PitchBook

EMERGING TECH RESEARCH

Data Analytics Report

VC trends and innovation spotlights

Q2 2024

REPORT PREVIEW
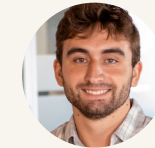The full report is available through the PitchBook Platform.

Contents

Institutional Research Group

Analysis

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

Ben Riccio Associate Analyst, Emerging Technology
ben.riccio@pitchbook.com
pbinstitutionalresearch@pitchbook.com

Data

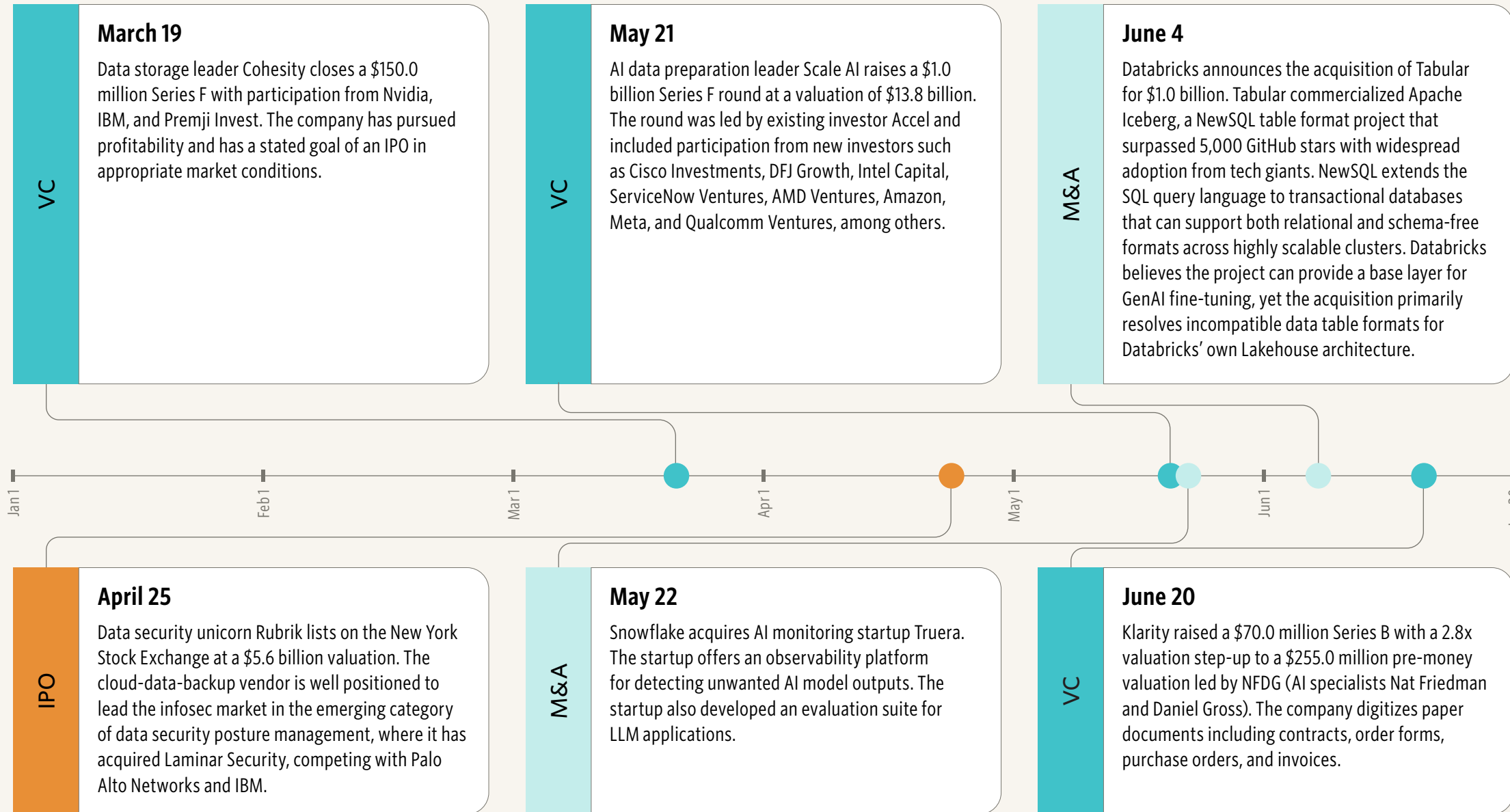**Matthew Nacionales** Senior Data Analyst

H1 2024 timeline

VC

March 19
Data storage leader Cohesity closes a $150.0 million Series F with participation from Nvidia, IBM, and Premji Invest. The company has pursued profitability and has a stated goal of an IPO in appropriate market conditions.

VC

May 21
AI data preparation leader Scale AI raises a $1.0 billion Series F round at a valuation of $13.8 billion. The round was led by existing investor Accel and included participation from new investors such as Cisco Investments, DFJ Growth, Intel Capital, ServiceNow Ventures, AMD Ventures, Amazon, Meta, and Qualcomm Ventures, among others.

M&A

June 4
Databricks announces the acquisition of Tabular for $1.0 billion. Tabular commercialized Apache Iceberg, a NewSQL table format project that surpassed 5,000 GitHub stars with widespread adoption from tech giants. NewSQL extends the SQL query language to transactional databases that can support both relational and schema-free formats across highly scalable clusters. Databricks believes the project can provide a base layer for GenAI fine-tuning, yet the acquisition primarily resolves incompatible data table formats for Databricks' own Lakehouse architecture.

Jan 1 · Feb 1 · Mar 1 · Apr 1 · May 1 · Jun 1 · Jun 30

IPO

April 25
Data security unicorn Rubrik lists on the New York Stock Exchange at a $5.6 billion valuation. The cloud-data-backup vendor is well positioned to lead the infosec market in the emerging category of data security posture management, where it has acquired Laminar Security, competing with Palo Alto Networks and IBM.

M&A

May 22
Snowflake acquires AI monitoring startup Truera. The startup offers an observability platform for detecting unwanted AI model outputs. The startup also developed an evaluation suite for LLM applications.

VC

June 20
Klarity raised a $70.0 million Series B with a 2.8x valuation step-up to a $255.0 million pre-money valuation led by NFDG (AI specialists Nat Friedman and Daniel Gross). The company digitizes paper documents including contracts, order forms, purchase orders, and invoices.

Q2 VC deal summary

375
total deals

$9.1B
total VC raised

-22.1%
deal value growth QoQ

2024 YTD VC summary

813
total deals

$20.7B
total VC raised

47.6%
deal value growth YoY

Data analytics landscape

1 Data infrastructure

2 Analytics platforms

3 Vertical applications

Data analytics VC ecosystem market map

This market map is an overview of venture-backed or growth-stage companies that have received venture capital or other notable private investments.

1 Data infrastructure

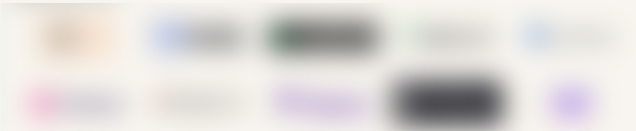**Datacenter**

Data pipelines

Data quality tools

DBMS
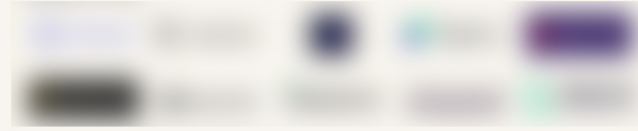
Data privacy & security software

Reverse ETL

2 Analytics platforms

AI core software

Analytics & BI

Unstructured data analytics

3 Vertical applications

Consumer

便利蜂 · JELLYSMACK · 花小猪 · STANDARD AI · mosyle
Pixellot · runway · VIP陪练 · 橙狮体育 · AiBEE

Fintech

JDT 京东科技 · Mission Lane · WeLab · 星化派 QUANT GROUP · LIQUIDITY
Socure · VNLIFE · GILION · Icertis · Perfios

Healthcare

innovaccer · insitro · immunai · DNAnexus · k health
神州医疗 Digital China Health · HUMA · Helix · Verana Health · Clarify
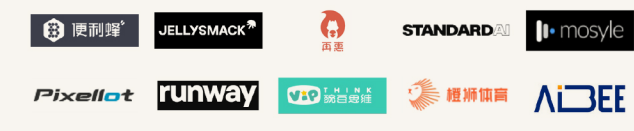
Industrial

IT

Mobility

VC activity

A rebound in VC funding applies to data analytics infrastructure and analytics platforms yet not vertical applications. AI funding pushed Big Data's vertical total to $20.7 billion in H1 2024. Stripping out the GenAI mega-deals in the AI core category, the infrastructure segment reached $3.7 billion after only $5.2 billion was raised in all of 2023. This pace suggests that data infrastructure VC activity is rebounding even as other verticals lag. Even so, the database management category only raised $324.5 million in VC funding in H1, on pace to fall 60.2% this year. The robust growth rates that fueled Databricks and Snowflake have tempered for earlier-stage challengers. Industrial data VC funding is also dwindling, on pace to fall 50.0% to $2.3 billion this year. VC activity is sustained by data center investments as more advanced computation comes to both AI and data science more broadly. This trend puts pressure on existing software vendors to extend their runways and compete with tech giants.

Data analytics exit activity is rebounding from 2023's lows yet remains in wait for major IPOs. We tracked $16.3 billion in disclosed exit value in H1 2024, up from $12.4 billion in calendar year 2023. Infrastructure and vertical applications drove deal value, with analytics continuing to produce smaller outcomes. We counted Tempus AI's IPO as a data analytics exit, given the company's database of clinical testing records. The other outlier was Databricks' $1.0 billion acquisition of Tabular, showing the outsize strategic value of infrastructure startups. Tabular commercialized Apache Iceberg, a NewSQL table format project that primarily resolves incompatible data table formats for Databricks' own Lakehouse architecture. Infrastructure is capturing exit count market share in a down market from analytics platforms and vertical applications, showing the resilient value of database innovation to incumbents.

Data analytics VC deal activity
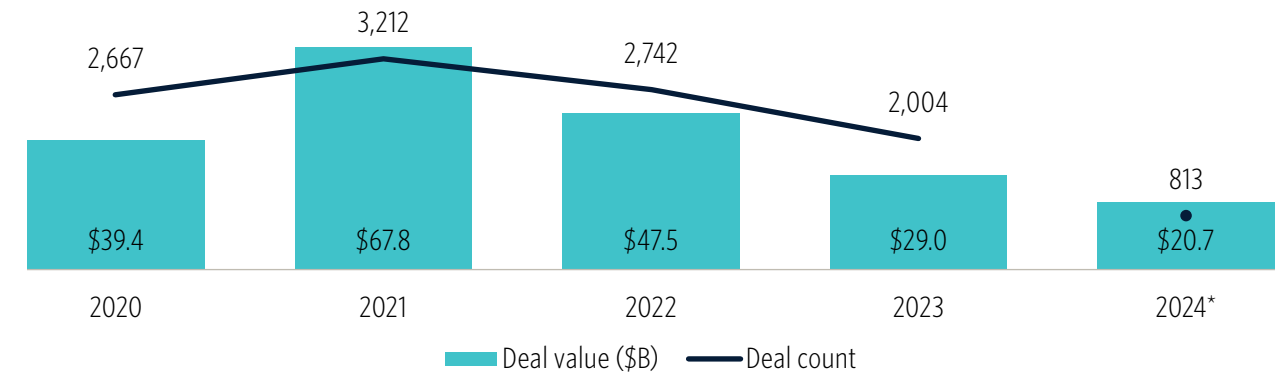
Source: PitchBook • Geography: Global • *As of June 30, 2024

Data analytics VC exit value ($B) by type
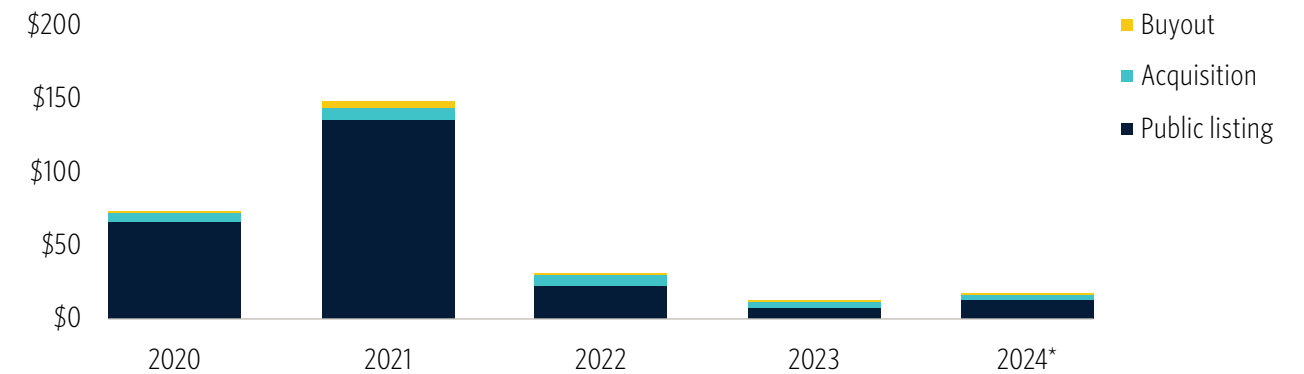
Source: PitchBook • Geography: Global • *As of June 30, 2024

Key data analytics early-stage VC deals by deal value in H1 2024*

Company	Close date	Category	Deal value ($M)	Post-money valuation ($M)	Lead investor(s)	Valuation step-up (post to pre)
Unstructured	March 14	Unstructured data analytics	$43.2	$223.2	Menlo Ventures	2.3x
Luma AI	January 9	Unstructured data analytics	$42.4	$220.0	Andreessen Horowitz	1.8x
Anomalo	January 24	Data quality tools	$33.0	$150.0	SignalFire	1.3x
Qdrant	January 23	Database management systems	$27.8	$142.8	Spark Capital	2.3x
Vectara	June 24	Unstructured data analytics	$25.0	$225.0	FPV Ventures, Race Capital	N/A
Tennr	March 26	Unstructured data analytics	$18.0	$78.0	Andreessen Horowitz	4.5x
Treefera	March 5	Analytics & business intelligence	$12.0	$71.5	AlbionVC	6.2x
Nile	January 30	Database management systems	$11.6	$52.8	Benchmark Capital Holdings	N/A
Upstash	February 8	Database management systems	$10.0	$40.0	Andreessen Horowitz	5.0x

Source: PitchBook • Geography: Global • *As of June 30, 2024

Key data analytics late-stage VC deals by deal value in H1 2024*

Company	Close date	Category	Deal value ($M)	Post-money valuation ($M)	Lead investor(s)	Valuation step-up (post to pre)
CoreWeave	May 2	Datacenter	$1,100.0	$19,000.0	Coatue Management, Magnetar Capital	7.1x
Scale AI	May 21	AI core	$1,000.0	$14,000.0	AlbionVC	N/A
AlphaSense	June 11	Unstructured data analytics	$650.0	$4,000.0	BDT & MSD Partners, Viking Global Investors	1.3x
Cdata	June 26	Unstructured data analytics	$350.0	N/A	Warburg Pincus	N/A
Cyera	April 9	IT	$300.0	$1,500.0	Coatue Management	2.4x
Abridge	February 22	Healthcare	$150.0	$850.0	Lightspeed Management Company, Redpoint Ventures	3.4x
Cohesity	February 7	Datacenter	$150.0	$3,200.0	NVIDIA, IBM, Premji Invest	1.2x
Atlan	March 19	Data quality tools	$77.5	$450.0	Insight Partners, Salesforce Ventures, Sequoia Capital	3.7x
Klarity	June 20	Unstructured data analytics	$70.0	$325.0	Daniel Gross, Nat Friedman	2.8x
PolyAI	April 22	Unstructured data analytics	$50.0	$500.0	Hedosophia, nVentures	N/A

Source: PitchBook • Geography: Global • *As of June 30, 2024

databricks

Overview

In Q2, Databricks continued to enhance its enterprise GenAI developer tools. At the Data + AI Summit in June, the company announced new features for its Mosaic AI platform. The platform, originating from the company's $1.3 billion acquisition of MosaicML in 2023, offers Databricks customers the tools to build custom LLM applications. The new features reflect the company's belief in compound systems to address quality and cost concerns in GenAI. Compound AI leverages several components, including multiple smaller models, retrievers, and vector databases to build applications rather than rely on one monolithic model.

Databricks announced five new Mosaic AI features at the summit, including Agent Framework for building custom RAG applications for enterprise data, and Agent Evaluation, an AI-assisted tool to evaluate the quality of developed GenAI applications. Other features enhance the training and governance abilities of Mosaic AI. Outside of custom GenAI developer tools, Databricks is also implementing AI solutions to its business intelligence offerings. The new business intelligence dashboard, AI/BI, provides users with a dashboard to build low-code data visualizations and an LLM to analyze large datasets through natural language queries.

Key company information

Founded
2013

Last financing valuation
$43.2B

Lead investors
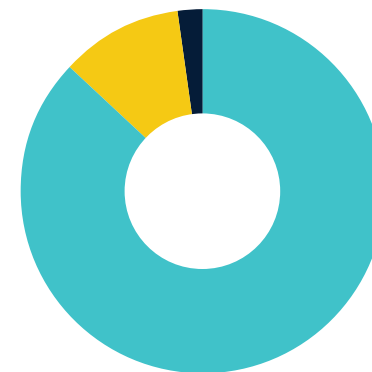NVIDIA, T. Rowe Price, Counterpoint Global, Franklin Templeton Investments, Andreessen Horowitz, NEA

Employees
9,189

Last financing
Raised $684.6M in a Series I

Total raised
$4.2B

December 2023
Leader in Gartner Magic Quadrant for Cloud Database Management Systems

Exit Predictor

■ **IPO:** 87% probability   ■ **M&A:** 11% probability   ■ **No exit:** 2% probability

Success: 98% probability

Note: Probability data is based on PitchBook VC Exit Predictor methodology.

Leadership update

Databricks' estimated $1.0 billion acquisition of Tabular in June brought in most of its 40 members, including CEO Ryan Blue.[14] Tabular commercialized Apache Iceberg, a competing data table format to Databricks' own Delta Lake. Blue and the former Tabular team will likely work on Delta UniForm, a format compatibility feature of Databricks that serves to increase data interoperability. The acquisition and integration of the Tabular team gives Databricks capabilities and expertise in two of the most popular open-source data table formats and can serve to support AI data preparation.

Competitive landscape

Databricks' new product announcements follow similar GenAI tools and data interoperability enhancements by competitor Snowflake. At a similar summit one week prior, Snowflake announced new features for its own GenAI service, Cortex. Cortex AI leverages leading LLMs and Snowflake's own open-source model, Arctic. Additionally, Snowflake announced its new open-source Polaris Catalog, which supports data stored in Apache Iceberg format. We believe these moves are reactive, and customers are more likely to use Databricks for AI. Databricks notes 210%

growth in the number of companies registering at least one AI model in 2024, with 1,018% growth in the volume of distinct AI models in the company's platform overall.[15] We have long believed that Databricks has a more comprehensive AI strategy than Snowflake does, and Databricks is now likely to surpass Snowflake in revenue.

Outlook

Databricks has achieved the necessary scale and growth to realize high performance in public markets. The company neared $2.4 billion in Q2 2024 with 60% revenue growth, putting it on pass to grow past Snowflake in revenue.[16] The company will need to defend itself against ML-optimized challengers that can carry out more efficient streaming data analysis and in-memory analytics within its database. Furthermore, it will benefit from custom AI models built on its database becoming essential to high-value industries, including IT and industrial. The company remains innovative and active in future-proofing its business, leading us to believe that, in the long run, Databricks will be valued more highly than relational database incumbents based on the growth opportunities of data lake architecture for streaming data.

14: "Databricks to Buy Data Management Firm Tabular for Over $1 Bln," Reuters, Jaspreet Singh, June 4, 2024.

15: Databricks, "State of Data+AI: Data Intelligence and the Race to Customize LLMs," May 29, 2024.
16: "Databricks Tells Investors Annualized Revenue Will Reach $2.4 Billion at Midway Point of Year," CNBC, Jordan Novet, June 12, 2024.

Financing history

Series A	Series B	Series C	Series D	Series E	Series F
September 24, 2013	**June 30, 2014**	**December 15, 2016**	**September 25, 2018**	**January 11, 2019**	**October 22, 2019**
Total raised	**Total raised**	**Total raised**	**Total raised**	**Total raised**	**Total raised**
$14.0M	$33.4M	$60.0M	$140.0M	$250.0M	$400.0M
Pre-money valuation	**Pre-money valuation**	**Pre-money valuation**	**Pre-money valuation**	**Pre-money valuation**	**Pre-money valuation**
$30.0M	$220.0M	$460.0M	$800.0M	$2.5B	$5.8B
Lead investor	**Lead investor**	**Lead investor**	**Lead investor**	**Lead investor**	**Lead investor**
Andreessen Horowitz	NEA	NEA	Andreessen Horowitz	Andreessen Horowitz	Andreessen Horowitz

Series G	Series H	Series I
February 9, 2021	**August 31, 2021**	**November 13, 2023**
Total raised	**Total raised**	**Total raised**
$1.0B	$1.6B	$684.6M
Pre-money valuation	**Pre-money valuation**	**Pre-money valuation**
$27.0B	$36.4B	$42.5B
Lead investor	**Lead investor**	**Lead investors**
Franklin Templeton Investments	Counterpoint Global	NVIDIA, T. Rowe Price

About PitchBook Industry and Technology 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 Industry and Technology 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.

PitchBook Data, Inc.

Nizar Tarhuni Executive Vice President of Research and Market Intelligence

Paul Condra Head of Emerging Technology Research

Additional research

Eric Bellomo
eric.bellomo@pitchbook.com
Gaming
E-Commerce

Brendan Burke
brendan.burke@pitchbook.com
Data Analytics
Information Security
Artificial Intelligence & Machine Learning

Aaron DeGagne
aaron.degagne@pitchbook.com
Medtech
Digital Health

Alex Frederick
alex.frederick@pitchbook.com
Agtech
Foodtech

Jonathan Geurkink
jonathan.geurkink@pitchbook.com
Supply Chain Tech
Mobility Tech

Kazi Helal
kazi.helal@pitchbook.com
Biopharma
Pharmatech

Derek Hernandez
derek.hernandez@pitchbook.com
Enterprise SaaS
Infrastructure SaaS

Ali Javaheri
ali.javaheri@pitchbook.com
Emerging Spaces

Robert Le
robert.le@pitchbook.com
Insurtech
Crypto

John MacDonagh
john.macdonagh@pitchbook.com
Carbon & Emissions Tech
Clean Energy Tech

Rebecca Springer
rebecca.springer@pitchbook.com
Healthcare Services
Healthcare IT
Pharma Services

Rudy Yang
rudy.yang@pitchbook.com
Enterprise Fintech
Retail Fintech