring venture capital pay off for the funded firms? A meta-analysis on the relationship between venture capital investment and funded firm financial performance. *Journal of Business Venturing* 28: 335–53. [CrossRef]
- Santos, Roberto S., and Lingling Qin. 2019. Risk Capital and Emerging Technologies: Innovation and Investment Patterns Based on Artificial Intelligence Patent Data Analysis. *Journal of Risk and Financial Management* 12: 189. [CrossRef]
- Schafheitle, Simon Daniel, Antoinette Weibel, Isabel Laura Ebert, Gabriel Kasper, Christoph Schank, and Ulrich Leicht-Deobald. 2020. No stone left unturned? Towards a framework for the impact of datafication technologies on organizational control. *Academy of Management Discoveries* 6: 455–87. [CrossRef]
- Siggelkow, Nicolaj. 2007. Persuasion with Case Studies. *The Academy of Management Journal* 50: 20–24. [CrossRef]

- Song, Jingli. 2019. ByteDance Reshuffles Management in Bid to Build a Third App with 100 Million DAU. Available online: <https://kr-asia.com/bytedance-appoints-toutiao-ceo-chen-lin-as-head-of-emerging-business-hoping-to-build-a-third-app-with-100-million-dau> (accessed on 18 March 2021).
- Tan, Xiaohan. 2019. Employee Poaching between Toutiao and Baidu. Available online: <https://tech.sina.com.cn/i/2019-01-10/doc-ihqfskcn5960166.shtml> (accessed on 18 March 2021).
- Thomke, S. 2020. *Experimentation Works: The Surprising Power of Business Experiments*. Boston: Harvard Business Press.
- Verganti, Roberto, Luca Vendraminelli, and Marco Iansiti. 2020. Innovation and Design in the Age of Artificial Intelligence. *Journal of Product Innovation Management* 37: 212–27. [CrossRef]
- Vial, Gregory, Jinglu Jiang, Tanya Giannelia, and Ann-Frances Cameron. 2021. The Data Problem Stalling AI. *MIT Sloan Management Review* 62: 47–53.
- Wadhwa, Anu, and Suresh Kotha. 2006. Knowledge creation through external venturing: Evidence from the telecommunications equipment manufacturing industry. *Academy of Management journal* 49: 819–835. [CrossRef]
- Wang, Xueyan. 2017. Wang Ye of Toutiao: This Is How Data-Driven Companies Use Data. Available online: <https://developer.51cto.com/art/201704/536607.htm> (accessed on 18 March 2021).
- Yang, Jing, and Tianyu Wang. 2020. ByteDance, an Algorithm-Backed Firm with Globalization Ambitions. Available online: <https://news.cgtn.com/news/2020-08-08/ByteDance-an-algorithm-backed-firm-with-globalization-ambitions--SNbGeeF0pq/index.html> (accessed on 18 March 2021).
- Yin, Robert K. 2009. How to do better case studies. *The SAGE Handbook of Applied Social Research Methods* 2: 254–82.
- Zider, Bob. 1998. How venture capital works. *Harvard Business Review* 76: 131–39.