PitchBook

EMERGING TECH RESEARCH

Artificial Intelligence & Machine Learning

Q4 2020

Contents

Credits & contact

RESEARCH

Brendan Burke
Senior Analyst, Emerging Technology
brendan.burke@pitchbook.com

DATA

Bailey York
Data Analyst

DESIGN

Kelilah King and **Julia Midkiff**

This Emerging Technology Research report is updated on a quarterly basis to reflect changes in venture capital deal activity and other market related updates deemed valuable by the research analyst. The previous quarterly report can be accessed **here.**

Q4 2020 news and highlights

Q4 2020 VC DEAL ACTIVITY

- Q4 2020 set a record for quarterly AI & ML VC funding with $16.5 billion invested, contributing to a new high of $52.1 billion invested in the segment in 2020.
- Deal activity in 2020 was led by healthcare, consumer, enterprise IT, and autonomous vehicles, with each category raising over $5 billion. Declines occurred in financial services, AI core, AI automation platforms, and computer vision.
- The median valuation step-up declined for a second consecutive year to 1.6x, in line with the VC market median, suggesting that valuation growth of AI startups is becoming more rational in a challenged economic environment.

Q4 2020 EXIT ACTIVITY

- VC exit value amounted to $49.6 billion in 2020, nearly matching the total VC investment, demonstrating the maturation of the AI theme.
- Pure-play enterprise AI debuted in public markets in Q4 with the listings of **Upstart** (NASDAQ: UPST) and **C3.ai** (NYSE: AI). These successful listings have set the stage for a wave of enterprise AI IPOs over the next 1–2 years.
- M&A activity was muted in 2020, with both deal value and deal count flat YoY. The pandemic has led to low exit values and tepid interest from big tech companies. We are looking for a resurgence in M&A activity in 2021.

NEWS

- **December 3, 2020: Facebook** (NASDAQ: FB) open sources a reinforcement learning model called ReBeL (Recursive Belief-based Learning) that achieves superhuman performance at imperfect-information games.
- **November 30, 2020: Google**'s **DeepMind** announced the success of its AI system AlphaFold in the Critical Assessment of Protein Structure Prediction (CASP), resolving a scientific problem related to the modeling of protein folding that had been unsolved for 50 years.
- **November 25, 2020:** South Korea's government pledged to invest $1.0 billion in AI semiconductor manufacturing with the goal of becoming a leader in AI.

TRENDS & OBSERVATIONS

In both public and private markets, leading AI-native companies are experiencing rapid valuation growth that reflects technical advantages and commercial adoption. Public markets are assigning outlier valuation multiples to companies with advanced AI capabilities, including **Tesla** (NASDAQ: TSLA), **C3.ai** (NYSE: AI), and **Palantir** (NYSE: PLTR), reflecting high growth expectations based on their competitive advantages. Exchange traded funds focusing on AI companies substantially outperformed the broader market in 2020. In private markets, horizontal platform unicorns are raising large private rounds with high valuation step-ups, including **Databricks**, **DataRobot**, **Dataiku**, **Dialpad**, and Scale AI. Overall, the median late-stage VC valuation in AI spiked from $45.0 million in 2019 to $60.0 million in 2020. We believe that after a period of uncertainty, AI's moment as a mainstream investment category has arrived.

Executive summary

Artificial intelligence (AI) is the area of computer science that focuses on creating intelligent machines that can perceive their environment and make decisions to optimize against a goal. Machine learning (ML) is a subfield of AI that aims to give computers the ability to learn iteratively, improve predictive models, and find insights from data without being explicitly programmed. In practice, AI & ML is a subfield of data science that can extract predictions from complex datasets.

We believe the commercial AI & ML market is in the first 10 years of a long-term shift in societal decision making, leveraging scientific progress made over the past century via improved computing power and enhanced datasets. As a result, we are just seeing the early growing pains of the technology in practical applications and barely gaining a glimpse into the potential of the technology to replicate human intelligence.

We estimate the AI & ML vertical will reach $111.5 billion in 2021, forecasting it to grow at a 22.1% CAGR into a $203.2 billion market in 2024. This estimate includes the following segments:

- **Horizontal platforms:** AI enablement platform-as-a-service (PaaS) layers that directly empower technical end users to make use of scientific advances in AI & ML
- **Vertical applications:** Solutions that address specific problems within industries with varying levels of AI & ML integration
- **AI & ML semiconductors:** Hardware and software that optimize semiconductors for computational efficiency and customized AI calculations
- **Autonomous machines:** Hardware systems capable of performing tasks in human-present environments without explicit human control

Each of these segments has separate but complementary markets to the others. This excludes services and non-semiconductor compute, including servers and storage, which we do not view as opportunities for startups. As a result, this estimate is lower than other market forecasts.

AI-first companies starting to achieve outsized exits on public markets: We believe technological innovation in AI & ML over the past decade is starting to produce customer outcomes that can build large companies. In Q4, **C3.ai** (AIaaS) and **Upstart** (financial services) listed on public markets with high growth and explicit disclosure of their AI architectures. AI & ML are becoming more common underlying technologies for tech unicorns, requiring investors to understand the level of competitive differentiation established by the technology.

Enterprise IT automation and AI in healthcare driving VC and M&A activity: These AI applications are maintaining strength through the pandemic as leading targets for VC and strategic investors. IT automation includes information security, sales & marketing, and HR applications. In healthcare, breakthroughs in machine learning effectiveness for diagnostics and drug discovery have caused rapid acceleration in VC activity. Generally, AI & ML startups can compound the effectiveness of their offerings through persistent learning, creating sustainable competitive advantages in categories where incumbents are underutilizing or not collecting data. We believe that investors must apply an AI-specific framework to evaluate opportunities in the segment.

Key takeaways

GPT-3 emerging as a platform for NLP startups. Natural language technology, including NLP, has caught up with computer vision in VC funding in 2020 after lagging the field historically. In parallel, the recent release of GPT-3 is leading to a new wave of startups leveraging **OpenAI**'s massive neural network via its distribution partner **Microsoft**. Independent research since GPT-3's launch has found that GPT-3 can perform well in text generation, code completion, and simple knowledge queries. Viable applications include email composition, digital avatar chatbots, copywriting generation, natural language database queries, and code completion. GPT-3 can thus be thought of as a computing platform upon which a new wave of enterprise and consumer applications can be built. The exponential gains in NLP's resemblance to human intelligence demonstrates the path forward for artificial general intelligence, although that breakthrough is not yet in sight. We believe the rapid innovation in the field will lead to natural language technology receiving the highest funding of horizontal platform categories in the near-term.

Winner-take-most opportunities to emerge in sales & marketing data integration, SMB fraud prevention, predictive maintenance, and core banking PaaS, among other fields: We believe the network effects of AI & ML can produce sustainable competitive advantages when predictions improve with more data, proprietary feature sets are generated, and emerging datasets are applied to more complex problems than incumbents can solve. With startups' ability to pull in data from multiple applications and customers and use it to develop proprietary features, we believe startups can build flywheels in large industries. We view banking, sales & marketing, predictive maintenance, and small- to medium-sized business (SMB) fraud prevention as areas where incumbents may not be focused and may be limited in their ability to develop feature sets that are extensible across the entire industry. The result of creating sustainable competitive advantages in these spaces should be high valuation premiums in public markets or strategic acquisitions. Other verticals and product categories may face barriers to entry that AI-first approaches may not be able to overcome.

Pressure for incumbent chipmakers to acquire FPGA, ASIC and distributed memory startups: We believe the dominant leaders in graphics processing units, central processing units, and field-programmable gate arrays (FPGAs) lack some of the necessary innovations in distributed memory to cost-effectively power an AI revolution. They will need to seek innovation externally to obtain market leadership. Our thesis is playing out with **Nvidia**'s announced acquisition of AI microcontroller design firm **Arm** and **AMD**'s announced acquisition of FPGA leader **Xilinx**, which suggest that incumbents are building diversified platforms of AI chips to address emerging customer needs. **Intel** has directly taken on **Nvidia** in AI processing with its $2.0 billion purchase of **Habana Labs**, creating a new battlefield in application-specific integrated circuits. Additionally, incumbents hungry for growth will look to application-specific chips that they can scale across emerging use cases including mobile robotics, 5G, and IoT devices. We have not seen AI semiconductor startups scale before selling to incumbents, making consolidation a probable outcome for companies within this market.

Multiple trends converging to stimulate deal activity in autonomous robotics: The autonomous robotics space has not been highly active in terms of venture or M&A activity, but we could see a revival in investor interest for a confluence of reasons. Factors include automation needs stemming from the pandemic, scientific improvements in AI & ML, underpenetration of robotics in non-automotive manufacturing globally, and a cyclical rebound from low manufacturing as a result of the US-China trade war. We believe a recovery will carry significant tailwinds for robotics and reveal the lagging positions of incumbents in industrial automation.

AI-first startup market opportunity

Background

The AI field is in the pilot stages of creating a new class of companies with different characteristics than the SaaS businesses that have defined the current era of tech giants. We believe that the industry is in the first 10 years of a 40-year era that can redefine multiple industries and product categories. This has coincided with a golden era of SaaS computing, though we believe the two technologies are computationally and functionally distinct. Investors have conflated the two and encouraged AI developers to commercialize their findings early, resulting in some market failures.

AI & ML has been a lucrative space for VC investors over the past seven years, belying the confusion in the market. **Previous PitchBook research** found that AI & ML has had outstanding financial returns for VC investors for companies funded between 2013 and 2018, with all series achieving an adjusted annualized return of at least 35.6%. These returns far surpass those for the technology industry more broadly, for which no series achieved over a 13.5% adjusted annualized return.

AI-first characteristics

Proprietary datasets: AI-first products derive their competitive advantages from the quality and exclusivity of their datasets. SaaS platforms claim to have data moats but often differentiate based on efficiency and user experience.

Produces decisions: SaaS typically produces efficiencies that decrease the amount of time needed to make mission-critical decisions, instead of the decisions themselves. The use of AI & ML to make decisions with better accuracy than humans increases both the value and the risk of AI platforms, necessitating ethics and governance frameworks not needed for SaaS.

Expanding network effects: AI companies can leverage data more effectively than SaaS startups can. SaaS platforms run algorithms that follow explicit rules over customer-provided data and update those algorithms based on customer feedback, resulting in flimsy data moats. AI companies use their own models to learn from their customers' data and compound the quality of their predictions, producing legitimate network effects. These network effects can constitute a long-term competitive advantage.

AI companies also bear more costs than conventional SaaS startups in product development:

- **Data labeling:** Andreessen Horowitz has estimated this cost to reach up to 15% of revenue in cases where customer data must be prepared for modeling.
- **Cloud computing costs:** Andreessen Horowitz has estimated this cost to reach up to 25% of revenue for AI & ML startups.
- **Customer model integration costs :** Professional services costs required to test and deploy models.
- **Separate teams of data scientists and AI developers in addition to application developers, DevOps engineers, and IT operations teams:** AI developers are scarce and do not displace more conventional software engineers.

These costs can escalate unpredictably based on the success of model training and complexity of model deployment. The development of a single model can amount to over $1 million, making them multiple times more expensive than SaaS apps and increasing the capital requirements to seed an AI & ML startup. As a result, AI-first startups have fundamentally different unit economics and growth trajectories than SaaS businesses, and the investment community is still early in defining how to evaluate them.

Public market performance

AI-native companies are debuting on public markets with greater disclosure into their AI architectures than their predecessors offered and are using their AI capabilities as justifications for competitive differentiation. In Q4 2020, **C3.ai** (NYSE: AI), an industrial AI-as-a-service (AIaaS) unicorn, and **Upstart** (NASDAQ: UPST), an AI-powered lender, debuted on public markets at high valuations and explicit descriptions of their AI advantages in their S-1 filings. We find that this level of disclosure has surpassed that of previous IPO candidates and indicates the increasing sophistication of AI-native companies.

C3.ai offers several AI KPIs that can be used to compare the legitimacy of vendors' AI approaches. These include:

- AI predictions made per day (1.1 billion)
- Number of machine learning models in use (4.8 million)
- Number of data integrations (770)
- Time to deploy models (34-36 weeks for financial services and oil & gas customers)
- Time for customer developers to build AI applications (26x faster)
- AI patents (for systems, methods, and devices for an enterprise AI platform)

The company also offers case studies for optimal use cases and compatible data sources. While there are neither clear industry-wide benchmarks nor audit standards for these KPIs, accounting for them demonstrates the legitimacy of the company's AI infrastructure. Disclosing that the company spent 10 years and $800.0 million to develop its platform underscores the challenge of developing a commercially viable AI platform-as-a-service.

We believe the company has been rewarded for its technical advances as the market has overlooked **C3.ai**'s high customer concentration and slowing revenue growth in granting the company an 81.4x EV/trailing revenue multiple as of December 31, 2020.

Upstart (NASDAQ: UPST) actualizes the controversial theory of using AI to make lending decisions. Doing so requires rigorous regulatory compliance. As such, the company discloses fine-grained details of its AI models, with additional data points including:

- Number of model features (1,600)
- Training data points (9 million)
- Model techniques (stochastic gradient boosting)

The company also offers insight into the regulatory processes needed to prove a lack of unlawful bias, or ethical AI. To gain regulatory approval from the US Consumer Financial Protection Bureau (CFPB), **Upstart** agreed to compare its AI model decisions to conventional underwriting processes and report the results. The company proved that its models resulted in more loan approvals at lower interest rates than conventional models. The result was the CFPB's first no-action letter, reducing regulatory uncertainty for **Upstart**.

Recent IPOs increasingly detail the barriers to starting an AI company, both financial and technical. AI model development is inherently iterative and does not fit the lean startup methodology common in enterprise software. Both **Upstart** and **C3.ai** required nearly 10 years to develop models they can bring to market today and are just now in the growth phase. The trajectory stated in **Palantir**'s Q3 2020 S-1 filing was similar and demonstrates that AI companies require long-term commitment to develop accurate and extensible models.

Figure 2.

Financial performance of recent publicly listed AI-integrated companies

COMPANY	SEGMENT	CATEGORY	ENTERPRISE VALUE ($B)	REVENUE GROWTH YOY	GROSS MARGIN (TTM)	EV/TRAILING REVENUE
C3.ai	Horizontal platforms	AIaaS	$13.4	71.0%	75.1%	81.4x
Lemonade	Vertical applications	Insurtech	$6.4	199.1%	N/A	65.9x
Bill.com	Vertical applications	Payments	$10.4	45.5%	75.0%	61.9x
CrowdStrike	Vertical applications	Information security automation	$45.9	92.7%	72.9%	60.2x
Datadog	Horizontal platforms	AIOps	$29.1	83.2%	78.7%	54.0x
ZoomInfo Technologies	Vertical applications	Sales & marketing automation	$19.7	103.3%	76.8%	46.1x
Palantir Technologies	Horizontal platforms	AIaaS	$42.6	24.7%	64.2%	42.6x
Accolade	Vertical applications	Personal health	$2.5	39.8%	43.6%	16.9x
Sumo Logic	Horizontal platforms	AIOps	$2.5	49.6%	70.2%	13.1x
Upstart	Vertical applications	Lending analytics	$1.5	70.5%	N/A	6.8x

Source: PitchBook | *As of December 31, 2020

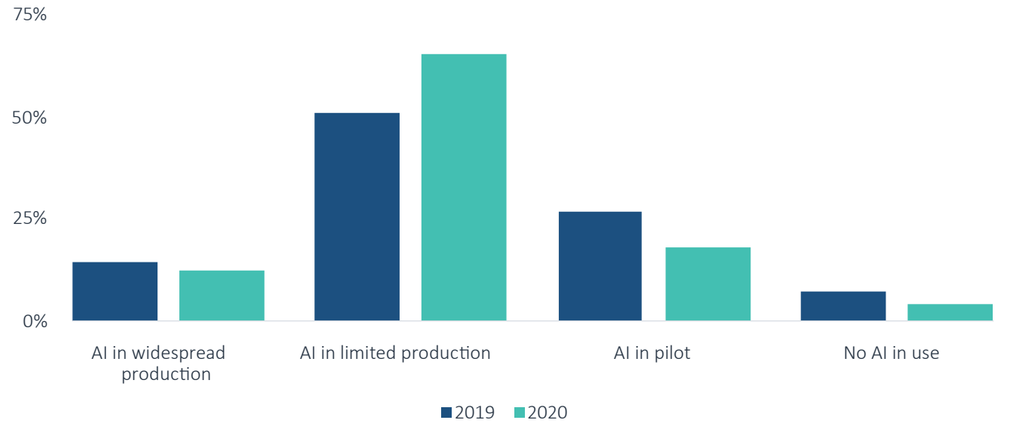
Adoption trends

Despite the heightened priority given to automation and data science initiatives, adoption of AI remains uneven among data science practitioners. According to a recent survey of machine learning adoption by dataset platform **Kaggle**, the percentage of data science professionals reporting usage of ML methods by their employers decreased from 2019 to 2020.[1] 51.5% of professionals used ML at work in 2019 compared to only 45.6% in 2020. Of adopters, 25.2% only use ML in test environments. This result is consistent with industry surveys we have previously cited. The uneven response rate reinforces our finding that ML is not widely used in practical applications by data teams, limiting the addressable market for horizontal platforms.

The percentage of widespread AI usage throughout an organization is also not increasing. While many cross-functional enterprise workflows are being brought to the cloud, AI is limited to specific use cases. According to a survey of chief data and analytics officers, the percentage of companies in which AI has gained widespread adoption also declined from 2019 to 2020 and remains around 10%. That limits the budgets that can be deployed for machine learning enablement software and privileges vertical applications that can sell directly to business line users rather than data scientists. Other industry surveys find that AI budgets remain low, and we believe further integration of internal AI efforts with business use cases will be needed to boost enterprise AI's portion of enterprise software budgets.

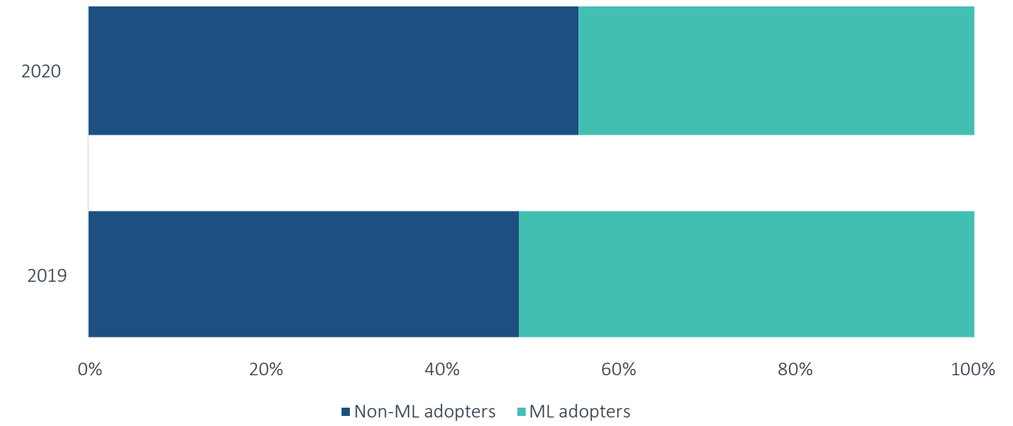
1: "2020 Kaggle Machine Learning & Data Science Survey," Kaggle, January 2021.

Figure 3. LEVEL OF AI ADOPTION FOR CHIEF DATA & ANALYTICS OFFICERS



Source: NewVantage Partners. Fields edited by PitchBook.

Figure 4. % OF DATA PROFESSIONALS REPORTING ML USAGE BY THEIR EMPLOYER



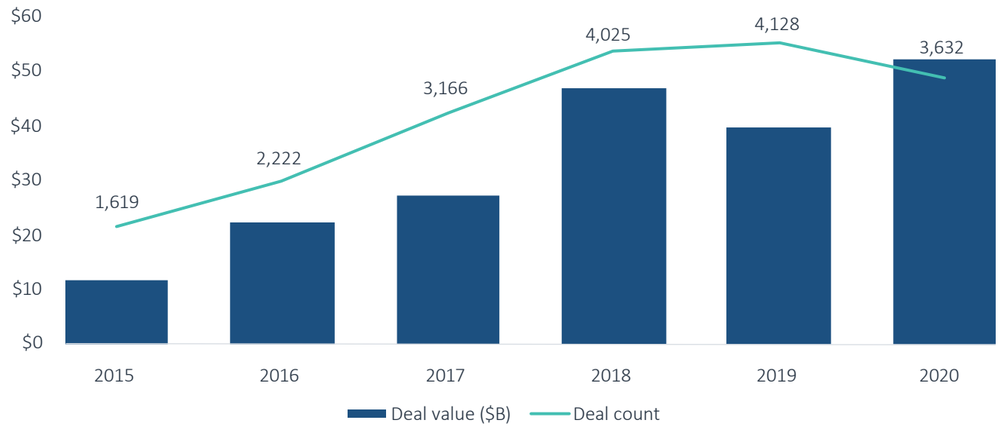
Source: Kaggle

VC activity

VC appetite for AI investment accelerated in 2021, driven by vertical applications. Q4 2020 set a record for quarterly AI & ML VC funding with $16.5 billion invested, contributing to a new high of $52.1 billion invested in the segment in 2020. Deal activity in 2020 was led by healthcare, consumer, enterprise IT, and autonomous vehicles, with each category raising over $5 billion. Declines occurred in financial services, AI core, AI automation platforms, and computer vision. Investment was concentrated in late-stage companies, with early-stage deal count declining as late-stage deal count increased. Picking winners led to a breakout in valuations, as median late-stage valuation rose from $45.0 million in 2019 to $65.0 million in 2020. In aggregate, median valuation step-up for the vertical declined for a second consecutive year to 1.6x, in line with the VC market median, suggesting that most startups struggled to grow their valuations during the COVID-19 pandemic.

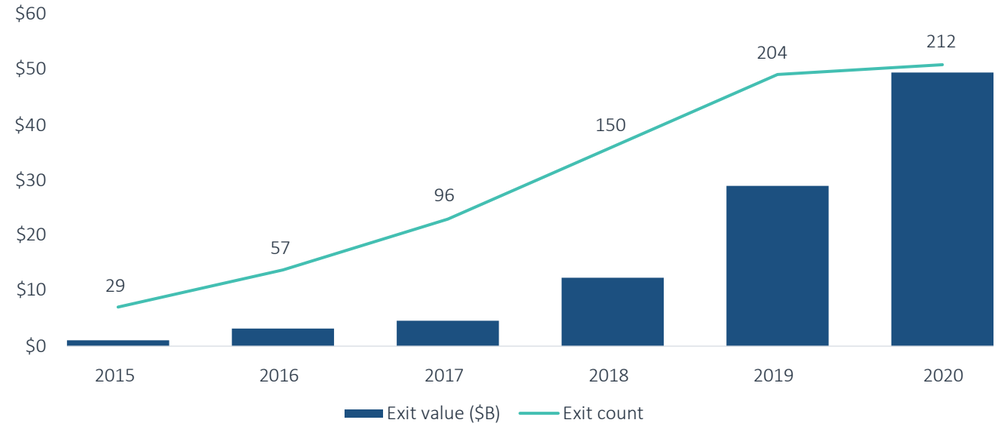
Exit value amounted to $49.6 billion in 2020, nearly matching total VC investment and demonstrating the maturation of the AI theme. IPOs drove this total, contributing $40.4 billion. Q4 saw listings from AI-powered lender **Upstart** (NASDAQ: UPST) and industrials-focused AIaaS unicorn **C3.ai** (NYSE: AI). These successful listings have set the stage for a wave of enterprise AI IPOs over the next 1–2 years. M&A activity was muted in 2020, with both deal value and deal count flat YoY. Deal value was significantly driven by a shakeout in autonomous driving, which featured $5.4 billion in exits for **Uber Advanced Technologies Group** and **Zoox**. Beyond those, we did not track an acquisition value over $1 billion after January. The pandemic has led to low exit values and tepid appetites from big tech companies. We did not observe an AI acquisition by a FAMGA (**Facebook**, **Apple**, **Microsoft**, **Google**, and **Amazon**) company after June 30, 2020, suggesting their focus is on internal R&D efforts. We are expecting a resurgence in M&A activity in 2021.

Figure 5. AI & ML VC DEAL ACTIVITY



Source: PitchBook | Geography: Global | *As of December 31, 2020

Figure 6. AI & ML VC EXIT ACTIVITY



Source: PitchBook | Geography: Global | *As of December 31, 2020
Note: Excludes Uber IPO

AI & ML VC ecosystem market map

Click to view interactive market map on the PitchBook Platform

Market map is a representative overview of venture-backed or growth-stage providers in each segment. Companies listed have received venture capital or other notable private investments.

Horizontal platforms

AI automation platforms

UiPath, MiningLamp, ThoughtSpot, Olive, coveo, QINIU, walkme, workato, HYPERSCIENCE, turbonomic

Natural language technology

Dataminr, mobvoi, AISPEECH, 云知声 Unisound, dialpad, SoundHound Inc., CLARABRIDGE, grammarly, interactions, uniphore

Computer vision

sensetime, Face++ 旷视 Cognitive Services, 云从科技 Cloudwalk Technology, trax, 依图 YITU, meero, Matterport, HOVER, ORCAM

AI core

databricks, 4Paradigm 第四范式, OpenAI, DataRobot, TRIFACTA, CloudMinds, scale, data iku, TRANSWARP, Lucidworks, neo4j, sparkcognition, BEYOND LIMITS, Cambridge Quantum Computing, roobo, Preferred Networks, H2O.ai, INSTABASE, DOMINO, StreamSets

Vertical applications

Consumer

ByteDance, 今日头条 Toutiao, 便利蜂 Bianlifeng, FAIRE, INMOBI, BERKSHIRE GREY, CONTENTSQUARE, Woxing Wosu, TEKION, THINK 豌豆思维

Financial services

JD 京东数科, MoneyLion, blend, Mission Lane, Tradeshift, konfio, 百信银行 AIBANK, Fundbox, ZEST AI, PRODUCEPAY

Healthcare

平安医保科技 Ping An, TEMPUS, zymergen, babylon, HUMAN LONGEVITY, INC., Recursion, XtalPi, Welltok, Valo, health, SYNTHEGO, LinkDoc, SOPHiA, Spring Rain Doctor, insitro

Industrial

indigo, Relativity, CONVOY, FARMERS NEW ENERGY, 能链集团, 千寻位置 Qianxun SI, UPTAKE, Bright Machines, SOLIDIA, BOWERY

Transportation

VIA, nauto, PATEO, Mogu Chelian

Vertical applications (cont.)

IT

FIGURE, SentinelOne, Lookout, ARCTIC WOLF, GONG, Globality, Checkr, Outreach, SimilarWeb, ASAPP

AI & ML semiconductors

Chips

Horizon Robotics 地平线机器人技术, BITMAIN, GRAPHCORE, Enflame 燧原科技, SambaNova, 墨芯科技 MOORE THREADS, Innovium, INUITIVE, cerebras, Xsight

Edge AI software

FOGHORN, ABEJA, Clobotics, LEAPMIND, StartDT

Intelligent sensors & devices

samsara, TERMINUS 特斯联, LedderTech, VERKADA

Autonomous machines

Autonomous vehicles

WELTMEISTER 威马汽车, WAYMO, AIWAYS, SINGULATO 奇点汽车, pony.ai, DiDi, Aurora, tusimple, WeRide 文远知行, QUANERGY

Robotics & drone design

nuro, UBTECH, Geek+, 奥比中光 ORBBEC, quicktron 快仓, Skydio, 竹间 EMOTIBOT, Vicarious, FLEXIV, SHIELD AI

Robotics & drone software

brain corp, AGILE ROBOTS Laiye, 来也, covariant, PERCEPTO

SEGMENT DEEP DIVE

Horizontal platforms

Overview

Horizontal platforms empower end users to build and deploy AI & ML algorithms across a variety of use cases. These platforms directly apply scientific advances in AI & ML research to commercial applications. Companies in this segment have differentiated AI & ML approaches and are built with AI & ML from the ground up (also referred to as AI-first). Furthermore, some horizonal platforms are used to improve AI & ML algorithms but do not use AI & ML themselves. Subsegments include:

AI core: Building blocks of AI & ML deployments, including DevOps tools needed to build and deploy models to production. Categories within this subsegment include: AI as a service (AIaaS), AI & ML developer tools, AI PaaS, autoML, cognitive computing, data preparation platforms, quantum AI, and TinyML.

Computer vision: The use of AI & ML to analyze visual data and make meaningful predictions about both the physical world and digital images. The technology can be used across use cases to label and make predictions about visual data. Key technologies utilizing computer vision across a range of verticals include: AI-enabled augmented reality, computer vision as a service, facial recognition, geospatial analysis, and visual data labeling software.

Natural language technology: Analysis and interpretation of human communications. Natural language technology uses computational linguistic techniques to learn from communications data and make predictions about the structure and content of language. Companies that apply novel developments in natural language technology research to

commercial use cases are included in this segment, even if they claim a specific focus such as customer service, legal automation or contract automation. This field can also be referred to as natural language processing (NLP), but we use a broader term to differentiate between companies using only NLP and those applying more recent developments in the field such as understanding and language generation. Categories within this subsegment include conversational AI, neural machine translation, natural language generation, natural language processing, and natural language understanding.

AI automation platforms: Software and services that enable enterprises across all verticals to leverage AI to automate critical business processes via predictive analytics. Categories include AI-first applications of the following products: AI for IT operations (AIOps), business intelligence, contract lifecycle management automation, database management, decision intelligence, and intelligent process automation.

Industry drivers

Shortage of data science talent: A lack of data scientists in every industry requires software solutions that can streamline the building and deployment of AI & ML models. 54% of enterprises with some AI deployments perceive a moderate to extreme skill gap facing their AI projects, presenting a barrier to becoming mature AI organizations. As a consequence, over 50% of all respondents in a recent Gartner survey reported an AI skills gap as the top barrier to AI adoption.[2] An Ernst & Young poll that confirmed 45% of senior AI professionals believed the lack of qualified AI professionals was the single biggest barrier to AI implementation across business operations.[3]

Security and data integration are common barriers to adoption: AI adopters struggle to leverage their existing databases for AI & ML model building. Common AI & ML frameworks lack integrations with popular business intelligence and analytics tools, requiring lengthy timelines for data wrangling that do not always result in satisfactory AI & ML models. Furthermore, the requirement of third-party model training produces security vulnerabilities that may not conform to enterprise policies. As a result, enterprises are encouraged to build internal solutions from a medley of open-source and commercial tools that can take months to over a year to coordinate, before a model is even to proven to be valuable.

Emerging NLP frameworks: In late 2018, **Google** released BERT (Bidirectional Encoder Representations from Transformers). The framework's innovation of bidirectional text analysis enabled better contextual understanding of text. Since that time, emerging algorithms including ALBERT, RoBERTa, and GPT-3 have achieved state-of-the-art results on a range of NLP tasks.

Improved computing power: Nvidia's recent A100 GPU exceeded the performance of its previous V100 model by 6x for NLP model training, enabling larger training runs and massive neural networks such as GPT-3.

Enterprise adoption of AI: According to a recent survey, 50% of enterprises across industries have reported using AI in at least one business function, creating a large addressable market for AI platforms.[4]

2: "State of AI in the Enterprise, 3rd Edition" Deloitte Insights, July 2020.
3: "Digital challenges: Overcoming Barriers to AI Adoption," MIT Technology Review Insights and EY, May 2019.

4: "Global survey: The state of AI in 2020," McKinsey Analytics, November 2020.

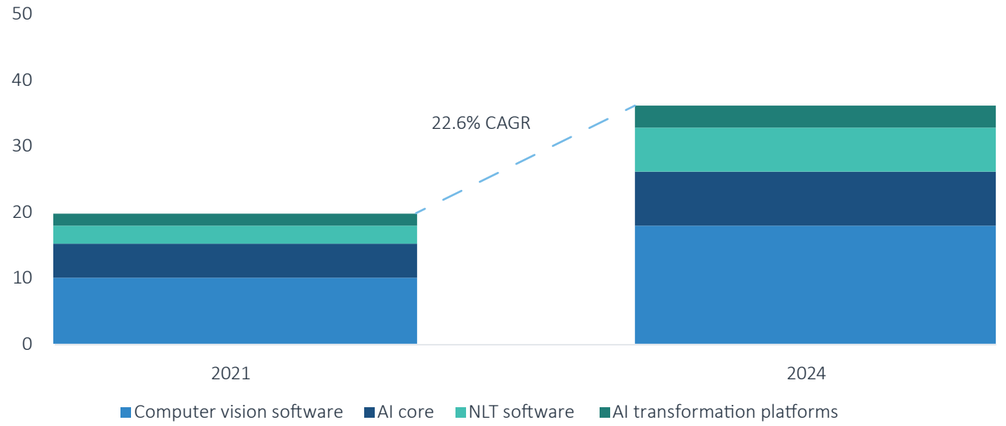
Market size

We estimate spending on the categories in this segment to reach $17.7 billion in 2021, growing at a 22.6% CAGR to $32.7 billion in 2024. This market size includes our estimate of spending on computer vision and NLT applications—even when they overlap with vertical application spending. Given that computer vision emerged earlier than NLT, we estimate its market will reach $10.0 billion in 2021, compared to only $2.7 billion for NLT. However, we expect NLT to grow faster than computer vision over the next three years at a 35.3% CAGR. Independent of these technique-specific platforms, we estimate 2021 will see $5.0 billion spent on AI core and $1.8 billion on AI automation platforms. Both markets should grow at nearly 20% over the next three years.

Disruption potential

AI software is currently dominated by tech giants including **IBM**, **SAS**, **Amazon**, **Microsoft**, **Google**, and **Facebook**. Despite the leadership of these giants, ML frameworks have faced implementation barriers with unsophisticated citizen developers and IT operations teams. Citizen developers are forced to use simple RPA, since open source ML frameworks require coding ability and IT operations teams face challenges deploying models and monitoring the performance of models in production. Incumbents are innovating rapidly but can lose market share to citizen developer-friendly frameworks that have built-in deployment and monitoring integrations. Open source libraries particularly can attract developer communities away from incumbent models. The market leader **IBM** has only around 10% market share, suggesting that companies with innovative approaches can become market leaders.

Figure 7. HORIZONTAL PLATFORMS MARKET SIZE ($B)



Source: PitchBook, Gartner, IDC, Omdia | Geography: Global

COMMON INDUSTRY KPIS

Operational

- Ease of use for citizen data scientists
- Stages of the AI & ML lifecycle addressed
- Number of frameworks, model varieties, and programming languages supported
- Scalability across neural network workloads
- Explainability
- Level of automation
- Financial
- Conventional SaaS metrics

Business model

Horizontal platforms typically contain a variety of computational costs, including the following components for a typical hyperscaler ML deployment:

Infrastructure costs

- **Storage:** Data is typically stored in a cloud server and carries nominal cost based on the scale of the data.
- **Network:** Data moved into and out of the cloud environment is charged on a volume-based rate.
- **Compute:** Model building includes separate costs for building, training, and deployment.

Operational costs: EC2 instances carry an hourly rate that can vary by the contract length and compute requirements, among other factors.

Security & compliance costs: AI & ML data must have an audit trail for various data privacy compliance frameworks.

Each of these components has separate costs for both training and deployment. As a result, training a single AI model can cost around $1 million in compute resources. Optional incremental costs include: breach discovery and log management, maintenance and support costs, data encryption, and secure web gateways. Startups can unbundle hyperscaler platforms and provide data pre-processing and monitoring on top of the hyperscalers' platform.

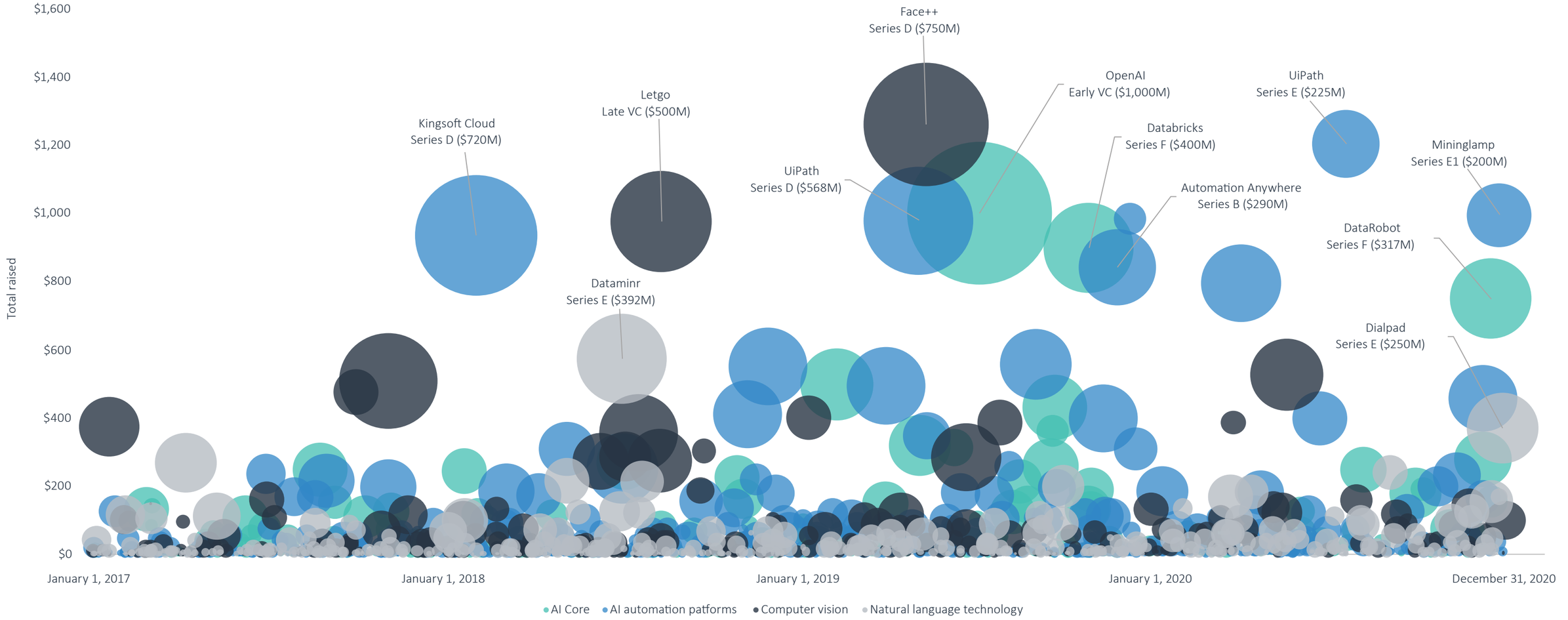
VC activity

Horizontal platform investment dropped off in 2020 in both deal value and deal count, with VC investment falling under $10 billion. We believe startups may have been challenged to grow revenue during the pandemic, causing deal count to drop precipitously in Q1 and rebound in Q4. The market for pure-play AI software tools may not yet be large enough to support a diverse ecosystem of vendors. Most categories declined in funding totals YoY with notable exceptions including conversational AI, NLP platforms, data preparation platforms, and facial recognition. However, leading private companies have achieved high valuation growth through the pandemic. In Q4, leading AIaaS unicorn **DataRobot** raised a Series F at a 1.9x valuation step-up as it pursues an IPO. Scale AI, a computer vision data preparation vendor, achieved a 3.3x valuation step-up into its Series D. NLP demonstrated commercial success as conversational AI startup **Dialpad** became a unicorn with its Series E, breaching the $100 million annual recurring revenue mark in 2020. The pandemic is putting pressure on the early stages of the ecosystem yet is creating winners at the high end.

Exit activity highlighted that enterprise AI has achieved scale commercially. **C3.ai** (NYSE: AI), an AIaaS platform focused on industrial applications, debuted on public markets with a $4.0 billion valuation. The company's revenue growth earned a 24.4x EV/trailing revenue multiple in its IPO, and the stock has traded well since its debut. Its growth demonstrates the opportunity to develop ground-up platforms to rewire legacy industries such as oil & gas and utilities with AI analytics. The company spent more than five years offering non-AI data analytics and built industry expertise before layering in AI capabilities. In M&A, however, the exit of **Element AI** demonstrated the challenges of commercializing AI as the company sold for less than the $257.0 million it had raised in an acquisition framed as an acqui-hire. The company was unable to hit milestones to justify its $102.0 million Series A in 2017 and raised what was likely a flat round in 2019. Other acquisitions did not disclose exit values, likely including some acqui-hires. We have seen low exit count in the segment in 2020.

Figure 8.

Horizontal platforms VC landscape ($B)



Source: PitchBook

Note: The left axis indicates total VC raised as of deal date. Bubbles indicate amount raised
Excludes Sensetime, Face++, and **Palantir** *for scale.*

Figure 9.

Notable horizontal platforms VC deals

COMPANY	CLOSE DATE	SUBSEGMENT	DEAL TYPE	DEAL SIZE ($M)	LEAD INVESTOR(S)	VALUATION STEP-UP
Dialpad	December 21, 2020	NLP	Series E	$250.0	OMERS Growth Equity	2.2x
DataRobot	December 9, 2020	AIaaS	Series F	$317.1	Altimeter Capital Management	1.9x
Scale AI	December 1, 2020	Data preparation platforms	Series D	$155.0	Tiger Global Management	3.3x
Chooch AI	November 16, 2020	Computer vision	Series A	$20.0	Vickers Venture Partners	15.3x
Abacus.AI	October 23, 2020	AI & ML development platforms	Series B	$22.0	Coatue Management	1.4x

Source: PitchBook

Figure 10.

Notable horizontal platforms VC exits

COMPANY	CLOSE DATE	SUBSEGMENT	EXIT SIZE ($M)	INDEX/ACQUIRER	VALUATION STEP-UP	VALUATION/TRAILING REVENUE
Element AI	December 19, 2020	AIaaS	$230.0	ServiceNow	N/A	N/A
C3.ai	December 9, 2020	AIaaS	$3,374.6	NYSE	1.0x	24.4x
Palantir Technologies	December 31, 2020	AIaaS	$21,000.0	NYSE	N/A	21.0x
Sumo Logic	September 17, 2020	AIOps	$1,845.6	NASDAQ	1.6x	12.0x
Letgo	March 25, 2020	Computer vision	$120.0	OfferUp	N/A	N/A

Source: PitchBook

HORIZONTAL PLATFORMS

Figure 11.

Key VC-backed horizontal platforms companies

COMPANY	TOTAL VC RAISED ($M)*	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	MOST RECENT LEAD INVESTOR(S)
SenseTime	$2,866.6	Horizontal platforms	Image recognition	Access to Chinese population data	SB China Venture Capital and Huaxing Fund
Face++	$1,258.0	Horizontal platforms	Face detection	High precision and accuracy	Bank of China Group Investment
UiPath	$1,202.2	Horizontal platforms	UiPath Studio	Vast partner network to support integrations	Accel
Databricks	$897.4	Horizontal platforms	Unified Data Analytics Platform	Scalable cluster management for ML models	Andreessen Horowitz
Automation Anywhere	$840.0	Horizontal platforms	Automation Anywhere Enterprise	Pre-built components for script writing	Salesforce Ventures

Source: PitchBook | *As of December 31, 2020

Figure 12.

Key horizontal platforms incumbents

COMPANY	TICKER	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	EV/TRAILING REVENUE*
Amazon	NASDAQ: AMZN	AI core	AWS Sagemaker	Integrated development environment for machine learning	36.4x
Alphabet	NASDAQ: GOOGL	AI core	TensorFlow	Best-in-class deep learning framework	21.0x
International Business Machines	NYSE: IBM	AI core	Watson Studio	Manages data engineering via Red Hat OpenShift	11.0x
Microsoft	NASDAQ: MSFT	AI core	Azure ML	Supports citizen developers via augmented analytics	22.6x
SAP	ETR: SAP	AI core	SAP Intellligence	Integrates with SAP's numerous data and analytics platforms	19.4x

Source: PitchBook | *As of December 31, 2020

Opportunities

Human-in-the-loop data preparation: Enterprise datasets currently require intensive manual labor to be used in AI & ML (i.e. "human-in-the-loop"). 39.5% of data scientists surveyed by Appen spend more than 50% of their time managing, cleaning, and/or labeling their data.[5] A recent survey of NLP practitioners found that 36% of active users use a data annotation tool while only 12% outsource data labeling. We expect those statistics to reverse in the next 3 years as AI operations become more efficient. Startups in this space have historically focused on preproduction of data and are increasingly supporting maintenance of datasets in production. In Q4, computer vision data preparation unicorn Scale AI achieved a 3.3x valuation step-up to a $3.3 billion pre-money valuation based on high demand during the pandemic for training data. Use cases achieving high growth included customer support, route optimization, and delivery robot perception. Scale AI also serves **OpenAI**, which trained the largest ever language model in 2020. Leading startups have built large teams of data labelers to support their customers in preparing datasets. We believe ML can itself be used to prepare data for further ML models, but human-in-the-loop solutions are currently required to prepare computer vision and language data for model building. This category is likely to experience high demand over the next three years.

MLOps: MLOps can streamline AI & ML development within organizations and achieve scale as enterprises develop internal AI & ML capabilities. There is currently a cultural gap between ML data scientists and IT operations professionals, similar to what was seen in software development over 10 years ago. Given the nascence of the field, prevalence of data scientists without experience shipping software, and emerging deployment processes, ML has not undergone the DevOps shift of more general software development. As a result, over 25% of data scientist time is spent on deployment efforts, an industry survey finds.[6] While MLOps spans a broad lifecycle from data preparation to model management, we believe these deployment challenges create opportunities to build deployment pipelines for trained models to a variety of runtime environments, depending on the model. In Q4, UK-based MLOps startup **Seldon** raised a Series A at a 2.2x valuation step-up from seed, achieving high open-source traction and reporting 38% monthly growth over the previous two years. The space can clearly support multiple vendors focused on different types of models and end users. We believe that increased enterprise adoption of ML will require widespread adoption of MLOps, enabling scale similar to DevOps tools vendors such as Datadog and New Relic over the medium term.

Machine learning security (MLSec): MLSec is underaddressed by startups and may create a new field of information security. ML carries its own security risks, including data poisoning, model theft, and reverse engineering. Current AI & ML processes integrate highly sensitive data in experimental and open source environments without security experts involved in the process. Contrary to popular wisdom, the leading risk for AI practitioners by far is cybersecurity, according to a recent survey.[7] **Microsoft** and Mitre recently launched a threat matrix for machine learning. Mitre's attack frameworks have been adoption drivers for cybersecurity tools more broadly and this framework might encourage ML operations teams to consider their defense-in-depth strategy, while giving vendors opportunities to align with the framework. **Microsoft** finds that a new suite of application security tools will need to be implemented by ML development teams. Much as ML has raced ahead without incorporating DevOps, so it has made its models vulnerable by default. We believe that entrepreneurs are just beginning to address this opportunity head on, as evidenced by the recent seed and incubator rounds for **Scanta**, **TrojAI**, **Neurocat** and **SafeRide Technologies**. In Q4, **Robust Intelligence**, an MLSec

5: "The State of AI and Machine Learning," Appen, September 2019.

6: "2021 Enterprise Trends in Machine Learning," Algorithmia, December 2020.

startup founded by Harvard Computer Science professor Yaron Singer, emerged from stealth with $11.0 million Series A led by Sequoia Capital to address the opportunity.

AutoML: AutoML refers to the automation of the ML lifecycle, including hyperparameter optimization, data pre-processing, and model architecture selection. AutoML has been commoditized by tech giants and large startups **Dataiku**, H2O.ai, and **DataRobot**. Startups are creating autoML platforms for use by citizen developers that can build low-code ML models and integrate them with RPA platforms to fully automate routine tasks. **dotData** has developed a competitive autoML platform that has distinct user interfaces for business users and data scientists, automating both data pre-processing and feature engineering. This unique feature set led to a $100.0 million pre-money valuation in the company's Series A, the highest among the autoML cohort we track. Previously, **SparkCognition** achieved a Series A pre-money valuation of $85.0 million. In Q3, public autoML vendor **Veritone** restarted its revenue growth from pandemic lows and noted increased demand from customers for low-code solutions, given the need to deploy ML solutions rapidly. The company noted particularly high growth from its advertising customers. We believe that RPA has scaled with simple rules-based automations that fail to provide significant business value, and yet the technology has grown into a $1.9 billion market. As a result, autoML can become part of the RPA tech stack if business user interfaces gain traction.

GPT-3. Developed by **OpenAI**, GPT-3 is an NLP model trained on internet text, Wikipedia, and books, resulting in an unprecedented 175 billion features that are used to interpret language and generate responses. Recent research has found that the model outperforms prior models on knowledge understanding, including subjective exam questions. However, the model performs only 20% better than random chance across a range of topics, meaning that it lacks mastery of simple language understanding tasks.[8] **Microsoft** invested $1.0 billion in **OpenAI** to build the model, conducted training on its cloud service Azure, and has since announced an exclusive agreement to commercialize the model. **Microsoft** may expect this partnership to, in part, offer an ROI on its investment in **OpenAI**. A wave of startups is developing to implement the model in unique commercial applications. Examples include automatic email composition (**OthersideAI**, Compose.ai, Magic Email), digital avatar chatbots (**Replika.ai**), web app development (Debuild), and copywriting generation (**Copy.ai**, **AI21,** Copysmith). Numerous other projects are emerging in natural language database queries, code completion, and edtech chatbots. In Q4, **OthersideAI** raised a $2.6 million seed round led by Madrona Venture Group, suggesting that some of these startups may be able to establish defensible business models. GPT-3 has the potential to become a platform for NLP startups, similar to the iPhone's iOS, that can democratize access to advanced computing. We believe these early-stage opportunities will lead to some startups scaling in NLP as **Cognex** did in machine vision to a market cap of over $10 billion.

Conversational AI for contact center as a service (CCaaS). Contact center automation is emerging as a killer app for conversational AI given growth in the market and disruption from COVID-19. CCaaS platforms provide cloud-native software infrastructure for voice and digital communications with both customer-facing and employee-facing components. The market reached $3.5 billion in 2019 with 29.7% growth, according to Gartner. COVID-19 is forcing enterprises to retool their contact centers for agent workflow optimization and automation. Valuation growth in this niche has been rapid and continued in Q4. E-commerce CCaaS startup **Gorgias** raised a Series B at a $300.0 pre-money valuation, the largest in this space

7: "Global Survey: The State of AI in 2020," McKinsey Analytics, November 2020.

8: "Measuring Massive Multitask Language Understanding," Dan Hendrycks et al, September 2020.

since Observe.AI's $250.0 million valuation in July. Virtual agents enable contact centers to scale in line with demand, which has become increasingly unpredictable during the pandemic. Current CCaaS leaders rely on third-party integrations for chatbots, creating the potential for a conversational AI platform to capture market share from them. We believe that an AI-first startup is likely to become a market leader over the medium-term.

Software 2.0: Deep learning can be used to create software applications that perform specific tasks without any human coding. By using the matrix multiplications of neural networks instead of scripts and binary code, ML-powered software can become more reliable and malleable than today's software applications. Its reliability stems from the consistency of model calculations, which provide an efficient path between instructions and outputs, compared to the brittleness of hand coded and compiled applications. Its malleability owes to machine learning models' ability to retrain themselves based on new information without requiring software development resources. Such software programs will require a new developer stack that includes interactive developer environments, data preparation platforms, model repositories, and monitoring.

In 2020, rapid revenue growth enabled AI developer platform Abacus.AI to grow its post-money valuation from $23.3 million to $102.0 million across two deals led by Index Ventures and Coatue Management. Its recently announced product suite offers unbundled ML model development modules including model hosting & monitoring, explainable AI, and a real-time ML feature store. These components are sufficient to build AI applications across ITOps, fraud, demand forecasting, sales & marketing, predictive analytics, and recommendation engines. We expect these use-case-focused AI & ML development platforms to close the gap between data science and business line teams. AI ultimately has the potential to overtake the $324.5 billion enterprise software industry.

Considerations

Early attempts at enterprise-wide automation have failed to deliver cost savings: RPA has demonstrated the constraints on enterprise automation. The category has grown quickly and promises to produce strong outcomes for startups. In practice, its implementation has met barriers in fully automating workflows and has been shown to enhance productivity rather than generate real cost savings. RPA implementation often requires specialized developers and scripts that fall apart when unforeseen cases arise. The segment has to continued to grow at a high rate, reflecting the priority of enterprises to improve slow and inefficient legacy processes, but highlights the barriers to automating processes across business units.

Public cloud hosts offer comprehensive capabilities for AI & ML model building and deployment: Public cloud hosts can offer AutoML, language and vision services. They also feature AI marketplaces that compete with horizontal platforms. AI marketplaces include APIs, microservices, datasets and prebuilt algorithms. Incumbents with offerings in this space include **Alibaba**, **Amazon**, **Apple**, **Baidu**, **Microsoft**, **Google**, **Tencent** and **IBM**. We believe enterprises with limited budgets can spin up limited AI & ML projects with a combination of open-source frameworks and cloud management tools. For this reason, scaling of multi-cloud and API-based MLOps tools may be delayed until ROI for existing AI & ML projects is strong enough to encourage heavier investment in AI & ML capabilities across the enterprise.

Explainability an expanding problem: Deep learning is compounding the black box nature of AI & ML as it becomes more effective, since its features are inherently fluctuating and obscure to even the data scientists training them. Models can lack clarity on critical metrics including privacy, security, ethics, and transparency, limiting their utility in sensitive use cases. The

Financial Stability Board has listed the interpretability of AI & ML models as a macroeconomic risk, encouraging regulation. Furthermore, in natural language, explainability is critical to understand the relationships identified by the algorithm but is not available without a ground-up focus during the requirements stage. We believe vendor differentiation will emerge based on the level of model transparency offered.

Outlook

Cloud hosts to limit market share for startups: The advanced and constantly improving AIaaS offerings of hyperscalers will make it cost effective for many enterprises to take their AI journey with their cloud provider. The lack of data scientists and operational challenges of establishing an AI center of excellence will make it difficult for startups to gain market share from incumbents, so they will be forced to build on top of them. Likely outcomes may resemble the DevOps market with several IPOs and a middle ground of trendy startups not achieving unicorn status and facing exits to incumbents or PE firms. The lofty initial valuations in the space, in part related to the 30-50x revenue multiples granted to RPA startups, may not be fully justified by commercial outcomes, although we believe that adoption rates will be high by 2023 and startups can capture some market share.

Natural language technology (NLT) to receive the highest VC funding of AI & ML horizontal platform categories. NLP is experiencing technical breakthroughs in 2020 that position the technology to become a building block for startups and enterprises alike going forward. The category has not historically been the highest-funded category of AI given the low performance of the technology relative to rules-based SaaS applications. As a result, it was lower funded than AI automation platforms and AI core in 2020, a year after being the lowest-funded horizontal AI category. Code automation, document discovery, and enterprise search are applications of NLP that should increase in adoption in the coming year. We have seen a consistent flow of late-stage deals in NLT that may create a wave of unicorns in the near-term.

C3.ai, Databricks, and DataRobot IPOs to attract VC investors to AIaaS platforms. The wave of ML-integrated IPOs in 2019 and 2020 has focused on vertical applications and AI automation platforms. Pure-play enterprise AI core has remained out of the public markets and has thus given little indication of the scale or scope of enterprise AI adoption. The recent valuation step-ups for **DataRobot** and **Dataiku** suggest that AIaaS is emerging as a standalone software category and is going to produce outstanding venture outcomes. Early-stage investors in this space tend to be corporate investors or thesis-driven VC firms. When combined with increased adoption of AI tools, we believe that investors will gravitate toward the opportunity to invest in the picks and shovels of the next era of software.

Hyperscalers to acquire MLOps startups. We believe that hyperscalers have fallen behind in user experience for ML data governance and model management and will be motivated to acquire startups to bolster their AIaaS capabilities. **Microsoft**'s and **Google**'s CVC arms have been active in supporting MLOps startups including **Dataiku**, **Pachyderm**, **Algorithmia**, and **Databricks**, in part indicating the strategic priority of the space to their parent companies. We believe that MLOps will emerge as a source of competitive differentiation in the compute-intensive AI market. The advanced and constantly improving AIaaS offerings of hyperscalers will make it cost effective for many enterprises to take their AI journey with their cloud provider. As a result, likely outcomes for startups may resemble the DevOps market with several IPOs and a middle ground of trendy startups not achieving unicorn status and facing exits to incumbents.

SEGMENT DEEP DIVE

Vertical applications

Overview

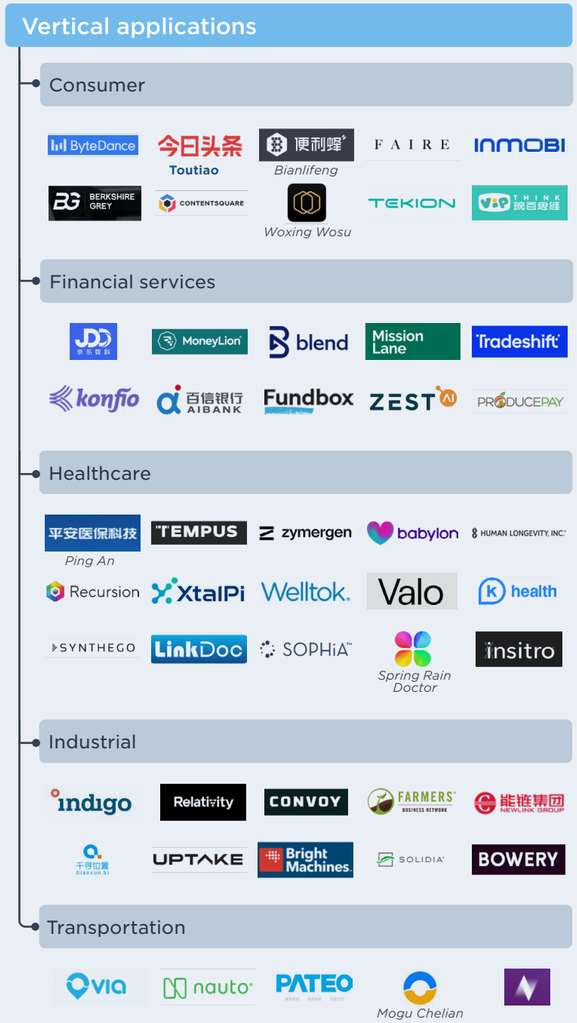
Vertical applications in AI & ML address specific problems within industries and are not always AI-first. Many startups in this category design a solution to an industry problem using software and integrate AI & ML to optimize some part of their product. These solutions typically differentiate based on the quality of the dataset used to train the industry-specific model and the industry expertise of the data scientists in identifying decision-making areas that can be enhanced by AI & ML models. As a result, many of these startups help automate specific functions within their industry but have limited ability to cross-apply their AI & ML to other industries.

Current subsegments include:

AI & ML in financial services: Includes technologies that embed AI & ML into existing financial services via advanced analytics, process automation, robo advisors and self-learning programs. Product categories include: Financial chatbots, intelligent banking, lending analytics, payment optimization, predictive underwriting, and robo-advisors.

AI in healthcare: Includes technologies that leverage AI & ML to improve medicine and the provision of care. Product categories include: AI-based drug discovery, clinical decision support, genetic analytics, healthcare administration, and personal health

Consumer AI: Includes technologies that use AI & ML to enhance B2C business models. Product categories include: AI in media & entertainment, AI & ML adtech, digital avatars and gaming, e-commerce recommendation engines, edtech, intelligent price optimization, and smart retail.



Industrial AI & ML: Includes technologies that automate industrial processes and unlock industrial data to find new efficiencies. Product categories include: Crop maximization, energy grid automation, geospatial analysis, heavy industry automation, IoT predictive analytics, supply chain optimization, and telecommunications optimization.

AI in IT: Includes enterprise software tools that optimize specific functions typically administered by IT departments, including both backend and frontend use cases. Product categories include: Human resource automation, information security automation, IT infrastructure management, legal automation, proptech, sales & marketing automation, and software development tools.

Industry drivers

Growing availability of industrial datasets: Two-thirds (66%) of surveyed companies are willing to share internal data externally to help develop new AI-enabled efficiencies, products, or even value chains.[9] In June 2019, 10 large pharmaceutical producers formed a consortium called the MELLODDY Project to share research data on an unprecedented 10 million chemical compounds on a blockchain that startup partners can support and use to train their AI algorithms. Startups can build preliminary models based on data and frameworks from sources including **Google** Dataset Search, **Kaggle**, UCI Machine Learning repository, **Microsoft** Coco and **GitHub**.

Industrial customers are achieving ROI with AI solutions: Deloitte's State of AI survey found that 81% of leading AI adopters achieve payback periods of two years or less on new AI projects.[10]

As a result of COVID-19, enterprises are shifting decreased IT budgets toward AI solutions: The crisis has given enterprises an opportunity to replace physical and logical computing systems with cloud-based AI systems. Additional layers of analytics can extract efficiencies needed to manage the downturn. The crisis accelerates an existing trend in enterprise budgets, as an Enterprise Strategy Group survey found that AI & ML would see the largest spending increase in 2020.[11] We expect AI spending to grow more than the IT market overall.

Market size

We forecast the vertical applications market to reach $50.2 billion in 2021, with a 27.7% CAGR out to 2024, resulting in a $104.5 billion market. The segment is heavily weighted toward enterprise IT applications, which we view as a $22.5 billion market, led by customer service automations in e-commerce and enterprise sales, growing to a $46.2 billion market by 2023 at a 27.0% CAGR. This estimate includes $2 billion to $4 billion markets in human resources, sales & marketing, information security, and fraud prevention. We estimate industrial AI to be the second-biggest category at $9.9 billion, focused primarily on manufacturing automation and predictive maintenance. Both consumer AI and AI in financial services are around $8 billion markets. Healthcare is a smaller market at $1.1 billion, but we believe it will grow more quickly than the overall segment at a 47.4% CAGR to a $3.5 billion industry in 2023 as more clinical and genetic data is unlocked for clinical decision support and drug discovery.

9: "The Global AI Agenda: Promise, Reality and a Future of Data Sharing," MIT Technology Review Insights, March 2020.
10: "Thriving in the Era of Pervasive AI: State of AI in the Enterprise, 3rd edition," Deloitte, July 2020.

11: "2020 Technology Spending Intentions Survey," Enterprise Strategy Group, February 2020.

Disruption potential

Vertical applications in AI have the greatest potential to disrupt industries with features including paper-based manual processes, unutilized datasets, complex logistical optimization problems, and low regulatory barriers to entry.

For this reason, we believe the following industries are most susceptible to disruption: accounting, construction, energy extraction and generation, and insurance underwriting.

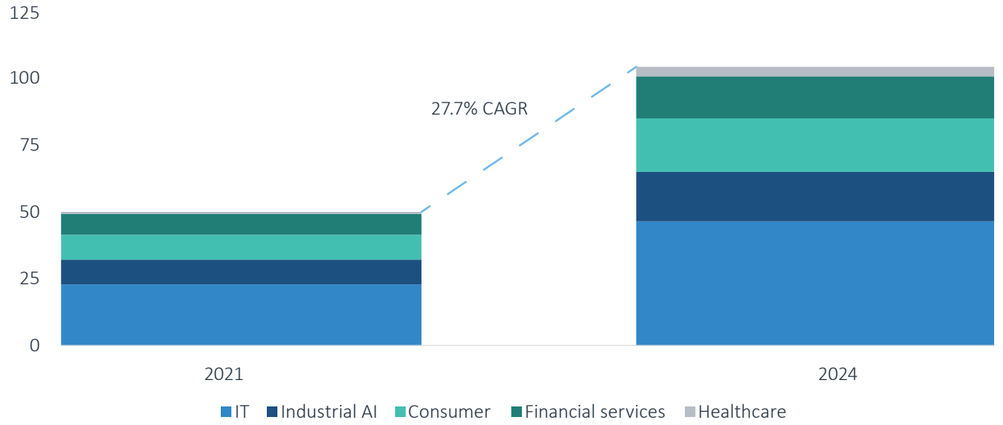
Industries with stringent regulations around data usage including healthcare, utilities, and retail may have higher barriers to entry.

Business model

The business models of enterprise AI vendors can vary widely and have considerable influence on their cash flow profiles. Software AI startups employ several key business models including:

- **SaaS:** A monthly subscription that bundles data engineering, model building and deployment. Additional setup and maintenance fees may be layered on top of a base package. Applications can be delivered as a cloud service through channel partners. In most cases, managed services are required via the vendor or a third-party consulting firm.
- **Co-development agreements:** An AI-first vendor partners with an industry specialist to co-develop an AI solution to be used by the industry specialist as a customer and licensed to the industry specialist to distribute. This often includes an upfront payment to the AI-first vendor to develop the initial models. This arrangement can be necessary to supply AI-first startups with

Figure 13. VERTICAL APPLICATIONS MARKET SIZE ($B)



Source: PitchBook and IDC | Geography: Global

COMMON INDUSTRY KPIS

Operational

- Quantity and quality of data
- Rawness of data
- Exclusivity of data
- Level of diminishing returns to more data
- Number of domain experts on team
- Number of AI & ML experts on team

Financial

- Gross margin
- R&D % of revenue

the data necessary to enter a new industry. It is a low-margin business model compared to IP ownership.

- **Joint venture:** An AI-first vendor and enterprise create a joint venture to develop vertical applications and market them. Typically, the customer is granted a choice to buy out the vendor at a future date.

In each case, AI-first vendors bear costs related to data engineering, model building and maintenance, creating a cost structure that exceeds those of SaaS business models. Consulting firms also arrange outcome-based or upfront capital expenses, though these are not commonly employed by startups. In each case, startups must require customers to consent to data contribution to improve their models without a royalty fee to the customer.

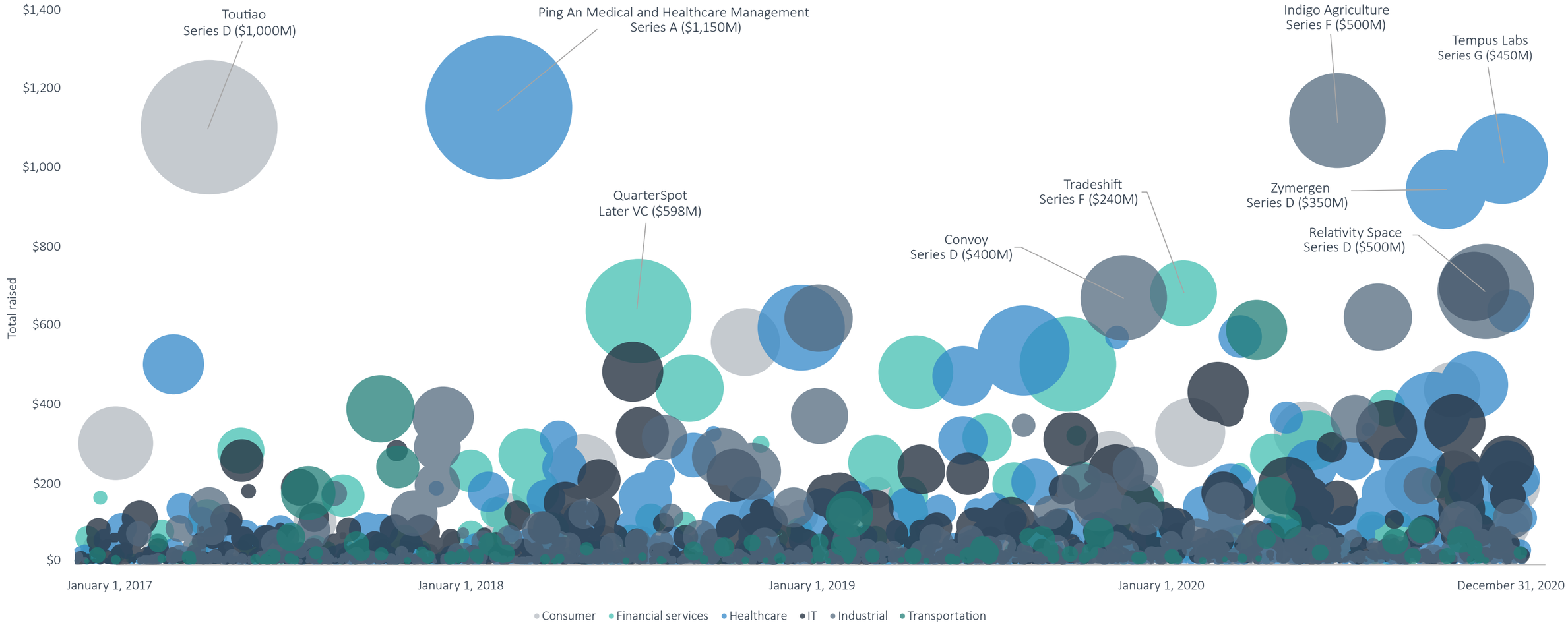
VC activity

VC investment remained robust in 2020, with more than $27.0 billion invested in the space—a record sum. The only categories to raise more than $2 billion were e-commerce recommendation engines, clinical decision support, AI-powered drug discovery, and information security automation, aside from media giant **Kuaishou**'s $3.0 billion round. Healthcare AI continued its surge in Q4, with mega-deals closed by clinical decision support company **Tempus Labs** and computational biology unicorn **Zymergen**. Healthcare is emerging as an ideal application for AI predictions and is achieving outstanding results. AI specialist fund Air Street Capital is focused on the opportunity as a result. In Q4, industrial automation set a record with $500.0 million invested into **Relativity Space**'s Series D to bring AI to spacecraft manufacturing. Founder talent continued to be a primary driver of venture deals, with former **Tesla** CIO Jay Vijayan's startup **Tekion** achieving unicorn status with a 6.7x valuation step-up in its Series C.

The first wave of vertical application IPOs continued with **Upstart**'s listing in December, although M&A activity continued to be light. **Upstart** (NASDAQ: UPST) brings AI to consumer loan analytics and achieved a $1.4 billion valuation upon listing at 6.7x EV/trailing revenue. The listing shows that AI can support scale for fintech startups, as seen previously by **Lemonade** and Bill.com in other niches. Aside from this exit and a robotics IPO in China, no exits in Q4 disclosed an exit value. This continues a concerning trend of low exit values among AI startups in 2020, as the last $1.0 billion acquisition was made in January for Shape Security. We have not seen AI startups take advantage of SPACs yet but expect to see activity given the successful debuts of **C3.ai** (NYSE: AI) and **Palantir** (NYSE: PLTR). Further, AI-focused private equity funds such as Symphony AI may increase their activity. The pandemic may have tempered acquisition appetite and limited the growth of startups. At this stage in the growth cycle of AI, winner-take-most opportunities are emerging, but only a few startups are achieving substantial financial outcomes.

Figure 14.

Vertical applications VC landscape ($M)



Source: PitchBook

Note: The left axis indicates total VC raised as of deal date. Bubbles indicate amount raised. Excludes Kuaishou *and* Bytedance *for scale.*

Figure 15.

Notable vertical applications VC deals

COMPANY	CLOSE DATE	SUBSEGMENT	DEAL TYPE	DEAL SIZE ($M)	LEAD INVESTOR(S)	VALUATION STEP-UP
Zenoti	December 15, 2020	Sales & marketing	Series D	$160.0	Advent International	2.0x
Relativity Space	November 23, 2020	industrial automation	Series D	$500.0	Tiger Global Management	4.5x
SentinelOne	November 11, 2020	Information security automation	Series F	$267.0	Tiger Global Management	2.5x
Recursion Pharmaceuticals	November 11, 2020	AI-based drug discovery	Series D	$245.9	Leaps by Bayer	1.1x
Eko	November 9, 2020	Clinical decision support	Series C	$65.0	Highland Capital Partners, Questa Capital	2.2x

Source: PitchBook

Figure 16.

Notable vertical applications VC exits

COMPANY	CLOSE DATE	SUBSEGMENT	EXIT SIZE ($M)	INDEX/ ACQUIRER	VALUATION STEP-UP	VALUATION/TRAILING REVENUE
Upstart	December 16, 2020	Lending analytics	$1,269.2	NASDAQ	1.7x	6.8x
nCino	July 14, 2020	Intelligent banking	$2,529.1	NASDAQ	N/A	18.2x
Lemonade	July 2, 2020	Insurtech	$1,273.0	NYSE	0.6x	16.1x
Accolade	July 2, 2020	Personal health	$819.8	NASDAQ	1.3x	7.4x
Shape Security	January 24, 2020	Information security	$1,028.0	F5 Networks	1.0x	N/A

Source: PitchBook

Figure 17.

Key VC-backed vertical applications companies

COMPANY	TOTAL VC RAISED ($M)*	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	MOST RECENT LEAD INVESTOR(S)
ByteDance	$8,115.0	Consumer	TikTok	NLP and computer vision video recommendations	Aglaé Ventures, Tiger Global Management, All Blue Capital
Kuaishou	$4,350.3	Consumer	Short-form video app	More upscale than TikTok	Tencent Industry Win-Win Fund
JD Digits	$3,141.3	Financial services	AI financial risk management	Asynchronous federated learning	JD.Com
Ping An Medical and Healthcare Management	$1,150.0	Healthcare	Intelligent Cognition	Multimodal healthcare AI models	SoftBank Investment Advisers
Indigo Agriculture	$1,116.6	Consumer	Indigo Marketplace	Satellite data for AI models	FedEx

Source: PitchBook | *As of December 31, 2020

Figure 18.

Key vertical applications incumbents

COMPANY	TICKER	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	EV/FORWARD REVENUE*
Bill.com	NYSE: BILL	Financial services	Intelligent Business Payments Platform	Accurately categorizes customer bills without manual input	61.9x
CrowdStrike	NASDAQ: CRWD	IT	Falcon OverWatch	Endpoint security knowledge graph	60.2x
Datadog	NASDAQ: DDOG	IT	Watchdog	Machine learning for application performance monitoring	54.0x
Mohawk	NASDAQ: MWK	Consumer	AI Mohawk E-commerce Engine	Predicts consumer goods trends via NLP	2.6x
Relay Therapeutics	NASDAQ: TXG	Healthcare	Dynamo Platform	Machine learning analyzes protein motion to sample a broader range of protein fragments	55.5x

Source: PitchBook | *As of December 31, 2020

Opportunities

AI-first banking. Banks are among the earliest adopters of AI technologies, although they have started with pilot projects in narrow fields including chatbots and back office RPA. We believe that 20%-40% of banking functions can be automated using existing datasets in banking, investment management, and insurance, with specific emphasis on the front and middle offices. The scale of this opportunity justifies a trillion-dollar value creation opportunity and thus a $100.0 billion revenue opportunity for AI vendors, which should be able to charge 10% of the value they create. Banks are updating their tech stacks to enable AI from their core systems through to their customer-facing portals. At the core level, banks are shifting to the cloud and outsourcing large-scale data integration projects, resulting in a Series B extension and 4.5x valuation step-up for cloud-native banking PaaS startup **Thought Machine** in Q4. At the customer level, banks are using AI to power consumer applications, leading to a unicorn valuation for lending analytics startup **Blend** in its Series F. The urgency of these projects has been heightened due to COVID-19, and we believe there is a dearth of startups addressing the opportunity relative to its scale.

ML-based chargeback guarantees. ML has become table stakes in fraud prevention, but legacy vendors struggle to tune their models to diverse customer environments and rely on human review and rules engines in many cases. ML models can be customized for different customer channels to achieve superior performance and offer guaranteed fraud reduction. Vendors with ensembles of multiple models can adapt to anomalous situations, and we believe startups are best positioned to address a range of customer types. Startups have already created a new category in chargeback guarantees, which is led by **Signifyd** and **Riskified**. Most vendors in the niche are private, are growing faster than the fraud prevention market overall, and have reached breakeven financially, according to market research.[12] **Signifyd** has proven that manual review offers no improvement over its ML. Our prediction that more unicorns would be created in the space has been borne out by **Forter**'s Q4 Series E. Emerging entrants in the space include **Bolt Financial**, **Apruvd**, **Vesta Payment Solutions**, and **ClearSale**.

Computational biology: AI research in biology is advancing rapidly, directly supported by VC. We believe computational biology offers a trillion-dollar opportunity for materials and healthcare innovation. Automation of CRISPR microbe editing can fundamentally change cellular composition to improve industrial materials, food quality, and health outcomes. The vast number of simulations needed to simulate cellular structure alterations require AI & ML to analyze and determine optimal combinations. In Q4, Valo Health raised a $190.0 million Series B to detect novel connections between genetic markers and disease. In the same field, pharmaceutical giant Bayer validated the deep learning approach of **Recursion Pharmaceuticals** with a $50.0 million investment and $30.0 million payment for an AI partnership. Successful discoveries in the category are being spun off as standalone drug developers, including **CereXis** from **Recursion** and **X37** from **Atomwise**. Computational biology is increasingly attracting leading computer scientists to genomics. China is making genomics a national priority, along with AI, creating robust VC activity in the space. 10X Genomics' and **Relay Therapeutics**' strong performances in public markets have given late-stage investors additional incentive to back companies without a clear path to profitability.

AI-enabled cancer screening: To date, AI-enabled cancer therapy has been held back by excessive costs and lack of data. Doctors are now using AI to decide on the right therapy for cancer based on detailed analysis of individual risk and data on deterioration. Researchers

12: "Aite Matrix: Global Chargeback Guarantee Vendors," Aite Group, November 2020.

are continuing to achieve outstanding results in breast cancer screening, with a recent model from MIT called Mirai setting new records compared to previous deep learning methods. Over the next few years, preventive risk screening may gain the data needed to become an essential process. Classifiers can be trained on samples of thousands of tumors to find new early warning signs in methylation profiles, along with a confidence score for inexact matches, with up to 150 cancer entities able to be detected currently. As an example, **PathologyWatch** has used $5.0 million in VC funding to acquire the 30,000 images of basal cell carcinoma needed to train a highly accurate AI model, providing a template for other startups to specialize in specific types of cancers. A raft of startups is pursuing this rapidly changing field, with some specializing in individual cancers, including **AI Therapeutics**, **PathAI**, **Paige**, **Kheiron Medical Technologies**, **Whiterabbit**, **Ezra AI**, **Linkingmed**, **Ibex Medical Analytics**, **NIRAMAI Health Analytix**, **DermaSensor**, **InterVenn**, **Elypta**, **VitaDX**, and **Iterative Scopes**. We believe the diverse nature of cancers may enable scalable companies in each variety. M&A activity in oncology has been robust, and diagnostics companies may comprise the next wave of acquisition targets for pharmaceutical companies and diagnostics providers.

RevOps: Sales & marketing automation is one of the leading sources of exits in AI & ML, but startups have struggled to scale given the strength of incumbents in the space. Increasingly, enterprises are creating a single department called revenue operations, which integrates sales, marketing, and customer success with a single IT backend and set of common industry KPIs. Customers can connect various data sources including CRM, social media, ERP, and third-party analytics platforms into a central platform that uses ML to normalize disparate data streams and uncover correlations among them. This enables organizational restructuring to unify sales, marketing, and customer success under a chief revenue officer. Market research studies find that organizations that pursue RevOps grow sales 3x faster than the control group.[13] The challenge with building this function is integrating disparate sources of data to build a department that is greater than the sum of its parts. Graph databases use ML to find correlations between data points and label data for predictive analytics. This area has been fruitful for startups since they build on top of sales & marketing leaders such as Salesforce and HubSpot and address a pain point for enterprises. Already in Q4 2020, Salesforce has announced its Revenue Cloud to address the RevOps opportunity, although it offers little new functionality. Among startups, **Zeotap** recently achieved a 5.0x deal size step-up in its Series C to merge CRM data with digital fingerprints in a graph database for predictive analytics. We believe the novelty of the technology can support further innovation.

Machine learning for e-commerce recommendations: In e-commerce, A/B testing often drives recommendations, which can take time and create misleading results. Effective recommendations are required as more retail activity moves online. In Q3 2020, US e-commerce sales grew 36.7% YoY to 14.3% of total retail sales, according to US Department of Commerce data. Machine learning can learn from customer behavior in real-time personalized e-commerce experiences through discounts, marketing automation, and recommendations without the need to conduct time-consuming and inefficient A/B tests. We believe that startups are predominantly pursuing this route, though **AWS** is making it available through Sagemaker and likely as part of its own product recommendations. Seed-stage startup **Machine Labs** uses machine learning for **Shopify** stores only and has achieved high traction through the pandemic. In Q4, visual AI recommendation startup **Syte** raised a $40.0 million Series C, citing 22% QoQ growth in 2020. We believe that further startups can be created in this area.

13: “Revenue Operations and the CMO: Game Changer or Game Over?” Sirius Decisions, June 2019.

Digital avatars: We believe that AI & ML has an important role to play in the creation of a metaverse, or virtual world in which gaming and social media can come together to create novel interpersonal and entertainment experiences. The high traction of digital avatar-based entertainment experiences during the pandemic, including virtual Fortnite concerts and Animal Crossing, shows that digital experiences can substitute for in-person events. AI & ML is necessary for the computer vision behind realistic avatars. Game development studio **Singularity 6** raised a $16.5 million Series A from Andreessen Horowitz to build virtual societies including AI characters, although the company has not released its first title yet. In Q3, **Hour One** emerged from stealth with a $5 million seed round to create AI replicas of people to act in their place in digital forums and are finding traction in sales & marketing applications. In Q4, video game unicorn **Roblox** recognized the opportunity in digital avatars with its acquisition of **Loom.ai**, which we previously identified as an innovator in this category.

Grid automation: A lucrative use case for AI & ML is improving the optimization of the energy grid by an order of magnitude due to the ability of automation to coordinate more resources in real time than existing market-based systems allow. Coordination of power generation resources is a good problem for ML given the predictability of supply and demand. Load and generation forecasting models can be trained on historical and real-time power transmission data, though these datasets are not exclusive. In 2017, **Google** subsidiary **DeepMind** proposed to UK utility **National Grid** to save 10% of its power generation through grid optimization, though the deal eventually fell through. In Q3, **Google**'s venture arm Gradient Ventures invested in a Series A for demand forecasting startup Myst.AI, demonstrating its parent company's strategic commitment to the opportunity. Among utilities, US utility **Duke Energy** plans to deploy AI & ML over the next five years. ML-first startup **Invenia** is engaged with Independent System Operators in the US and receives payment when it makes an accurate prediction. **Opus One** is pursuing the opportunity to create models for clean energy resources and partnered with a Scottish utility in Q4 to integrate clean energy production data into a grid planning model. We believe this opportunity will progress slowly but could see a winner-take-all opportunity.

Predictive maintenance: Equipment maintenance efficiency in many applications has been demonstrated to improve by over 10x due to ML algorithms, which could contribute to a quantum leap for the IoT value proposition. For example, cognitive analytics developed at **SparkCognition** improved the speed of predictive analytics for pump maintenance by 20x with accuracy of 99% relative to existing sensor systems. AI & ML algorithms from **Clobotics**, deployed via drone inspections, increased wind turbine inspection efficiency by 10x with 95% accuracy. Industry forecasts anticipate the market growing at a 39% CAGR out to a $23.5 billion market in 2024.[14] Predictive maintenance has been a leading AI use case for **Palantir** and **C3.ai** in their investor presentations to public markets. In Q4, corporate investor Yamaha Motor Ventures led a Series A for no-code AI startup **Canvass Analytics**, demonstrating the strategic value of the technology to industrial incumbents.

Considerations

Advanced AI & ML approaches can have excessive fixed costs: We believe AI companies commonly spend 25% or more of revenue on cloud resources. Training a single AI model can cost around $1 million in compute resources. Models are often retrained with new data, making this a recurring cost. Additionally, images, audio or video data incur storage costs. Lastly, transferring trained models across cloud environments incurs transfer costs.

14: "Predictive Maintenance Market Report 2019-2024" IoT Analytics, October 2019.

These costs can flow to costs of goods sold if model training is required for each customer engagement. Investors should question how startups utilize cloud resources in their model training and inference.

Overfit models: With large datasets, models can be trained extensively to the point where it has too many features to make predictions on novel datasets. As a result, some models trained on quality datasets may not have "model-market fit," even if the software solution itself appeals to customers. Startups should engage in early stopping for neural network training and utilize small datasets to avoid this problem.

Models trained on racially biased data can contribute to racially biased decisions in financial services, HR automation, healthcare, transportation optimization and consumer AI: Racial bias can emerge in models from unexpected sources, such as healthcare spending data, due to sampling bias and unforeseen features, such as correlating health spending to health risk. For startups, the public relations consequences of a biased model are material, and racially biased models must be retrained on fresh data with additional feature engineering and testing. The effect of such a restart can amount to millions of dollars. Thus, it can be cost effective to integrate explainability and ethical AI governance frameworks from an early stage to ensure that AI algorithms do not incorporate racial bias in their features nor predictions. The negative consequences of bias will gain more public attention as AI & ML are increasingly used in lending and hiring decisions.

Privacy concerns: The retail and advertising industries are another AI & ML end market that is heavily rooted in the use of personal data. With recent lapses in protection and governance of personal data, there will likely be more concerns surrounding data-science products in consumer-centric use cases. Regulatory changes, such as the EU's implementation of GDPR this year, have far-reaching effects into how AI & ML-focused companies will have to operate, especially if other jurisdictions follow suit. Adherence to these higher standards will likely add operational costs and could slow implementation in some geographies. This can also affect individual behavior without the intervention of government, as potential customers may choose an alternative because of privacy concerns.

COVID-19 has exposes practical limitations for healthcare applications: Relative to AI, statistics and rules-based models have proven more useful for pandemic measurement and response. Leading AI researchers have found success in studies with small samples, but these have not formed a core part of pandemic response. Understandably, the accuracy of natural language processing models for virus tracking that were able to detect the virus early have broken down at scale. We believe that AI & ML will form a core part of vaccine research but its utility as a decision-making tool for global leaders is still far off.

Outlook

Winner-take-most opportunities to emerge in numerous verticals: These verticals include sales & marketing data integration, small to medium-sized business (SMB) fraud prevention, predictive maintenance, VR-based social media, and core banking PaaS. We believe the network effects of AI & ML can produce sustainable competitive advantages when predictions improve with more data, proprietary feature sets are generated, and emerging datasets are applied to more complex problems than incumbents can solve. With the ability to pull in data from multiple applications and customers and use it to develop proprietary features, we believe that startups can build flywheels in large industries. We view banking, sales & marketing, predictive maintenance, and SMB fraud prevention in particular as

areas where incumbents may not be focused and may be limited in their ability to develop feature sets that are extensible across the entire industry. The result of creating sustainable competitive advantages in these spaces should be high valuation premiums in public markets or strategic acquisitions. Other verticals and product categories may face barriers to entry that AI-first approaches may not be able to overcome.

M&A trends to continue in ad targeting, information security, sales & marketing, and HR automation: Incumbents recognize that internal AI systems may be easier to buy than build and are purchasing AI startups at significant premiums. **Cisco**, **Medtronic**, **Johnson & Johnson**, **Shell**, and **Verizon** have all made acquisitions in the past year, demonstrating the appetite for AI automation among legacy providers. In information security, incumbents are already AI-enabled and are actively adding to their product suites in adjacencies. In each of these categories, acquisition activity has been consistent throughout 2019 and 2020, and we believe there is a strong pipeline of further targets.

Venture investors to become more comfortable with AI-first business models in a downturn: AI-first software startups have suffered from the perception that gross margin profiles are inferior to those of SaaS startups with lower customization and data requirements. The decrease in growth for some SaaS startups in 2020 may encourage investors to look to AI-first business models with lower gross margins but higher business value and winner-take-all potential. Early-stage investors may be more patient with the time and cost needed for seed-stage startups to develop AI models given the time needed for the economy to recover from the pandemic. Economic downturns tend to reveal which solutions provide the best value for customers, and we believe AI-first software platforms are likely to win over rules-based SaaS in the long-term.

SEGMENT DEEP DIVE

AI semiconductors

Overview

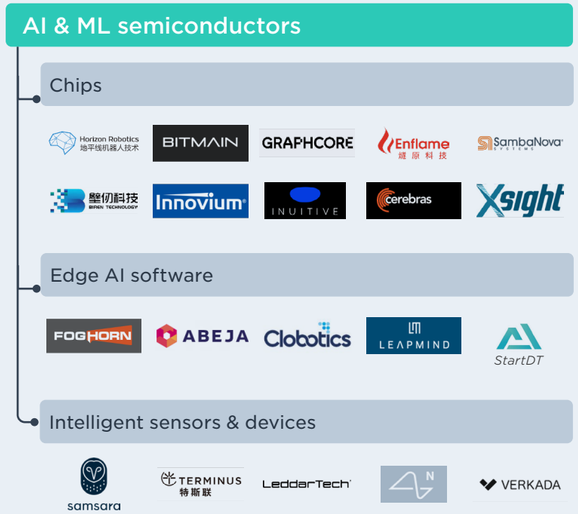
AI & ML model training and inference require hardware with maximum computational efficiency and customized processing for AI calculations. AI requires highly parallelizable, predictable computations that benefit from specialized chips. Using general purpose chips can cost tens to thousands of times more than using an AI chip for these purposes. As a result, a segment has developed for both the design and software-based optimization of computing hardware, including both semiconductors and sensors. Subsegments include:

AI chips: types of computer chips that attain high efficiency and speed for AI-specific calculations. AI chip optimizations include parallel calculations, low precision calculations to reduce transistor count, AI & ML algorithm compression, and using novel programming languages built specifically to efficiently translate AI computer code for execution on an AI chip. AI chips are tens or even thousands of times faster and more efficient than CPUs for training and inference of AI algorithms. Technologies within this category include:

- Graphics processing units (GPUs)
- Field-programmable gate arrays (FPGA)
- AI-specific application-specific integrated circuits (ASIC)
- Neural network processors

Edge AI: Compression algorithms that optimize AI & ML models for deployment within various semiconductor environments and edge devices. Technologies in this category include:

- Post-training neural network binarization and quantization algorithms
- Quantization aware training



- Ensemble boosting algorithms
- Edge AI application programming interfaces (APIs) and software development kits (SDKs)

Intelligent sensors & devices: Devices that are optimized to run AI & ML models. This segment contains technologies including:

- **Sensors:** Devices that measure specific parameters in real-world conditions including environmental conditions, motion, images and chemical levels
- **Sensor systems:** Assemblages of sensors and chipsets that empower specific IoT use cases such as smart buildings, industrial IoT and connected vehicles

Industry drivers

Massively expensive computation: The amount of compute being used in large AI training runs has been increasing exponentially. In 2018, **OpenAI** calculated that the compute needed to train breakthrough models including AlphaGoZero and Neural Machine Translation was doubling every 3.4 months. The cost of the hardware needed for **Google** to train AlphaGoZero has been estimated at $35 million.[15] **Google** benefited from its proprietary tensor processing units, but similarly ambitious and practical projects undertaken by other companies may require even higher costs to replicate.

Research finding that deep learning improves at scale: The more compute that is used, along with optimal model architecture and large data training sets, the more accurate deep learning algorithms become. This result has been empirically demonstrated by ML researchers and means that additional breakthroughs in accuracy will scale in parallel with computing power.[16]

Demand shifting from discrete to integrated processors: AI chips have focused on the data center but are increasingly shifting to the edge to support emerging use cases including smartphones, tablet computers, wearable devices and connected automobiles. This is leading to integrated processors growing faster than discrete processors.

Market size

We estimate the AI & ML semiconductor market to reach $20.8 billion in 2021, achieving 44.5% growth over 2020 as a result of COVID-19. Over half this total can be attributed to mobile phone ASICs, which will decline as a percentage going forward. We expect the market to grow at a 17.3% CAGR from 2021 to 2024, resulting in a $33.7 billion market. Our estimate includes all AI semiconductor varieties, including field-programmable gate arrays, graphic processing units, microcontrollers, and ASICs. To be included in this estimate, the primary function of these chips must be to enable AI processing. We forecast the highest growth segments in this market to be field-programmable gate arrays and microcontrollers, all of which we anticipate will grow at CAGRs over 50.0% from 2021 to 2024 based on edge AI use cases emerging.

Disruption potential

The graphic processing unit market faces disruption as purpose-built chips for AI & ML training are just entering the market. The Chinese government perceives the AI chip

15: "How Much did AlphaGo Zero Cost," Dan Huang, 2018

16: Joel Hestness et al, "Deep Learning Scaling is Predictable, Empirically," 2017

industry as a clean slate where it can compete for global supremacy in a field where it has historically been a net importer. While graphic processing units have become cornerstones of modern AI training efforts, they are power-intensive and thus costly. Edge computing is a trend that we believe will push AI calculations closer to where they are needed, including local data centers and edge devices. As a result, we believe that graphic processing units could lose market share to chips with higher efficiency and greater utility at customer sites.

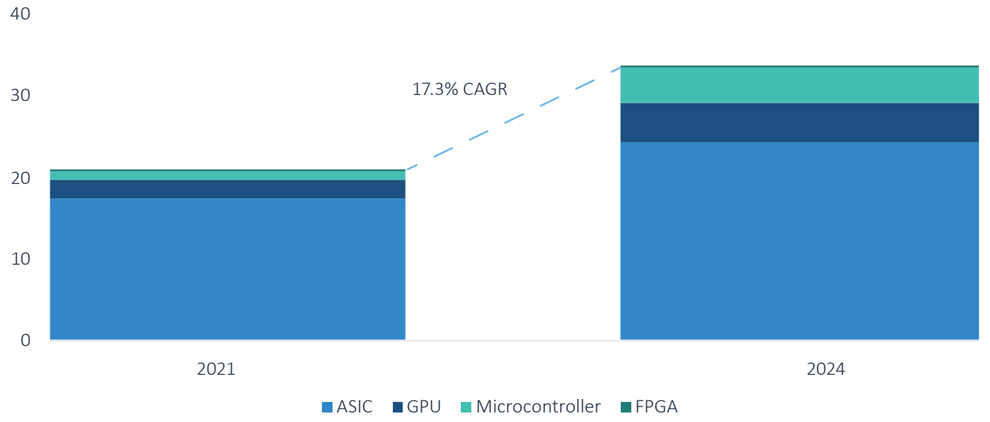
Business model

AI semiconductors are currently deployed through unit sales and usage-based pricing from cloud hosts. For example, **AWS** sells usage-based instances of **Nvidia Tesla** V100 GPUs for AI & ML workloads. The usage-based pricing of GPU instances enables around a one-year payback for **AWS** for the underlying hardware, assuming constant usage of these GPUs, given current pricing of around $11,000 for a standard **Nvidia Tesla** V100. AI & ML developers can thus access GPUs without high capex.

VC activity

VC investment in AI chips has remained muted relative to the large M&A deals for **Arm** and **Xilinx**. VC deal count has declined for two straight years, and 2020's deal value of $3.4 billion has barely exceeded 2018's total of $2.9 billion. In Q4, ASIC design startup Graphcore achieved a 1.4x valuation step-up with a $217.4 million round, bringing the company's cash on hand to $444.0 million, which is sufficient to tape-out a chip at the leading node, according to our research.[17] Chinese investment remained high, with

Figure 19. AI SEMICONDUCTORS MARKET SIZE ($M)



Source: PitchBook and Gartner

COMMON INDUSTRY KPIS

Chip performance

- Watts/GFLOPS
- $/GFLOPS
- Inference accuracy
- Transistor density

Unit economics

- Foundry sale price

Benchmarks

- Training efficiency and speed
- Inference efficiency and speed
- Generality

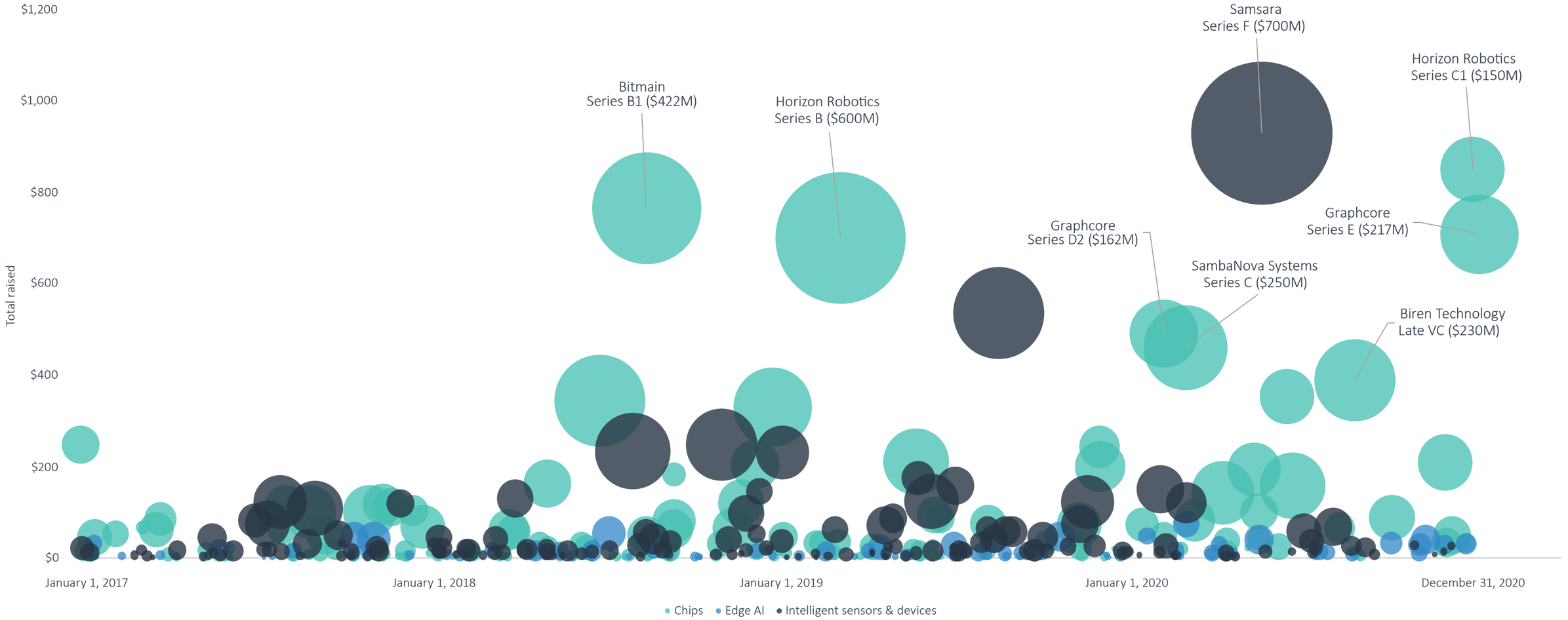
automotive ASIC maker **Horizon Robotics** raising a $150.0 million round and China-based investors leading a $106.0 million round in Israeli startup Inuitive. We believe there are numerous underfunded opportunities in this segment and burgeoning commercial traction.

The announced acquisition of **Xilinx** (NASDAQ: XLNX) by **AMD** (NASDAQ: **AMD**) promises to reshape the AI FPGA (field programmable gate array) landscape. **AMD** plans to buy **Xilinx** for $35.0 billion or 38.2x EV/trailing EBITDA. The acquisition is part of a horizontal integration strategy that moves **AMD** further into AI inferencing, among other markets, via **Xilinx**'s AI Inference Acceleration chips. **AMD** is currently a market leader in data center GPUs used for training. Thus far in Q1 2021, startup FPGA maker **Achronix Data Acceleration** has exited via SPAC at 32.1x EV/forward EBITDA, further acknowledging growth forecasts in the space. AI ASIC makers in China continued to achieve outsized exits, with **Bestechnic** (SHG: 688608), a voice recognition ASIC startup, achieving a valuation of $3.0 billion in its IPO after **Cambricon Technologies**' lofty exit in Q3 2020. China-based startups are achieving a low cost of capital via public markets while US and European startups are struggling to achieve significant commercialization or exits.

17: "Graphcore Raises $222 Million in Series E Funding Round," AiTHORITY, December 29, 2020.

Figure 20.

AI semiconductors VC landscape ($B)



Source: PitchBook

Note: The left axis indicates total VC raised as of deal date. Bubbles indicate amount raised.

Figure 21.

Notable AI semiconductors VC deals

COMPANY	CLOSE DATE	SUBSEGMENT	DEAL TYPE	DEAL SIZE ($M)	LEAD INVESTOR(S)	VALUATION STEP-UP
Graphcore	December 29, 2020	Chips	Series E	$217.4	Ontario Teachers' Pension Plan	1.4x
Horizon Robotics	December 22, 2020	Chips	Series C1	$150.0	Capital Today, 5Y Capital, Hillhouse Capital Group	N/A
Biren Technology	August 20, 2020	Chips	N/A	$230.0	Hillhouse Capital Group	3.4x
Syntiant	August 4, 2020	Chips	Series C	$35.0	Applied Ventures, M12	1.6x
Enflame Technology	May 7, 2020	Chips	Series B	$99.0	SummitView Capital	1.8x

Source: PitchBook

Figure 22.

Notable AI semiconductors VC exits

COMPANY	CLOSE DATE	SUBSEGMENT	EXIT SIZE ($M)	ACQUIRER/INDEX	VALUATION STEP-UP	EV/TRAILING REVENUE
Cambricon Technologies	July 20, 2020	Chips	$3,290.8	Shanghai Stock Exchange	1.3x	59.4x
Xnor.ai	January 8, 2020	Edge AI	$200.0	Apple	3.2x	N/A
Habana Labs	December 16, 2019	Chips	$2,000.0	Intel	N/A	N/A
Mnubo	July 12, 2019	Edge AI	$77.6	Aspen Technology	N/A	N/A
DeePhi	July 17, 2018	Edge AI	$251.9	Xilinx	N/A	N/A

Source: PitchBook | *As of December 31, 2020

Figure 23.

Key VC-backed AI semiconductors companies

COMPANY	TOTAL VC RAISED ($M)*	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	MOST RECENT LEAD INVESTOR(S)
Samsara	$930.0	Chips	VS2 Vision System	Custom machine vision cameras	Dragoneer Investment Group, Warburg Pincus and General Atlantic
Horizon Robotics	$850.0	Chips	AI Brain Processing Unit (BPU)	Customized for ADAS and autonomous vehicles	SK Hynix, SK Group
Bitmain	$764.8	Chips	Antminer	Custom cryptocurrency miner	Crimson Capital China
Graphcore	$708.1	Chips	Intelligence Processing Unit	High efficiency on small batches of data	Mayfair Equity Partners
Terminus Technologies	$536.3	Chips	Zhihui X9 server	Edge computing server for object recognition	China Everbright
SambaNova Systems	$460.6	Chips	Software-defined devices	Algorithmic optimizations and custom board-based hardware	BlackRock

Source: PitchBook | *As of December 31, 2020

Figure 24.

Key AI semiconductors incumbents

COMPANY	TICKER	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	EV/FORWARD REVENUE*
Advanced Micro Devices	NASDAQ: AMD	AI chips/Edge AI	Epyc Processors	Competitive AI GPUs	12.6x
Intel	NASDAQ: INTC	AI chips/Edge AI	FPGAs for Artificial Intelligence	Leading AI FPGAs	2.8x
Lattice Semiconductor	NASDAQ: LSCC	AI chips/Edge AI	sensAI Stack	Relevant AI FPGAs	15.6x
Nvidia	NASDAQ: NVDA	AI chips/Edge AI	Tesla V100 GPU	Leading AI GPUs	21.7x
Xilinx	NASDAQ: XLNX	AI chips/Edge AI	Vitis AI	Competitive AI FPGAs	11.3x

Source: PitchBook | *As of December 31, 2020

Opportunities

Field-programmable gate arrays (FPGAs): FPGAs enable efficient inferencing and can be a steppingstone toward custom AI chips for specific applications. Semiconductor product lines are bifurcated into separate chip designs for inferencing and training so as to optimize for each of the workload types. GPUs have commonly been deployed for AI training, but we believe FPGAs are around 4x and up to 10x more efficient and comparable in speed for inferencing when produced at the same node, due to their high volume of distributed memory blocks.[18] Intra-chip memory access provides major efficiency and speed improvements compared to communication with off-chip memory. Furthermore, interconnections between memory and compute can be configured by programmers to accommodate specific algorithms. FPGAs are more generalizable than ASICs, which are more efficient for specific use cases and 100x faster and more efficient than CPUs produced at the same node.[19] **Intel** acquired FPGA provider **Altera** for $17.0 billion in 2015, but it relies on its Xeon CPUs for AI inferencing. FPGA design startup **FlexLogix** reports that 2020 presented an inflection point in FPGA adoption, with its chipmaker customer count on the verge of doubling from a sizable base. It has also announced a proprietary chip that competes against **Nvidia**'s Jetson line.

Intel Capital has also invested in software for FPGAs via University of Toronto spinoff **LegUp Computing**, showing that the FPGA titan has interest in ongoing innovations in the area. Given the recent acquisitions by **Intel** and **AMD** (**Xilinx**) in this space, **Nvidia** may be motivated to make an acquisition to compete. We believe AI FPGAs can scale to $1.0 billion in sales by 2024.

Neural network ASICs: ASICs have lower barriers to entry than GPUs or FPGAs. Custom-built ASICs can achieve high efficiencies based on hardwired circuitry that cannot be configured by developers after production. Gartner estimates that 15% of AI chips will be dedicated AI ASICs by 2024, up from under 1% in 2019. We believe data center AI ASICs can scale to $4.0 billion in sales by 2024. This field features startups that have developed leading designs, including **Cambricon Technologies** and **Horizon Robotics**. In Q4, Graphcore's VC mega-deal demonstrated it has raised sufficient funding to produce its intelligent processing unit (IPU) and accelerate AI-specific workloads. **Google** has also entered the market with its tensor processing unit. In China, conglomerates **Baidu**, **Alibaba**, **Tencent**, and **Huawei** are leading innovators. We believe startups in this space may be able to scale given the lack of incumbents and the unique needs of diverse use cases including on-premises AI training and inference, drones, object recognition, and sound recognition.

Neuromorphic hardware: This novel architecture is inspired by neurobiological architectures to utilize spiking neural networks and run deep neural networks at low power. The approach can be used to accelerate image and sound recognition in edge devices. Deep learning's resemblance to the brain's patterns can benefit from hardware that mimics neurons and presents the opportunity for a category of chips not yet dominated by incumbents. In particular, memory is an area that can mimic the function of synapses with high interconnectivity. Neuromorphic designs remain in the early prototype stage and may not be commercialized for four more years. Industry estimates have optimistically pegged the neuromorphic hardware market at $5.0 billion by 2029. **IBM**, **Intel**, and publicly traded **BrainChip** are currently developing versions of this chip. Startups making advances in this

18: "AI Chips: What They Are and Why They Matter," Center for Security and Emerging Technology, April 2020.
19: Ibid.

space include **aiCTX**, **Nepes**, **Vicarious**, **Celepixel**, Robosensing, **Memry**, Symetrix Weebit, **Knowm**, and **Weebit Nano**. We believe the nascent nature of this architecture may make it an opportunistic area for startups.

Edge AI chips leveraging emerging TinyML algorithms: IoT has struggled to scale across enterprises given the uneven ROI for individual projects and challenges integrating sensors with the cloud. Commercialization of miniaturized ML algorithms could unlock value for common IoT use cases, including predictive maintenance, supply chain optimization, and connected healthcare. Existing IoT hardware lacks the combination of energy efficiency—which we believe should be under 50 µA/MHz for TinyML algorithms—and memory—which we believe should be over 20 MB. Emerging approaches to energy-efficient ML edge computing by recently seeded startup **Edge Impulse** have claimed to reduce the memory needed for neural network classifiers to under 10 KB, allowing ML algorithms to be embedded on common ARM Cortex microcontrollers. In Q3, South Korea-based startup **Nota** raised a Series A from a group including Samsung Venture Investment to compress AI models for mobility, drones and robotics. Use cases gaining traction include mobile device voice recognition and smart cameras, in addition to mobile phone processors more generally.

Hardware companies innovating in this space include UniSound, which has built a chip that can interpret voice using deep neural networks in edge applications, and **GreenWaves**, which is developing a chip that can communicate a variety of high-bandwidth signals with low power requirements. In August 2020, California startup Kneron announced an edge AI chip that outperforms similar **Intel** and **Google** edge chips in power efficiency by 2-4x, with focuses on image and audio recognition. The company that can process AI algorithms in battery-powered devices with no need for cloud connectivity can unlock a $4.2 billion market for AI microprocessors by 2024. Furthermore, incumbents are actively investing in the space via the TinyML Foundation including **Arm**, **Alphabet**, and **Qualcomm**, and we expect ample M&A opportunities in the medium term.

Considerations

The slowing of Moore's Law threatens deep learning progress: Moore's Law observes an empirical and finite relationship in which transitory density in an integrated circuit doubles every two years, but the time this actually takes has extended. The costs of continuing Moore's Law are justified only because it enables continuing chip improvements, such as transistor efficiency, transistor speed and the ability to include more specialized circuits in the same chip. The past 10 years have seen improvements in transistor density slowing. Today, leading chips contain billions of transistors, but they have 15x fewer transistors than they would have if Moore's Law had continued unabated. For this reason, there are constraints on how much innovation is possible in raw processing power.

Chip development is expensive: The cost to design a leading semiconductor at a 7nm node is $300 million, including software, physical design, testing and tapeout. Chip companies are accustomed to releasing new chips on a regular cadence, resulting in an astronomical burn rate for startups.

The ongoing trade war between the US and China threatens the scalability of semiconductor startups: Due in part to the trade war between the US and China, the semiconductor market shrunk in 2019. Leading fabricators are based in China and Taiwan and fell victim to tariffs. Growth in semiconductors is rarely linear, however, so this decline may have been a short-term blip, especially if fabricators are able to diversify operations across geographies.

Semiconductor startups may not scale without a commercially viable architecture: Some advanced semiconductor startups lack the ability to manufacture a cost-competitive chip due to a lack of commercially viable architecture. We believe chip designs with the potential to commercialize and compete with products on the market today must feature the following characteristics:

- Pricing of under $0.01 per giga floating-point operations per second (GFLOPS)
- Sufficient funding to bring a chip to production
- Interconnect architectures to optimize data flow between processing and memory
- Compatibility for training and inference software because some chips require separate software for both processes, which has knock-on effects for software design
- A fabrication node at 7nm or below
- Integrated developer kits
- Validation from cloud hosts

Large incumbents present high barriers to entry: Even with a novel design and enough capital support, companies will face competition from large, established hardware manufacturers. **AMD** and **Nvidia** have a duopoly over the GPU market, and the FPGA market is dominated by **Xilinx** and **Intel**. China-based chipmakers specialize in ASICs. **Nvidia** has gained a lock-in effect for AI developers via its native developer environment CUDA, for translation of AI code to machine code, and RAPIDS for hardware acceleration. Other technology giants have competitive training offerings, with **Google**'s tensor processing units that are optimized to run functions on its TensorFlow framework. **Tesla** has announced a customized chip that will more efficiently run its cars, and **Amazon** launched an AI inferencing chip called Inferentia in December. The strong position of these chips creates barriers to entry for the cloud data center and healthcare computing markets.

Outlook

Incumbent chipmakers to be pressured to acquire FPGA, ASIC and distributed memory startups: We believe that the dominant leaders in GPUs, CPUs and FPGAs lack some of the necessary innovations in distributed memory to power an AI revolution cost-effectively. They will need to seek innovation externally to obtain market leadership. Our thesis is playing out with **Nvidia**'s announced acquisition of AI microcontroller design firm **Arm** and **AMD**'s announced acquisition of FPGA leader **Xilinx**, which suggests that incumbents are building diversified platforms of AI chips to address emerging customer needs. **Intel** has directly taken on **Nvidia** in AI processing with its purchase of **Habana Labs**, creating a new battlefield in ASICs that may pressure **Nvidia** and **AMD** to gain a foothold. Additionally, incumbents hungry

for growth will look to application-specific chips that can scale across emerging use cases including mobile robotics, 5G and IoT devices. We have not seen semiconductor startups scale before selling to incumbents, making consolidation a probable outcome for companies in this market.

China is likely to develop leading custom ASICs, giving it a leg up in the race for AI semiconductors: China-based conglomerates and startups have been able to engineer chips to perform specific tasks at a world-class level, supplying outsized investment totals to startups relative to the US and Europe. **Horizon Robotics** has developed a custom ASIC for computer vision for ADAS and autonomous vehicles that we believe is globally competitive for original equipment manufacturer and Tier-1 automakers. **Bitmain** has outcompeted **Nvidia** in crypto mining, and the company is working on AI ASICs. China-based startups are also focusing on sound processing, smart home and IoT applications. These chips may become core parts of new use cases, since US innovation lies primarily in data center GPUs and training. China's relative lack of price sensitivity in chip development may lead to more rapid development in these niche areas and pressure on the US government to support domestic chipmakers.

Cerebras to pursue independent growth: The future of AI chip startups has been guided by **Intel**'s $2 billion acquisition of **Habana Labs** and will be tested by the commercialization of **Cerebras**' Wafer-Scale Engine, the largest chip ever built with 1.2 trillion transistors. **Cerebras** has made several scientific breakthroughs to build its chip designed for deep learning and has begun delivering prototypes. To this point, the chip is being used in research applications. If it can find a home with hyperscalers, then it may be able to gain self-sufficiency. Given **Nvidia**'s dominant market position and recent acquisition of **Mellanox**, we believe the most likely outcome is an acquisition in the scale of **Habana Labs**.

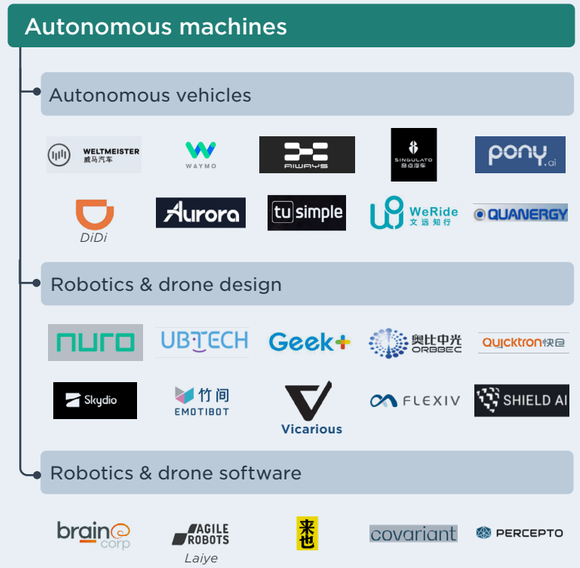
SEGMENT DEEP DIVE

Autonomous machines

Overview

Autonomous machines can perform tasks in human-present environments without explicit human control. These machines synthesize ML, computer vision, and datasets of the physical world such as navigation. The segment requires the design of complex hardware with software “brains” and is therefore valued differently from AI & ML software as well as semiconductors. Current applications of AI in robotics lack “common sense” to learn from their environment. This segment does not include optimization of existing robots, including those used in manufacturing and supply chains. Subsegments include:

- **Autonomous vehicle software and design:** software and hardware solutions that enable self-driving or driver assistance capabilities for cars, trucks and other on-road vehicles.
- **Intelligent robotics & drone design:** robotic systems and unmanned aerial vehicles that can operate without human input. AI & ML can be used for learning, control and adaptation of robots. Common applications of autonomous robotics include swimming, carrying objects, picking up objects and putting them down.
- **Intelligent robotics & drone software platforms:** operating systems for autonomous robots. These platforms can include fleet management and predictive maintenance.



Industry drivers

COVID-19 driving autonomous delivery demand: Because of the closure of retail outlets, last-mile delivery requires additional capacity with less cost sensitivity from merchants.

AUTONOMOUS MACHINES

Companies offering autonomous delivery are stepping in to expand the capacity of overstretched delivery networks.

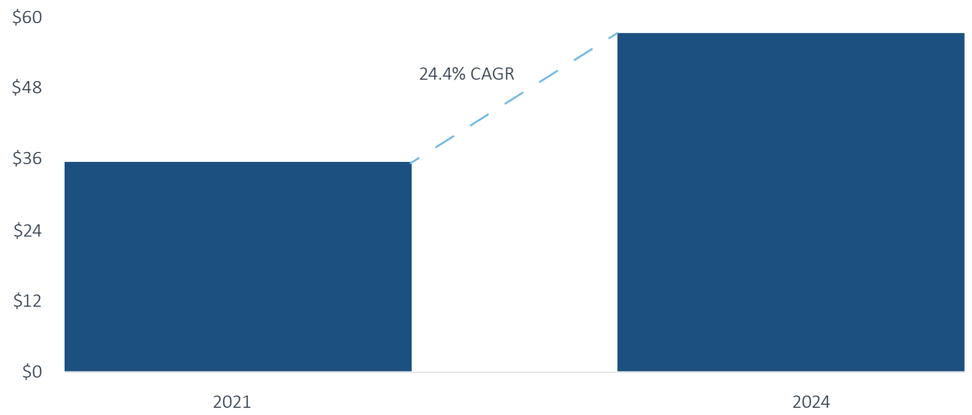
Improving regulatory environment for autonomous vehicles: In March 2020, The National Highway Traffic Safety Administration proposed a new rule for easier certification of autonomous vehicles.

E-commerce growth driving autonomous robotics implementation: E-commerce is projected to grow from 14% of retail to 50% over the next five years, which means warehouses must become more efficient.

Market size

We estimate the autonomous machines market to be $35.5 billion as of 2021 and forecast it to grow at a 17.2% CAGR to $57.1 billion in 2024. This estimate includes commercial and delivery drones, autonomous vehicles and autonomous robots. This estimate has been revised upward significantly based on existing applications of AI in Level 2 and 3 autonomous driving. While the industrial robotics market is larger on its own at over $100 billion, we focus on the smaller opportunity for autonomous robots within the manufacturing, healthcare and supply chain industries. Autonomous driving is expected to grow at a 20.7% CAGR through 2024, higher than for robotics given the challenge of displacing existing industrial robotics in manufacturing and warehouse settings.

Figure 25. AUTONOMOUS MACHINES MARKET SIZE ($B)



Source: PitchBook and Gartner

COMMON INDUSTRY KPIS

Operational

- Number of robots deployed
- Patents
- Product lifecycle
- R&D lifecycle
- Regulatory approvals
- Government partnerships
- OEM joint ventures

Financial

- Addressable markets
- Gross Margin
- Unit sales
- Revenue mix

Disruption potential

Autonomous machines can achieve disruption akin to Uber's displacement of the taxi industry via more efficient coordination of resources and reduction of overhead. In the short term, they can be a force multiplier for human efforts, making existing workforces more efficient. In the long term, they could create entirely new use cases that disrupt existing supply chains and vehicle fleets. Autonomous vehicles, drones and mobile robots may enable 10x increases in equipment investments that displace human-operated versions due to increased uptime and predictive fleet optimization.

Business model

Beyond traditional hardware sales & service business models, autonomous machine suppliers are increasingly delivering robotics as a service to operate with multiple types of hardware. AI & ML algorithms trained for machine intelligence can be deployed on an edge device via software development kits and paid for via recurring revenue. We believe this model is becoming more prevalent among incumbents and is table stakes for startups.

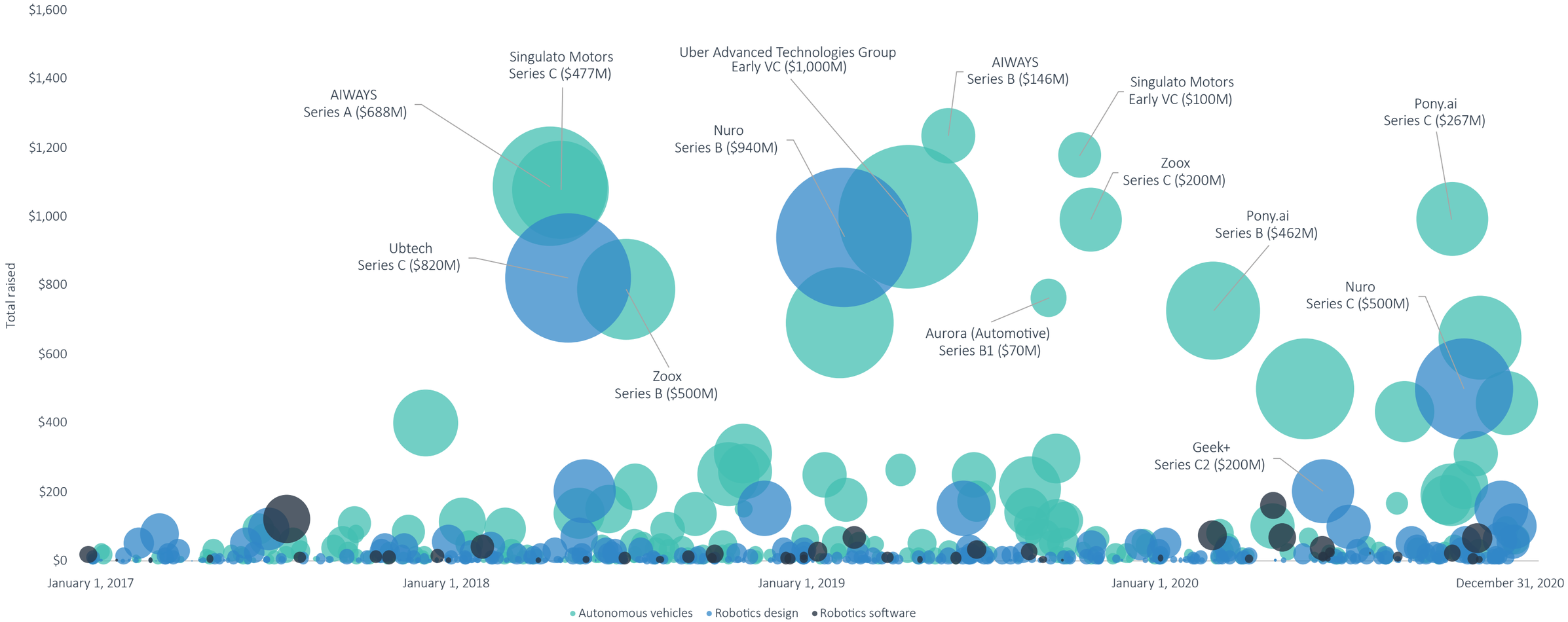
VC activity

Autonomous machines VC deal value had flat growth YoY at more than $6.0 billion, excluding an outlier $3.0 billion investment in **Waymo**. Intelligent robotics companies raised nearly $2.0 billion of this total, growing slightly over 2019. Intelligent robotics had a strong Q4 with VC mega-deals raised by warehouse robotics startup Quicktron, intelligent robot arm maker Flexiv, and delivery robotics unicorn **Nuro**. The Flexiv and Quicktron deals were led by strategic investors, demonstrating both high growth and strategic overlap. **Nuro** nearly doubled its valuation to $5.0 billion after regulatory approvals and continued deployment of its autonomous delivery robots in 2020. Autonomous drones escalated in early-stage funding, with strategic investors backing **Iris Automation**, **Aerobotics**, **Percepto Robotics**, and **vHive** at substantial early-stage valuations. Despite divesting **Boston Dynamics**, SoftBank continued its commitment to autonomous robots with an investment in China-based service robot startup **Keenon**.

Robotics M&A activity ticked up in Q4, beginning what we believe will be a growth phase, while the autonomous vehicle industry continued consolidation. Online grocer **Ocado Group** (LON: OCDO) acquired intelligent robotics startup **Kindred AI** for $262.0 million, a significant premium to its Series B post-money valuation of $85.0 million. **Kindred AI** is applying reinforcement learning—a cutting-edge AI technique that holds promise for robots to learn continuously and function in complex human-present environments—to warehouse robots. In a transaction adjacent to intelligent robotics, Hyundai acquired robotics innovator **Boston Dynamics** for $1.1 billion with a significant strategic premium, after the company sold for $100.0 million in 2017. In autonomous vehicles, Uber (NYSE: UBER) divested its autonomous driving unit in a sale to **Aurora**, a further sign that market leaders are seeking de-risk their substantial investments in the technology.

Figure 26.

Autonomous machines VC landscape ($B)



Source: PitchBook

Note: The left axis indicates total VC raised as of deal date. Bubbles indicate amount raised. Excludes **Waymo** *and* **Weltmeister** *for scale.*

Figure 27.

Notable autonomous machines VC deals

COMPANY	CLOSE DATE	SUBSEGMENT	DEAL TYPE	DEAL SIZE ($M)	LEAD INVESTOR(S)	VALUATION STEP-UP
Flexiv	December 30, 2020	Industrial robots	Series B	$100.0	Meituan, New Hope Group, Yunfeng Capital	N/A
Iris Automation	December 15, 2020	Drones	Series B	$13.0	Sony Innovation Fund, Verizon Ventures and Bee Partners	N/A
Nuro	November 9, 2020	Industrial robots	Series C	$500.0	T. Rowe Price	1.9x
Applied Intuition	October 22, 2020	Autonomous vehicle software	Series C	$125.0	General Catalyst, Andreessen Horowitz, Lux Capital Management	8.8x
Skydio	July 13, 2020	Drones	Series C	$98.3	Next47	1.5x

Source: PitchBook

Figure 28.

Notable autonomous machines VC exits

COMPANY	CLOSE DATE	SUBSEGMENT	EXIT SIZE ($M)	ACQUIRER/INDEX	VALUATION STEP-UP	EV/TRAILING REVENUE
Uber Advanced Technologies Group	December 7, 2020	Autonomous vehicle software	$4,000.0	Aurora	0.6x	N/A
Zoox	June 28, 2020	Autonomous Vehicle Design	$1,300.0	Amazon	N/A	N/A
EHang	December 11, 2019	Autonomous vehicle design	$623.0	NASDAQ	1.6x	58.7x
CTRL-Labs	September 23, 2019	Robotics platforms	$1,000.0	Facebook	4.0x	N/A
Aeryon Labs	January 28, 2019	Autonomous vehicle software	$200.0	FLIR Systems	N/A	N/A

Source: PitchBook

Figure 29.

Key VC-backed autonomous machines companies

COMPANY	TOTAL VC RAISED ($M)*	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	MOST RECENT LEAD INVESTOR(S)
Weltmeister	$3,915.5	Autonomous vehicle design & software	Living Pilot	Strategic partnership with Baidu	SAIC Capital China, Shanghai International Group
Waymo	$3,000.0	Autonomous vehicle design & software	Waymo Driver	First-mover advantage in data collection	Silver Lake Management, Canada Pension Plan Investment Board, Mubadala Capital-Ventures
Nuro	$1,532.0	Intelligent robotics	Self-driving delivery bot	Compact, cabless, on-road solution	SoftBank Investment Advisers
AIWAYS	$1,234.0	Autonomous vehicle design & software	CARL	Autonomous charging robot	Mingchi Fund
Singulato Motors	$1,178.3	Autonomous vehicle design & software	TBD	Acquired Toyota IP	ITOCHU International
Uber Advanced Technologies Group	$1,000.0	Autonomous vehicle design & software	Neuropod DL Inference Engine	Uniform interface for deep learning models from multiple frameworks	Toyota Motor, Denso, SoftBank Investment Advisers

Source: PitchBook | *As of December 31, 2020

Figure 30.

Key autonomous machines incumbents

COMPANY	TICKER	SUBSEGMENT	KEY PRODUCTS	PRODUCT DIFFERENTIATION	EV/TRAILING REVENUE*
ABB Group	NYSE: ABBN	Intelligent robotics	ABB Robotics	Manages manufacturing and production processes	2.3x
Cognex	NASDAQ: CGNX	Intelligent robotics	Vision Guided Robotics	Proprietary machine vision algorithms	18.0x
iRobot	NASDAQ: IRBT	Intelligent robotics	Roomba	AI-powered vacuum cleaner	1.5x
Stryker	NYSE: SYK	Intelligent robotics	Mako Robotic-Arm Assisted Surgery	3D model-based planning, haptic technology, and data analytics	6.9x
Omron	TKS: 6645	Intelligent robotics	Vision Sensors/Machine Vision Systems	Drag and drop programming interface	2.7x

Source: PitchBook | *As of December 31, 2020

Opportunities

New datasets solving autonomous driving edge cases: Vehicle ML algorithms struggle to deal with edge cases including unfamiliar objects and scenarios. The first wave of startups collected much of their data from the perspective of the car's sensors, adding real-world data over time through simulation and testing. Some edge cases, including climate, human behavior, stray objects, and the built environment, have separate datasets that can be used to generate their own models. **Isee** recently raised a $15.0 million Series A at a $50.0 million pre-money valuation from Founders Fund to bring cognitive computing to autonomous trucking, with the goal of replicating humanistic common sense. These solutions require more onboard compute than existing solutions but may be much more effective than the limited neural networks generated by perception data and hard coding. In Q4, **Applied Intuition**, a vehicle simulation startup, achieved an 8.8x valuation step-up into its Series C. The company enables autonomous vehicle design companies to customize simulation environments using real-world and synthetic data to mimic edge cases. We believe point solutions such as these may be easier to commercialize relative to full-stack autonomous driving startups.

Domain randomization: Robots that have been long-deployed in isolated environments must make better use of computer vision to operate in human-present environments. Computer vision models have historically required labeled data from real-world cameras and are trained in real-world environments, but they are increasingly able to simulate camera-based perception in virtual environments, facilitating training across millions of samples. **OpenAI**'s Dactyl hand can complete the Rubik's Cube with a single hand after training in a simulated environment. **OpenAI**'s approach uses reinforcement learning and domain randomization to transfer neural networks learned in simulation to a real robot. We believe domain randomization holds promise for robots to function in complex human-present environments and allows for multiple use cases, including autonomous vehicles and industrial automation. In September, a team of Korean and German researchers used domain randomization to train a robot in curling. The robot beat expert human teams in three of four matches, offering promise for real-world operation in dynamic environments. While this technique is still under development, autonomous robotics startup **Covariant** is commercializing domain randomization for warehouse optimization. Pieter Abbiel founded **Covariant** while serving as Director of the Berkeley Robot Learning Lab at UC Berkeley and after studying under Andrew Ng as a PhD student. In Q2, the company raised a $40.0 million Series B at a valuation step-up, demonstrating a positive outlook during the pandemic. **Covariant** has partnerships with automation incumbents **ABB** and **Knapp**, demonstrating the priority for legacy robotics vendors to upgrade their product suites with AI & ML. We believe the approach can be utilized by other startups for other targeted use cases.

Automated mobile robot retrofits: Warehouse operators are looking for ways to bridge the manual labor-intensive gap between brick-and-mortar stores and last-mile delivery. **Amazon** Mobile Robotics, formerly Canvas Technology, has touted this as a competitive advantage based on its vision-based navigation. Technology developer **Humatics** has also focused on facilitating seamless navigation from indoor to outdoor environments through ultra-wide band beacons that allow for localization accuracy down to millimeters.

Robotics startups **MIR** and **Vecna** are also notable entrants into the space. In Q3, Portugal-based startup MOV.AI raised $4 million to embed autonomous decision-making into any industrial robot, addressing the retrofit problem in existing robotics deployments. In heavy industry, autonomous robots have been deployed in mining and construction for over 10 years, and improvements in neural networks are encouraging further automation for dynamic environments. Early-stage startup **SafeAI** has already partnered with **Ford** and **Caterpillar** to retrofit industrial vehicles with autonomous driving equipment COVID-19 is pushing logistics operators to consider automation, accelerating this theme.

Autonomous drones: Autonomous drones can unlock new use cases in field monitoring and extreme environments. We believe agricultural field monitoring is an area of high spending growth, with additional connectivity becoming available in rural areas. Autonomous drones can form a central part of its growth story. Startups **Exyn**, **Embention**, **vHive**, **PINC**, and **PreNav** have sensor-enabled drones that address a range of use cases from supply chain to subterranean mine inspection. **Exyn** has developed autonomous drones with multiple embedded sensors that can enter GPS-denied environments such as subterranean mines. **Skydio**'s $98 million Series C in Q3 demonstrate that drones are shifting from early adoption by consumers to widespread adoption by enterprises for diverse use cases.

Considerations

Products killed before commercialization: There is commonly a lag between product announcements and launch dates, which can last over a year. During this time, technical roadblocks can emerge, and cost overruns can reduce the economics of already low-margin hardware products. Robotics teams should have extensive experience with go-to-market strategies to overcome these commercialization hurdles.

Robots in human-present environments failing in the field: As new use cases emerge, robotics companies are struggling to turn proofs of concept into sticky revenue. SoftBank Robotics has faced high churn on its robot-as-a-service contracts, pressuring it to shorten the three-year length of its contracts. Henn na Hotel in Japan “fired” half of its robot fleet in early 2019 and replaced it with human workers. Humanoid robots may not be ready to interact with humans at present.

Regulatory risk: In autonomous vehicles and mobile robotics, uncertainty around regulatory frameworks may limit the growth of emerging technologies. The National Highway Traffic Safety Administration currently does not exempt self-driving cars from occupant safety standards, limiting the benefits of driverless technology. In robotics, regulation might be reactive to the introduction of robots in human-present environments, creating risk along with innovation in domain randomization.

Outlook

Multiple trends converging to stimulate deal activity in robotics: The autonomous robotics space has not been highly active in terms of venture or M&A activity, but we could see a revival in investor interest for a confluence of reasons. Factors include automation needs stemming from the pandemic, scientific improvements in AI & ML, underpenetration of robotics in non-automotive manufacturing globally, and a cyclical rebound from low manufacturing activity as a result of the US-China trade war. While an economic recession will chill deal flow in 2020, we believe a recovery will carry significant tailwinds for robotics and reveal the lagging positions of incumbents in industrial automation.

Emerging autonomous vehicle startups to surpass first wave: We believe startups that have formed recently are learning from the failures of leaders within the self-driving vehicle space, assuming lower levels of technical debt and less inflated expectations. Automakers and original equipment manufacturers are entrenched in the industry, and we believe they will remain committed to the technology. The maturation of AI & ML will necessitate more advanced approaches to this problem; novel datasets are already presenting fundamentally better results for complex environments. We anticipate another wave of unicorns in the space based on improvements in computing hardware and the introduction of multimodal datasets. Use cases may also be optimized as startups pursue low-hanging fruit such as trucking and middle-mile transportation. We believe the space is still promising.

Delivery drones to grow quickly: Urban air mobility has been fast-tracked due to the COVID-19 crisis. We believe this category can grow into a $1.0 billion market by 2025 at a high growth rate, setting it apart in this space. We do not believe that AI & ML drone startups have just begun to receive funding in line with this opportunity, in part due to regulatory concerns and the looming presence of **Amazon**. Delivery drones may present an emerging opportunity like that of micromobility companies in terms of growth, depending on the ability of the startup to win government approvals.

Supplemental materials

Select company analysis

Business overview

Founded in 2013, **Databricks** offers a data science platform including AIaaS functionality with a suite of data science tools for data engineering, data warehousing, and machine learning algorithms. The company grew out of the open source Apache Spark data science community, creating an extensible Unified Data Analytics Platform that can pull in data from across enterprise silos and prepare it for cluster-based computing. Once an effective open source product was in place, the company moved to a closed-source model and rapidly grew revenue beginning in 2016. **Databricks** now offers a product suite on top of Apache Spark, including notebooks for ML model collaboration, a data lake, and data security.

The company has achieved success in increasing the efficiency of data scientists and data engineers. The Forrester Total Economic Impact study it commissioned found that some data scientists and engineers report efficiency gains of over 20% by using its platform.[20] Those gains derive from less time spent building data pipelines and data preparation, tasks that we find can make up a majority of data scientists' time. The study suggests that **Databricks** is increasing its share of customers contributing over $1 million in ARR.

Management

Databrick's management has high technical expertise capable of generating product-led growth. The company was founded by seven early contributors to Apache Spark, a popular open source data science framework, and all of the cofounders remain at the company. CEO and co-founder Ali Ghodsi earned a PhD in distributed computing and was VP of engineering before being promoted to CEO at the start of the company's growth phase in 2016. CFO David Conte previously took **Splunk** public in the same role, suggesting the company's accounts will be ready for scrutiny by public markets. The board features co-founders and VC investors, including a16z co-founder Ben Horowitz, indicating that additional board appointments with public company experience will be required before IPO.

Competitive differentiation

Databricks stands out in its support for MLOps on top of its data warehouse, separating from legacy vendors **SAP**, **Microsoft**, and **AWS** while also being more AI-oriented than next-gen data warehouse vendor **Snowflake**. The company gains credit in the market for scaling ML model deployment and management across a variety of compute clusters, a critical problem

20: "The Total Economic ImpactTM of the Databricks Unified Data Analytics Platform: Cost Savings and Business Benefits Enabled by Databricks Platform," Forrester, April 2020.

Select company analysis

for data scientists new to ML. This advantage positions the company to capitalize on AI core adoption growth, which we forecast to be 16.7% out to 2023. For the time being, **Databricks** is capturing market share from incumbents such as **Alteryx**, **IBM**, and **SAS**. **Databricks** runs on **Microsoft** Azure and **AWS**, though those hyperscalers would likely prefer to natively support their ML workflows and may improve their offerings in the future.

Outlook

The company has achieved the scale and growth to realize high performance on public markets. Its most recent investment values the company at $28.0 billion, or 65.9x ARR of $425.0 million, on the back of 75% YoY ARR growth in Q3 2020. **Snowflake** debuted in 2020 with similar scale and higher growth and has traded up to 167.8x EV/trailing revenue. For this reason, the appeal of public markets will encourage an IPO in 2021 while the company is still in a high-growth phase. More broadly, **Databricks** will likely be encouraged to disclose the usage of its Notebook product for machine learning, which can give clear insight into the adoption of ML tools overall and level of spending within enterprises. An IPO would be important for the AI & ML horizontal platform market as a result.

Financing history

SERIES E

February 5, 2019

Total raised ($M):
$250.0

Pre-money valuation ($M):
$2,500.0

Investors:
Andreesen Horowitz (lead), Coatue Management, Founders Future, Geodesic Capital, Green Bay Ventures, Microsoft, NEA

SERIES D

August 22, 2017

Total raised ($M):
$140.0

Pre-money valuation ($M):
$800.0

Investors:
Andreesen Horowitz (lead), A.Capital Ventures, Australia Future Fund, Microsoft, NEA

SERIES C

December 15, 2016

Total raised ($M):
$60.0

Pre-money valuation ($M):
$453.0

Investors:
NEA (lead), Andreesen Horowitz, B7, In-Q-Tel, SineWave Ventures

SERIES B

June 30, 2014

Total raised ($M):
$33.0

Pre-money valuation ($M):
$214.0

Investors:
NEA (lead), Andreesen Horowitz, Australia Future Fund, Data Collective

SERIES A

September 24, 2013

Total raised ($M):
$14.0

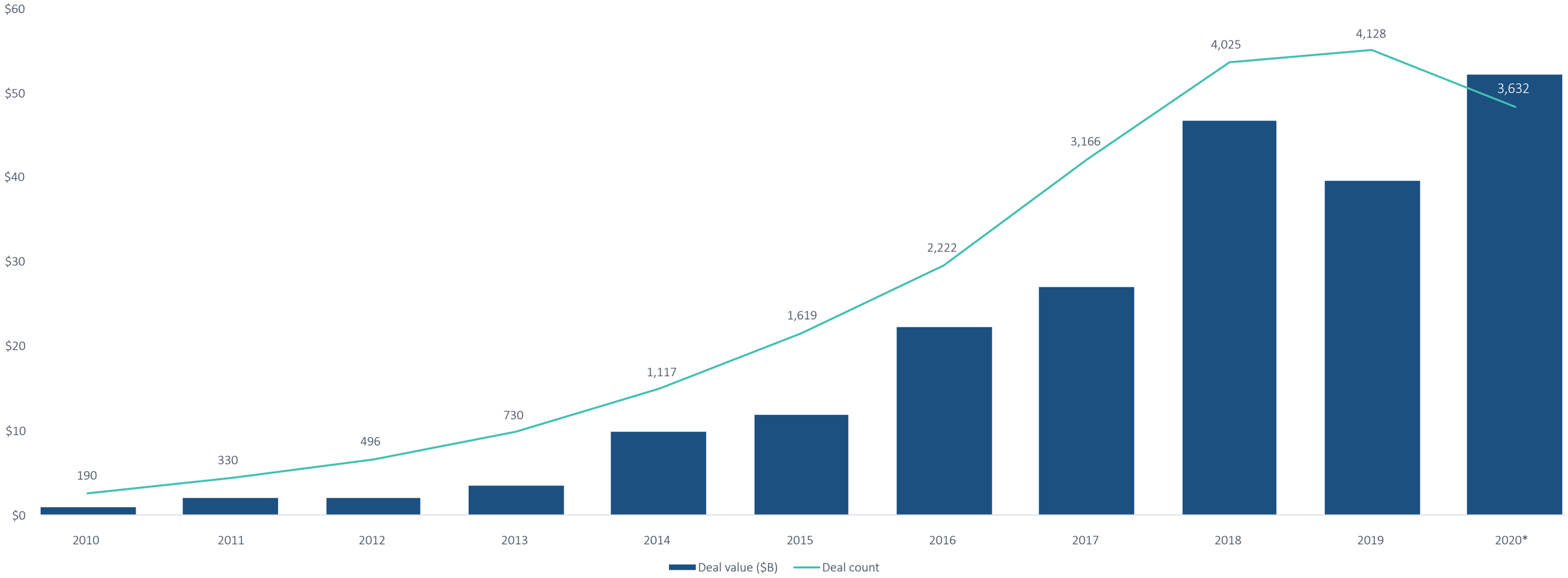
Pre-money valuation ($M):
$34.0

Investors:
Andreesen Horowitz (lead)

Additional VC data

Figure 33.

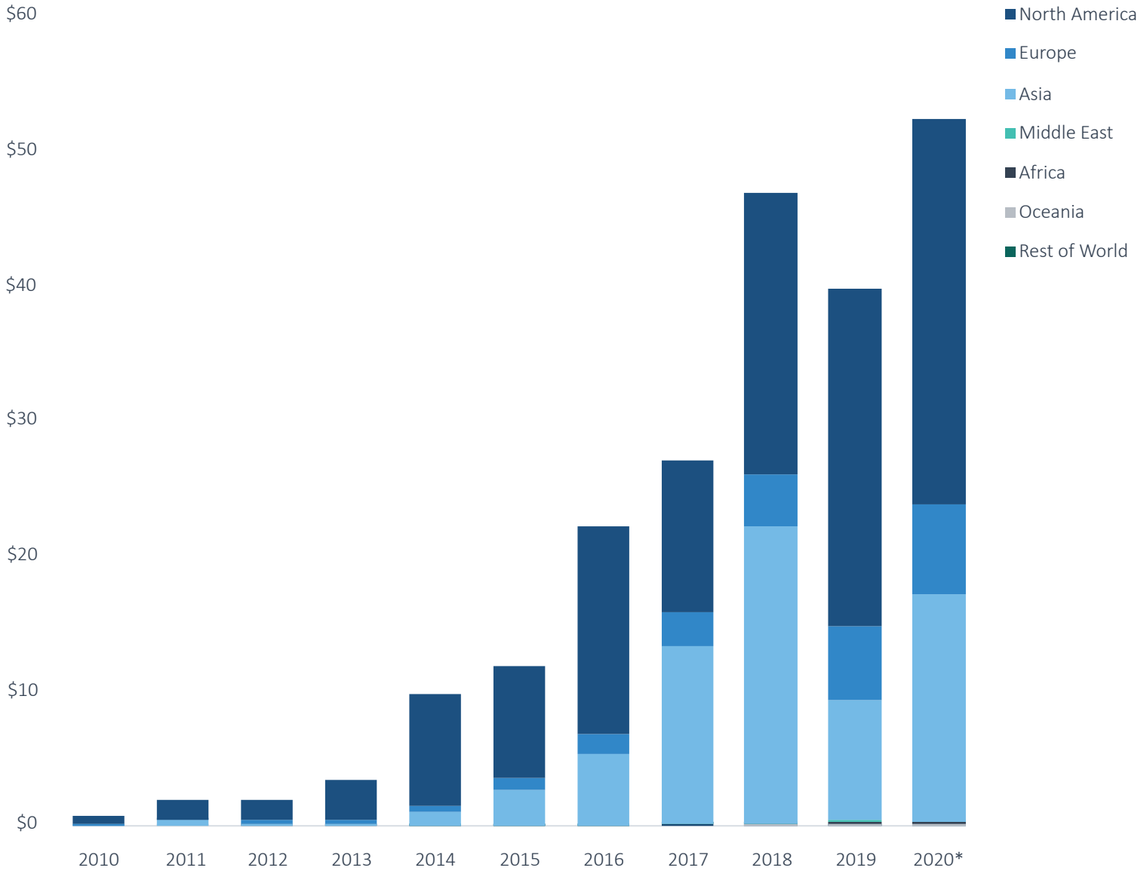
AI & ML VC deal activity



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 34.

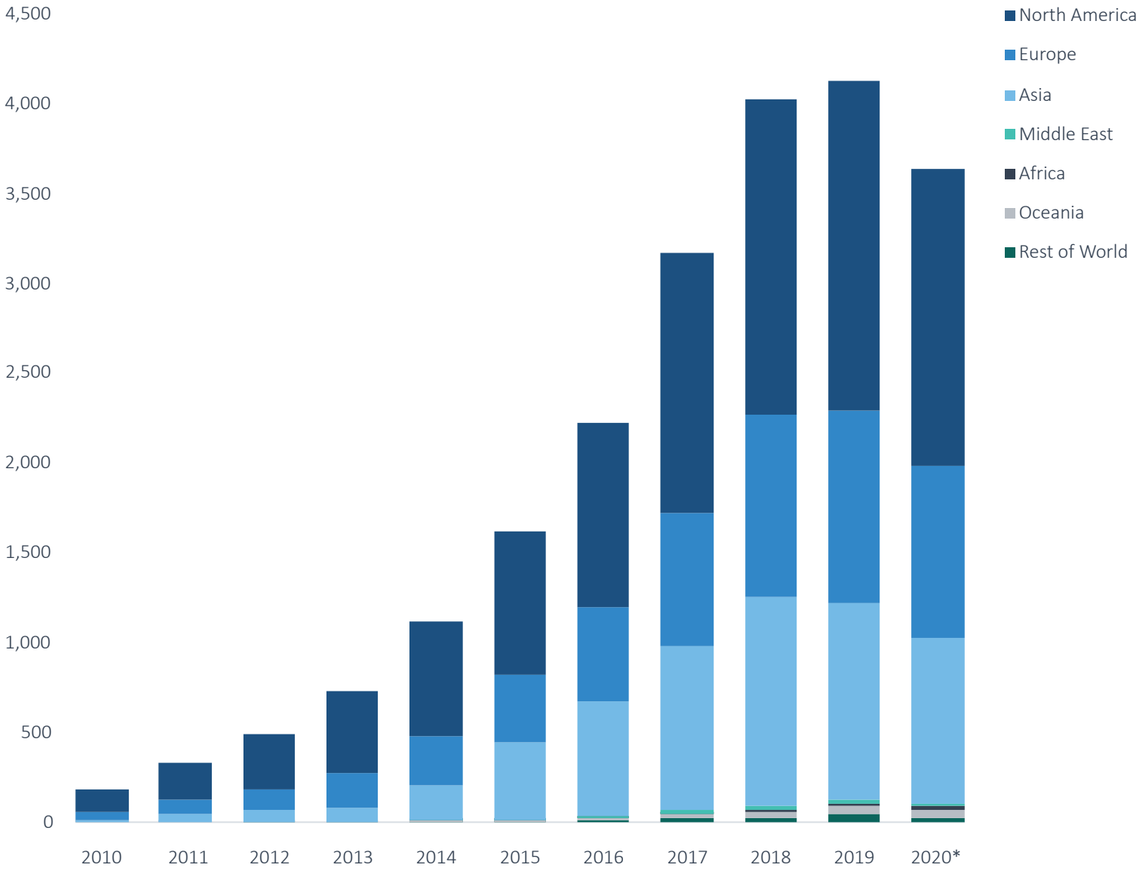
AI & ML VC deals ($B) by region



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 35.

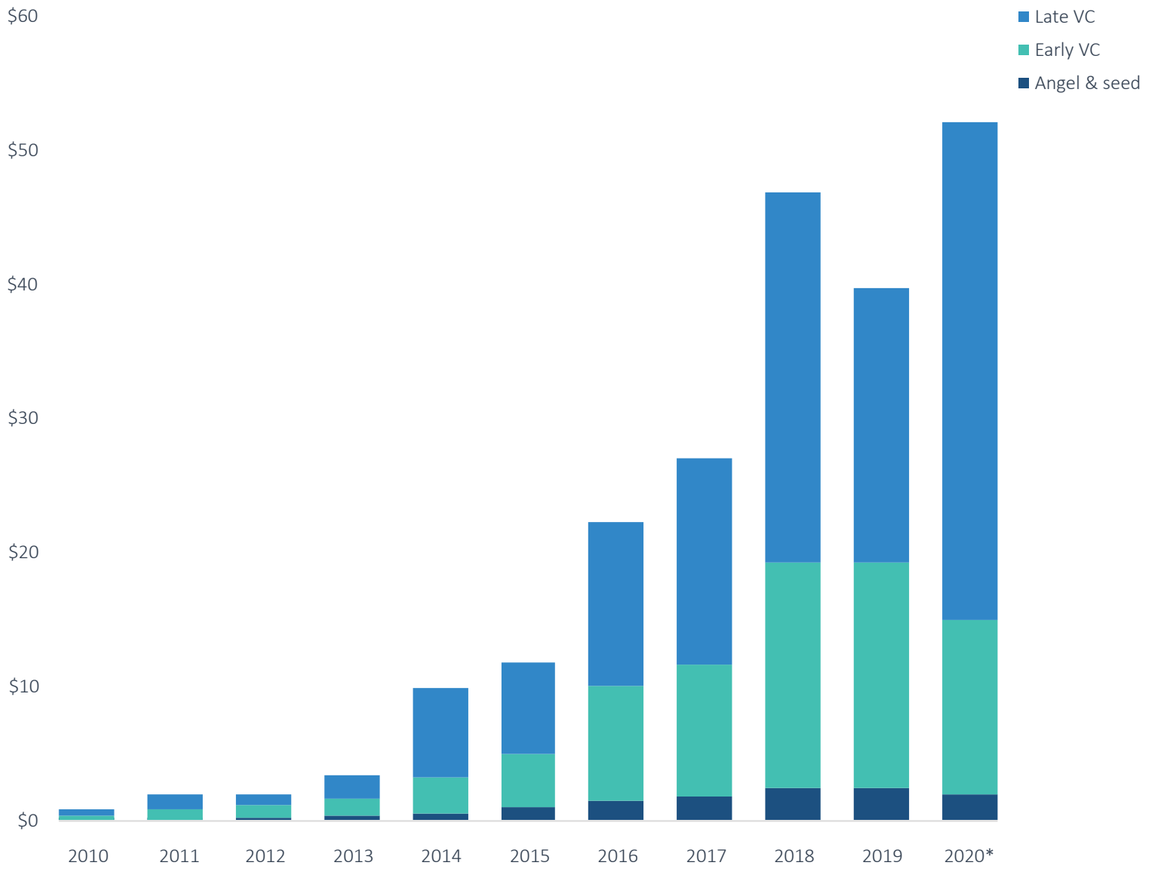
AI & ML VC deals (#) by region



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 36.

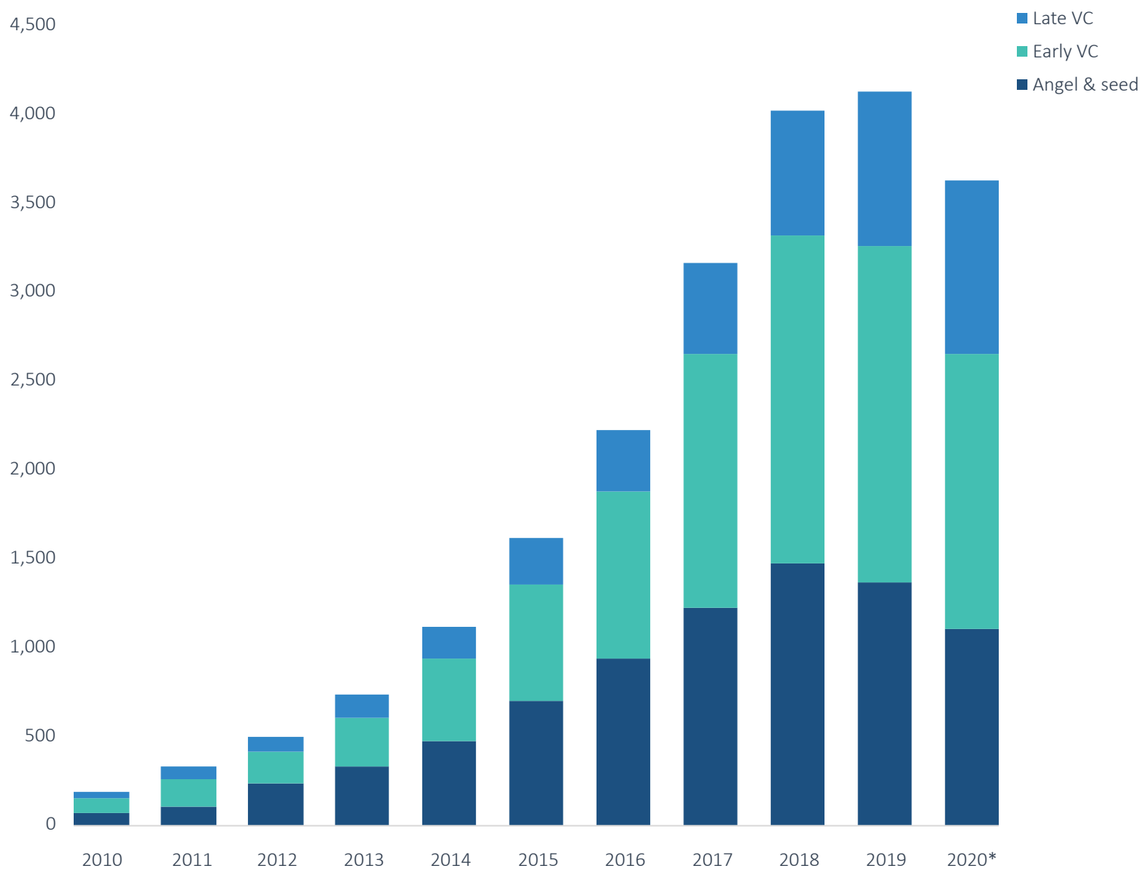
AI & ML VC deals ($B) by stage



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 37.

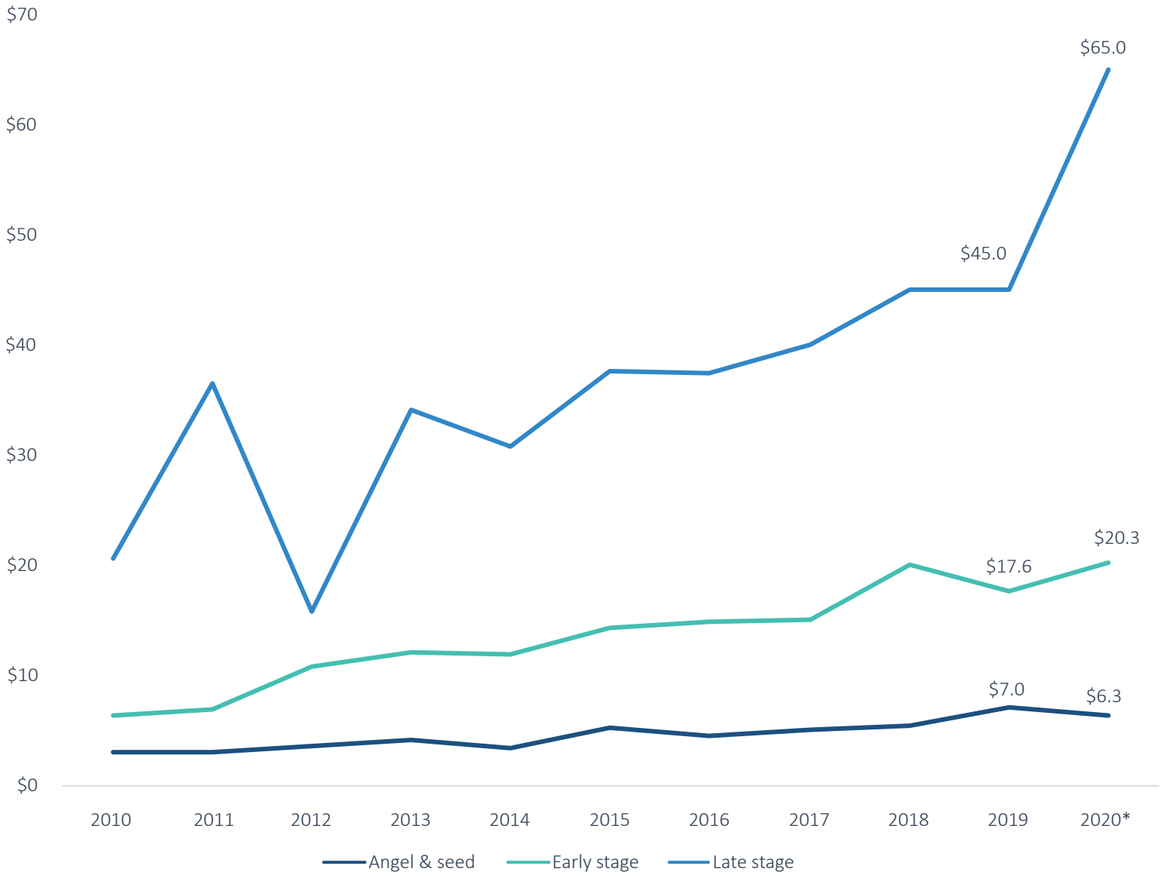
AI & ML VC deals (#) by stage



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 38.

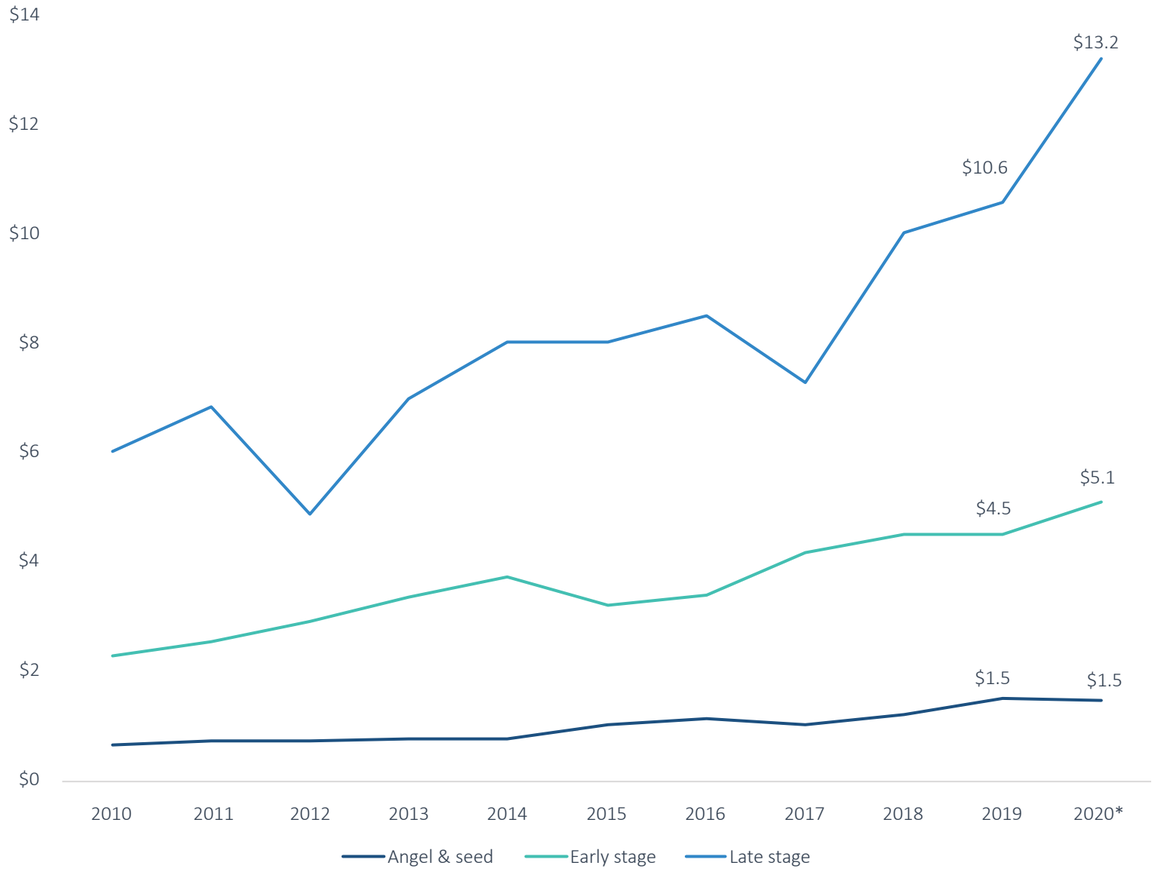
Median AI & ML VC pre-money valuation ($M) by stage



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 39.

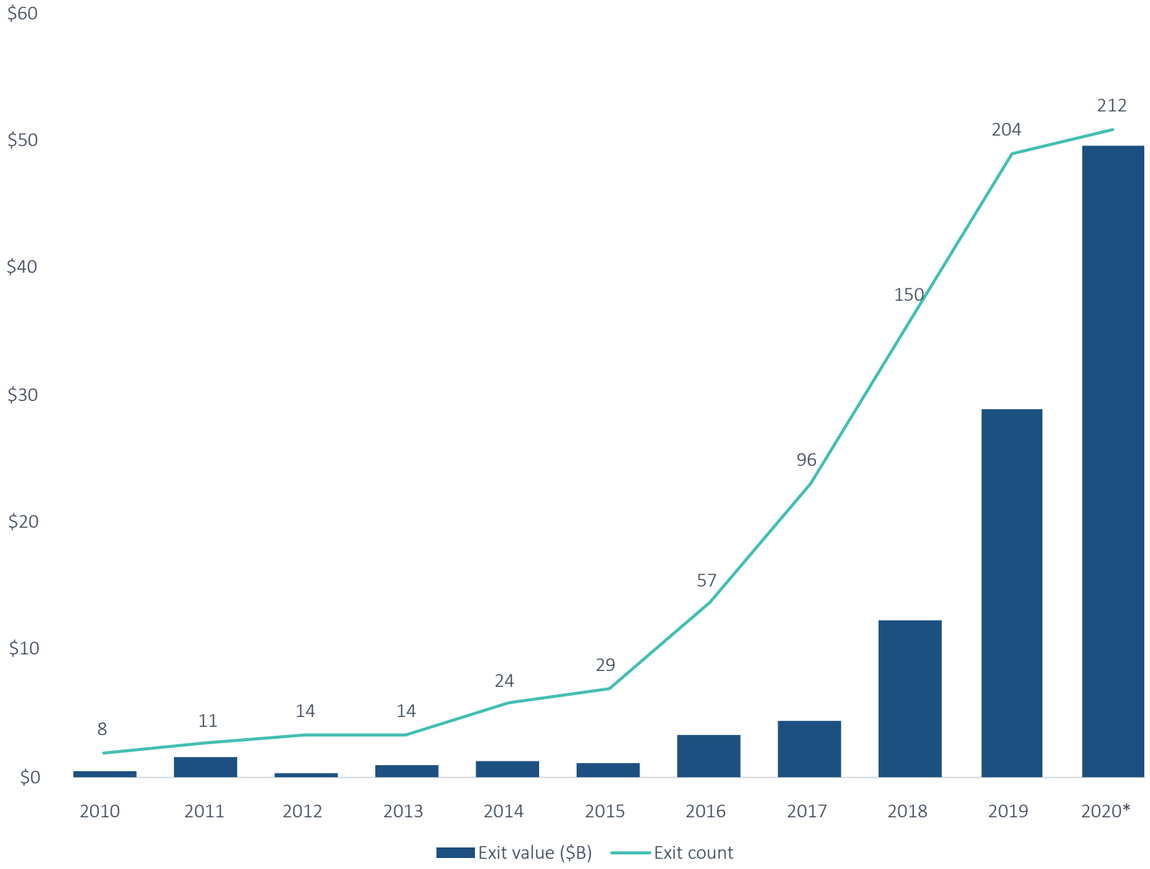
Median AI & ML deal size ($M) by stage



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 40.

AI & ML VC exit activity



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 41.

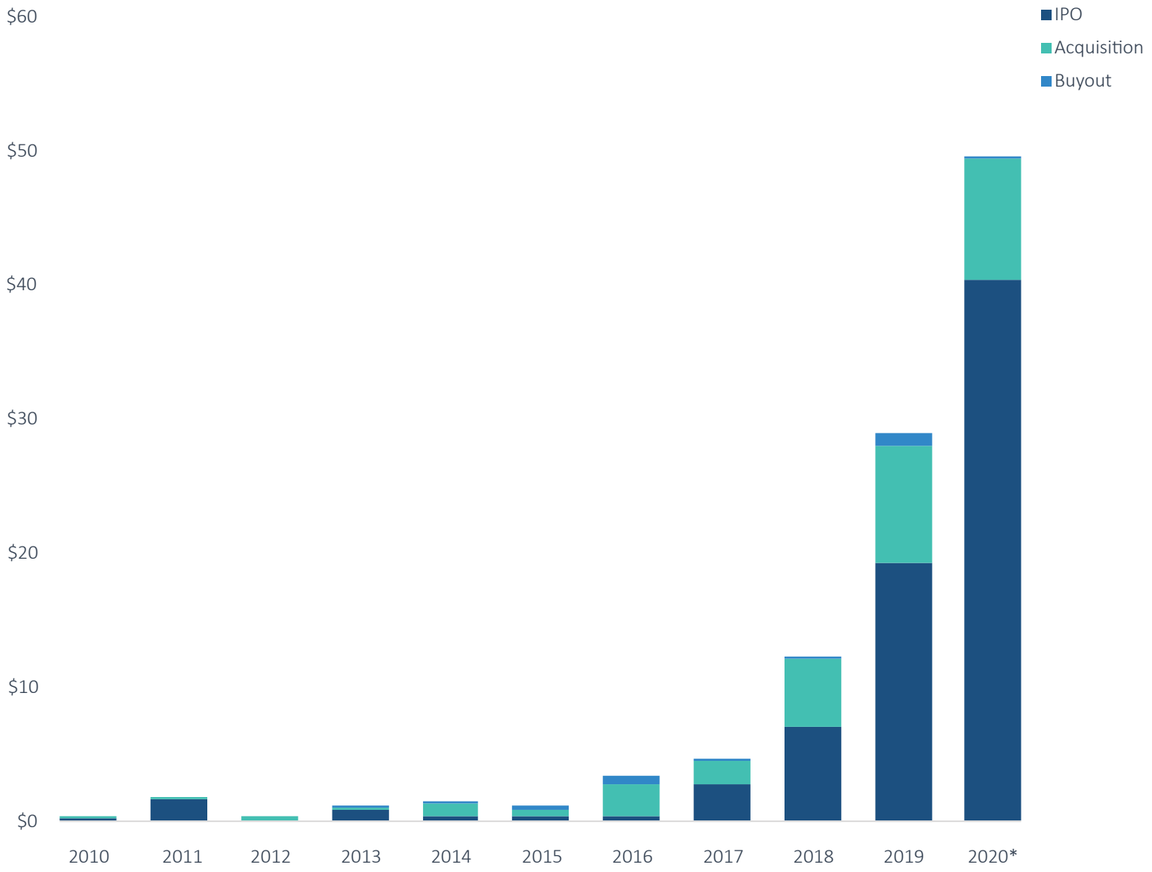
Top AI & ML VC exits by size

COMPANY	CLOSE DATE	EXIT SIZE ($M)	POST-MONEY VALUATION ($M)
Palantir Technologies	December 31, 2020	$21,000.0	$21,000.0
Datadog	September 19, 2019	$7,177.5	$7,825.5
CrowdStrike	June 12, 2019	$6,075.4	$6,687.4
Nio	September 12, 2018	$5,421.2	$6,422.8
Uber Advanced Technologies Group	December 7, 2020	$4,000.0	$4,000.0
C3.ai	December 9, 2020	$3,374.6	$4,025.6
Cambricon Technologies	July 20, 2020	$3,290.8	$3,657.4
nCino	July 14, 2020	$2,529.1	$2,778.9
Livongo	July 25, 2019	$2,190.5	$2,545.8
Habana Labs	December 16, 2019	$2,000.0	$2,000.0

Source: PitchBook

Figure 42.

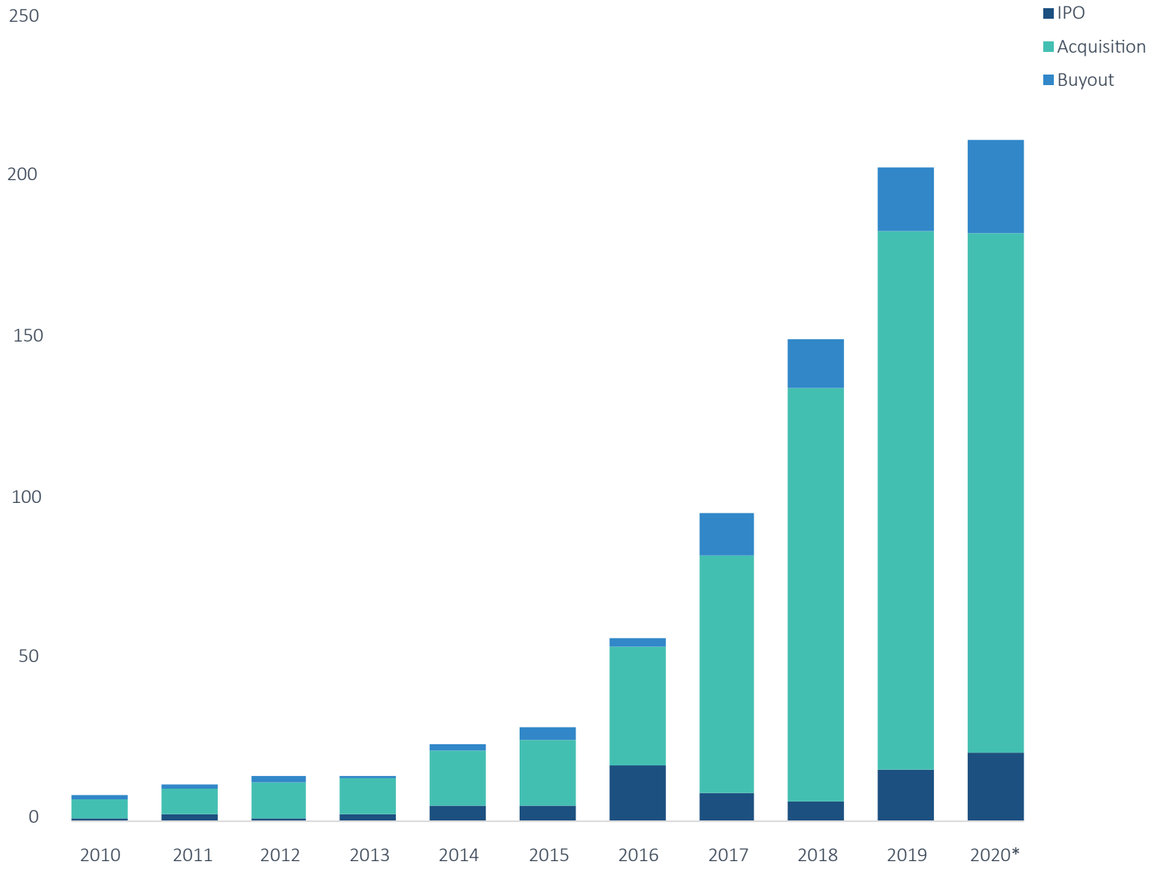
AI & ML VC exits ($B) by type



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 43.

AI & ML VC exits (#) by type



Source: PitchBook | Geography: Global
*As of December 31, 2020

Figure 44.

Top VC investors in AI & ML since 2017 by deal count

INVESTOR NAME	DEAL COUNT
Right Side Capital Management	123
Sequoia Capital China	118
500 Startups	105
Alumni Ventures Group	92
Intel Capital	88
Khosla Ventures	86
SOSV	83
Sequoia Capital	82
GGV Capital	82
Accel	81

Source: PitchBook | *As of December 31, 2020

Figure 45.

Top PE investors in AI & ML since 2017 by deal count

INVESTOR NAME	DEAL COUNT	PRIMARY INVESTOR TYPE
Vista Equity Partners	7	PE/buyout
The Carlyle Group	6	PE/buyout
Permira	5	PE/buyout
Insight Partners	5	Growth/expansion
Thoma Bravo	5	PE/buyout
TA Associates Management	5	PE/buyout
Warburg Pincus	5	PE/buyout
K1 Investment Management	5	PE/buyout

Source: PitchBook | *As of December 31, 2020

Buyers list

Figure 46.

Top strategic acquirers since 2017

INVESTOR NAME	DEAL COUNT	ENTITY TYPE
Apple	18	Corporation
Accenture	8	Corporation
Facebook	7	Corporation
Microsoft	7	Corporation
ServiceNow	7	Corporation
DataRobot	6	VC-backed company
Cisco Systems	6	Corporation
Twitter	4	Corporation
Intel	4	Corporation
Oracle	4	Corporation

Source: PitchBook | Geography: Global | *As of December 31, 2020

Glossary

Horizontal platforms

AI as a service (AIaaS): Cloud-hosted platforms that offer a mix of customizable AI & ML services including autoML, natural language processing and computer vision. These platforms support the development of custom AI & ML models and deploy their models into a runtime environment. Common features include AI marketplaces that offer data science tool APIs, deployment microservices, datasets and prebuilt algorithms.

AI & ML developer tools: Software libraries for data science and software developers involved with the AI & ML lifecycle. These can often be open source.

AutoML: an automated process of building, deploying and/or managing machine learning models. These processes are often low- or no-code and include automated model selection based on a single input of labeled data.

Cognitive computing: computing systems that are inspired by the function of the human brain. They are often modified neural networks using techniques such as reinforcement learning to replicate the learning process observed in humans. Artificial general intelligence is an end state of cognitive computing.

Data preparation platforms: Software that facilitates the labeling of data for usage in AI & ML models.

Quantum AI: Usage of quantum computers to carry out AI algorithms. Quantum computers' multifold improvement in computing speed can facilitate currently unfeasible AI & ML applications, though the technology is in development.

AI-enabled augmented reality: applications of computer vision to interpret visual data for real-time interpretation in augmented reality visualizers.

Computer vision as a service: Software platforms that enable computer vision inference on customer data via a cloud-hosted or on-premise application.

Facial recognition: The use of AI & ML to identify facial patterns and infer their resemblance to existing databases of facial structures.

Geospatial analysis: Modeling of satellite and map data to make predictions about the physical world including transportation, climate and biological systems.

Visual data labeling software: Automated data tagging software to produce data used in computer vision models.

Conversational AI: Application of natural language processing for interactive AI models including chatbots and virtual assistants.

Neural machine translation: Usage of deep neural networks to identify relationships among translated language datasets to automatically translate natural language with high accuracy and speed.

Natural language generation: AI models that synthesize textual content to create rich descriptions of insights found in data. NLG models can increasingly write complex content based on and related to simple prompts.

Natural language processing: AI models intended to read, decipher, understand, and make sense of language.

Natural language understanding: NLP models that can interpret phonetic spelling, dialects, cultural terms and domain-specific terminology.

AI for IT operations (AIOps): platforms that combining big data and machine learning functionality to analyze complex data resulting from digital transformations.

Contract lifecycle management automation: Software used to extract data and clarify the content of contracts for contract generation, negotiation and execution.

Robotic process automation (RPA): RPA is often purported to incorporate AI & ML optimizations but this capability is not common or formalized across the industry.

AI & ML semiconductors

Graphics processing units (GPUs): Chips that use parallel computing to break complex problems in separate tasks and execute them simultaneously.

Field-programmable gate arrays (FPGA): Chips based around a matrix of configurable logic blocks that can be programmed after manufacturing. Gate array refers to the manufacturing method in which layers of the chip are predefined.

AI-specific application-specific integrated circuits (ASIC): chips designed specifically for a client to provide a function required by the client's end product.

Neural network processors: Specialized circuit that implements all the necessary control and arithmetic logic necessary to execute machine learning algorithms, including parallel matrix multiplications and convolutions.

Post-training neural network binarization and quantization algorithms: Algorithm compression methods that can shrink the parameters of machine learning models to binary values or smaller data points that can be used with lower-precision calculations in edge devices.

Quantization aware training: ML model training using lower-precision calculations to generate model features, producing quantized outputs optimized for edge devices.

Edge AI: The ability to run AI algorithms locally on a distributed device. This includes the hardware components and software, so is inclusive of TinyML.

TinyML: Compressed ML algorithms that can operate within the hardware constraints of edge devices.

About PitchBook Emerging Tech Research

Independent, objective and timely market intel

As the private markets continue to grow in complexity and competition, it's essential for investors to understand the industries, sectors and companies driving the asset class.

Our Emerging Tech Research provides detailed analysis of nascent tech sectors so you can better navigate the changing markets you operate in—and pursue new opportunities with confidence.

Additional research

Agtech
Alex Frederick
alex.frederick@pitchbook.com

Artificial Intelligence & Machine Learning
Brendan Burke
brendan.burke@pitchbook.com

Cloudtech & DevOps
Paul Condra
paul.condra@pitchbook.com

Fintech
Robert Le
robert.le@pitchbook.com

Foodtech
Alex Frederick
alex.frederick@pitchbook.com

Health & Wellness Tech
Kaia Colban
kaia.colban@pitchbook.com

Information Security
Brendan Burke
brendan.burke@pitchbook.com

Insurtech
Robert Le
robert.le@pitchbook.com

Internet of Things (IoT)
Brendan Burke
brendan.burke@pitchbook.com

Mobility Tech
Asad Hussain
asad.hussain@pitchbook.com

Supply Chain Tech
Asad Hussain
asad.hussain@pitchbook.com