Dataloop raises $16 million for data annotation tools

VB venturebeat.com/2020/10/14/dataloop-raises-16-million-for-data-annotation-tools

Kyle Wiggers

October 14, 2020


AI data management and annotation startup Dataloop today announced it has raised $16 million in funding through an $11 million series A round and a previously undisclosed $5 million seed round. A spokesperson says the funds will enable Dataloop to increase its recruitment efforts and grow its presence in the U.S. and Europe.

Training AI and machine learning algorithms requires plenty of annotated data. But data rarely comes with annotations. The bulk of the work often falls to human labelers, whose efforts tend to be expensive, imperfect, and slow. Dataloop claims to solve the annotation challenge with a platform for automating data prep and data operations. It specializes in high volumes, high variance, and complex data, helping a range of companies create AI development and production pipelines.

According to cofounder and CEO Eran Schlomo, the pandemic has accelerated digitalization within enterprises and increased demand for automation and AI applications. Dataloop has seen a surge of interest in its platform from health and medical sectors. The autonomous vehicle space has slowed with the onset of the pandemic, Schlomo says, but there are "promising signs" that development has resumed.

"Many data labeling companies were forced to move to a work-from-home model with their data labeling workers due to the pandemic, which presented innumerable challenges in terms of collaboration, communication, data guideline training, quality, and supervision," a spokesperson told VentureBeat via email. "The Dataloop platform enables data labeling processes and data management for distributed teams with collaboration and communication tools built-in. Many data labeling companies found and continue to see the value in our model."
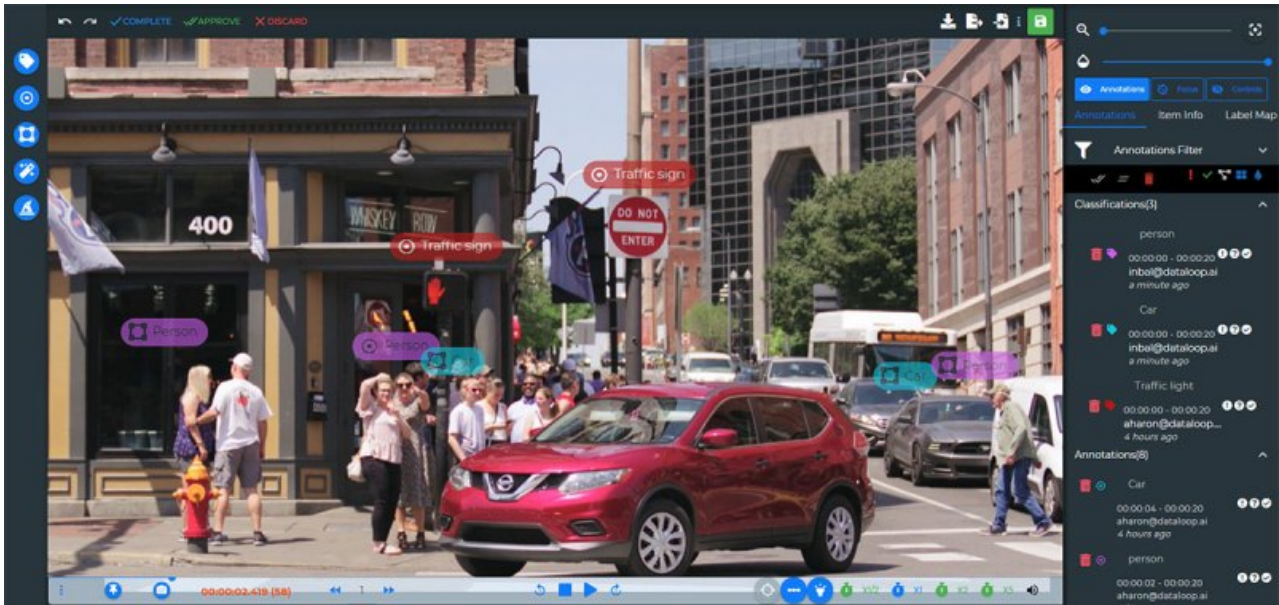
Above: Video scene classification in Dataloop.
Image Credit: Dataloop

Dataloop's annotation platform lets customers create an unlimited number of datasets with guidelines and ontologies. Algorithms, guidance tools, and models ostensibly reduce labeling time, while feedback and quality assurance components provide visibility throughout the annotation process.

Dataloop offers human-operated tools and plugins from item classification to pixel-level segmentation for images, videos, and more. The platform's trackers and models — which handle things like video interpolation, scene classification, smart object tracking, object identification, and occlusion — run automatic annotations on items before they reach human annotators, who only have to fix or validate labeled data instead of working from scratch.

From a backend dashboard, Dataloop customers can see per-user metrics regarding processes, performance, annotation, and usage. They're also afforded access to quality and validation tools embedded in the annotation process. Managers can assign feedback and tasks using Dataloop's "data bug" dialog module, annotations issue tracking, and item review. And they can support internal annotation teams or order labeling on-demand, thanks to Dataloop's managed workforce network of labeling providers.

Above: Dataloop's automatic annotation tools.
Image Credit: Dataloop

Dataloop's data operations toolkit is as at least as extensive as its annotation tools. Customers can use it to search, edit, and query data by item info, metadata, annotation status, and more. They can also sort and filter datasets before distributing items for annotation and quality assurance, combining input from multiple sensors to receive data outputs like images, videos, and lidar point clouds. And they can opt to use Dataloop's Python software development kit to build custom data pipelines for model training, active learning, and model integration.

Amiti Ventures led Tel Aviv-based Dataloop's series A round, with participation from F2 Venture Capital, OurCrowd, NextLeap Ventures, and SeedIL Ventures.

Dataloop is in a category adjacent to companies like Scale AI, which has raised over $100 million for its suite of data labeling services, and CloudFactory, which says it offers labelers growth opportunities and "metric-driven" bonuses. That's not to mention Hive, which netted $10.6 million in November 2019; Alegion, which nabbed $12 million in August 2019; Appen; SuperAnnotate; and Cognizant.

But Shlomo says Dataloop has 5,000 users on its platform, including Fortune 100 companies. One unnamed pharmaceutical client is reportedly using Dataloop to detect and track cancer cells in microscopic images and video frames. A client in the defense industry is leveraging the platform to build an autonomous on- and off-road vehicle that will be used for guarding and patrol in areas where it is unsafe for soldiers to be posted.

"The closing of this round will allow Dataloop to bring its next-generation cognitive cloud capabilities to the forefront, a cloud in which mixed human-machine intelligences are weaved seamlessly together to create the next generation of smart applications,"

Shlomo said. "With this opportunity, we hope to continually improve data labeling, making it increasingly rewarding both financially and professionally for those in the profession."