

FORTUNE

Data-labeling company Scale AI valued at \$7.3 billion with new funding

BY [JEREMY KAHN](#)

April 14, 2021 3:00 AM GMT+10

Scale AI, a San Francisco-based startup that specializes in helping companies label and curate data for artificial intelligence applications, has been valued at \$7.3 billion in its latest funding round, announced Tuesday.

[The company](#) said it has raised \$325 million in the latest investment round. The new financing brings the total amount of venture capital it has raised since its founding in 2016 to more than \$600 million.

The latest valuation figure more than doubles the one the company received during [its last \\$155 million funding round in December 2020](#). Alexandr Wang, the company's chief executive officer, told *Fortune* that "there are no current timelines or current plans" to take the company public, but that "we are always paying attention to the market."

The latest investment round was led by Dragoneer and Greenoaks Capital, both San Francisco-based investment firms, and Tiger Global, the prominent technology-focused investment company based in New York. Wellington Management and Durable Capital also joined the funding round. Existing investors Coatue Management, Index Ventures, Founder's Fund, and Y Combinator also participated, the company said.

Jeff Wilke, a former top [Amazon](#) executive, will join the company as a special advisor to the CEO, Scale said.

The startup is among a handful of companies that have sprung up to help businesses deal with the data preparation needed to train A.I. systems, a process that often requires tens of thousands of examples to be annotated with labels. Others that also offer data labeling services include Samasource, Labelbox, Hive, Cloudfactory, and DataLoop.

Data labeling and curation has become a popular "picks and shovels" play for investors hoping to cash in on the current A.I. boom. Scale has grown to become perhaps the largest in the sector, having branched out from an initial focus on image and video data, primarily for companies working on self-driving cars. The company now offers support for a wide variety of both vision and natural language data for businesses in industries ranging from finance to logistics to government.

Wang tells *Fortune* that the company is currently on track to earn \$100 million in revenue in a 12 month period and that sales have doubled in the past year. Notable customers include PayPal, Pinterest, and the U.S. Air Force. The company also works with major automakers such as Toyota, General Motors.

The company has also grown to 300 employees, up from about 100 a year ago.

Machine learning is increasingly replacing traditional software programming as the way companies automate tasks. "Data is the new code," Wang says.

"Data is so fundamental and critical to the building of these systems, the training, the testing," Wang continues. "Companies need to be able to utilize and manipulate data just as they have utilized and manipulated code in the past."

Scale has moved from just labeling data to a suite of software-based services. Its offerings help companies to gather, annotate, curate and clean-up their data as well as to build and monitor machine learning models trained on that data.

One software package, which it calls Nucleus, allows customers to quickly search through their data to find mislabeled examples that may be degrading an A.I. algorithm's performance. It can also easily apply new labels to existing and new data to help an A.I. system improve by giving it more training in areas where it is weaker.

"When someone talks about 90% or 95% accuracy, that failure rate is not uniformly distributed," says Russell Kaplan, a former senior machine learning engineer for Tesla who now heads the Nucleus team at Scale. All A.I. systems have statistical biases—a tendency to make mistakes. These errors often involve "edge cases," rarer events that are not as well-represented in their training data. Kaplan compares Nucleus to software debugging tools, but for data instead of software code. The system was built from technology he developed at a startup called Helia that Scale acquired in November 2020.