
Case Study

Zebra Medical Vision accelerates proactive discovery of medical conditions using Coiled



Published July 2021

Coiled customer Zebra Medical Vision was named one of the most innovative companies of 2020 by *Fast Company* and with good reason. Its AI-powered software can scan medical images and automatically detect chronic undiagnosed medical conditions that affect the general public.

Founded in 2014, Zebra Medical Vision performs opportunistic screening of medical images to detect indicators of musculoskeletal, cardiovascular, and pulmonary problems. Its AI discovers underlying chronic conditions, such as clotted coronary arteries, using imagery originally captured for other purposes (like a CT scan that explored breathing issues).

Zebra Medical Vision initially used data analysts to identify patterns in the medical images. However, as their business grew they needed a more scalable approach and moved to machine learning. But because they use an incredible amount of data to train and validate their machine learning models, it became apparent in 2020 that further changes to how they operated were necessary. Before working with Coiled — and utilizing Dask — Zebra Medical Vision's data analysts mined and queried databases that physicians and other annotators constantly updated. This process was slow, cumbersome, and extremely costly.

A Pressing Need

Eyal Toledano, CTO and co-founder of Zebra Medical Vision, envisioned utilizing computer vision and deep learning models along with a new approach he calls data-as-code. This process creates datasets as artifacts with versioning, configuration management, and data quality checks. To enable data-as-code, he searched for an environment that could query a large amount of data and process it quickly.



Toledano also wanted an open-source solution that fit into his existing architecture and supported both on-prem and a cloud environment. His initial search proved disappointing, as the options were either cloud-only or on-prem-only. And while he knew Python could do some of the work, his team hit limits in their experimentations with Pandas.

Toledano liked the APIs and how Python was synergistic with his existing code base, but he thought the library limitations were hard to overcome. Then he discovered Coiled, the Dask-powered cloud environment which met his architectural needs. It appeared to be the ideal environment to grow with the business as customer demand increased.

“Using Coiled, we were able to automate large scale data curation and trigger training pipelines automatically. It’s exactly like having CI/CD for data. We call it data-as-code.”

Eyal Toledano, CTO and co-founder, Zebra Medical Vision

Toledano knew transitioning away from database systems to the Python-based Dask would present his team with a few obstacles. Toledano thought being able to leverage the large, active Dask user community would help ease the transition. Having access to Dask experts is a key reason Zebra Medical Vision selected Coiled as our CEO Matthew Rocklin is the initial author of Dask and our team includes many other core Dask maintainers.

A Minimal Learning Curve

In choosing Coiled Cloud, Toledano found an environment that was simple and straightforward. He achieved peace of mind knowing the Coiled team would be there whenever he needed them. While there were some early challenges around network configuration, Toledano was surprised by how easy it was for Zebra Medical Vision to bridge its knowledge gaps.

Using tutorials provided by Coiled — and leaning on the expertise of our Dask engineers, Toledano and his team quickly deployed on Dask without a lot of DevOps work. Reproducing and linking datasets is crucial to the clinical



validation of its models. Now, thanks to Coiled, Zebra Medical Vision creates these data assets as part of a reproducible and version-controlled CI/CD process for data, reducing both time and cost.

“We were able to segue from never-ending SQL queries on 2TB of metadata and 2,500 tables that cost hundreds of dollars per query to 15 – 30 minute exercises that were a small fraction of the previous cost.”

Eyal Toledano, CTO and co-founder, Zebra Medical Vision

An Incredible Transformation

According to Toledano, Coiled Cloud and Dask allowed him to achieve his end goal: to be focused on the ML models, not the tech that curates the datasets. He believes that his previous approaches negatively impacted Zebra Medical Vision’s resources, finances, and overall time-to-market.

“Coiled is amazing technology. We were able to get an initial reduction across all of our pipelines – data curation to automated experiments. We saw our processing time drop from 66 hours to 35 minutes and, with additional tuning, down to 15 minutes. We gained 64 hours and 45 minutes using Coiled and Dask.”

Eyal Toledano, CTO and co-founder, Zebra Medical Vision

CHALLENGE

Could only process limited sets of data, which took days and wasn’t reproducible.

SOLUTION

Migrated to Coiled Cloud and Dask for a supercharged Python experience.

RESULTS

Got answers back in 15 minutes vs. 66 hours.



About

Coiled scales Python to the cloud for data professionals. Based on Dask, the leading Python-native solution for distributed computing, Coiled has hosted more than 100M tasks for data professionals, scientists, and researchers around the globe including Capital One, Anthem Health, and the Air Force to solve challenges in business, research, and science.

Coiled is a remote-first company with the best and brightest working from around the globe.

Founded by the initial author of Dask, Coiled is on a mission to create a platform that gives Data Scientists the power of the cloud and machine learning, freeing them from today's limitations so they can solve important problems.

[Learn More](#)

[Try It Now](#)

Follow Coiled

