## 1 Scale AI, Inc. (Scale) Introduction

Scale is the market leader in ML data curation and human-machine teaming. Scale's metaprogramming combines an AI engine with human-in-the-loop verification to label computer vision and surveillance data to build faster, cheaper, and higher-quality annotated datasets and achieve our mission of accelerating the development of AI applications for the U.S. Government.

Deploying AI products is a process, not an event. We've learned from real-world teams doing AI/ML at scale that they need a better way to manage their data to rapidly iterate on datasets. Scale provides a simple API to collect, annotate, and manage datasets — all in one place.

Scale follows a data-centric ML lifecycle shown below in Figure 1, which allows our partners to achieve faster time to deployment. We can accelerate AI readiness by keeping your engineers focused on more impactful and differentiated projects than on data annotation. Scale is committed to applying the expertise we developed in the commercial sector for the government sector. Scale also provides improved operational efficiency because you have a singular relationship with both the technology and labor provider.

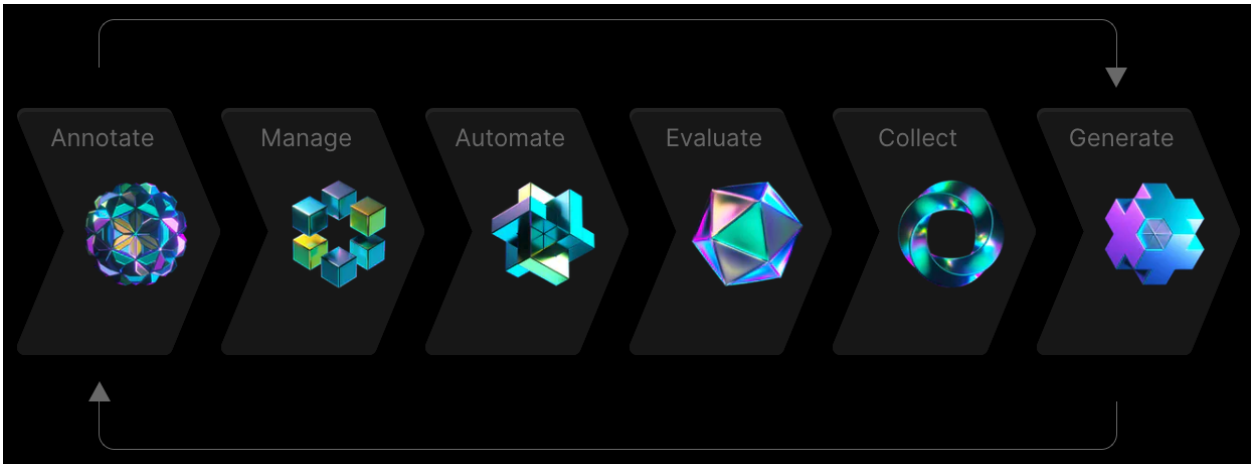



**Figure 1: Scale's data-centric ML lifecycle**

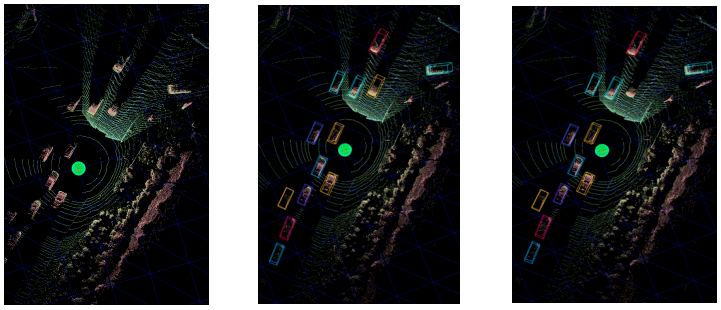
We can deploy object classification, obstacle avoidance, target tracking, BVLOS, facial recognition and 3D-reconstructed map models on a variety of aerial platforms. We can also accelerate multi-modal situational awareness, including vision-based object detection and classification & tracking.

Scale has a proven track record of understanding and meeting highly technical and nuanced customer needs — including many of the world's top machine learning teams and leading autonomous vehicle (AV) companies. For example, Scale annotates more ground-based autonomy perception data than any other company in the world. Scale currently executes billions of annotations per month, maps 3,800km of roads per week, and is the industry leading AI perception stack data solution across the globe.

# 2 Capabilities

## 2.1 Sensor Fusion:

Scale AI focuses on an end-to-end labeling solution that takes in individual task data through a REST API (Representational State Transfer Application Programming Interface) call. The API call can be customized according to the labeling format and taxonomy that the customer requires. The first annotation on a task is partially completed by a deep learning model trained on previous data. The task is then passed to a series of labelers, reviewers, and quality assurance staff before returning the annotated data back to the customer via a callback URL (for ease of technical integration) or manual download from the web application's dashboard. By leveraging a programmatic means of data submission and receipt, Scale enables complex use cases like Active Learning where our humans-in-the-loop can validate model predictions in real-time and ensures our API can be deeply integrated within developer workflows.For example, as shown below, even in complex 3D LIDAR scenes, auto-labeling can help drive significant efficiency and accuracy.

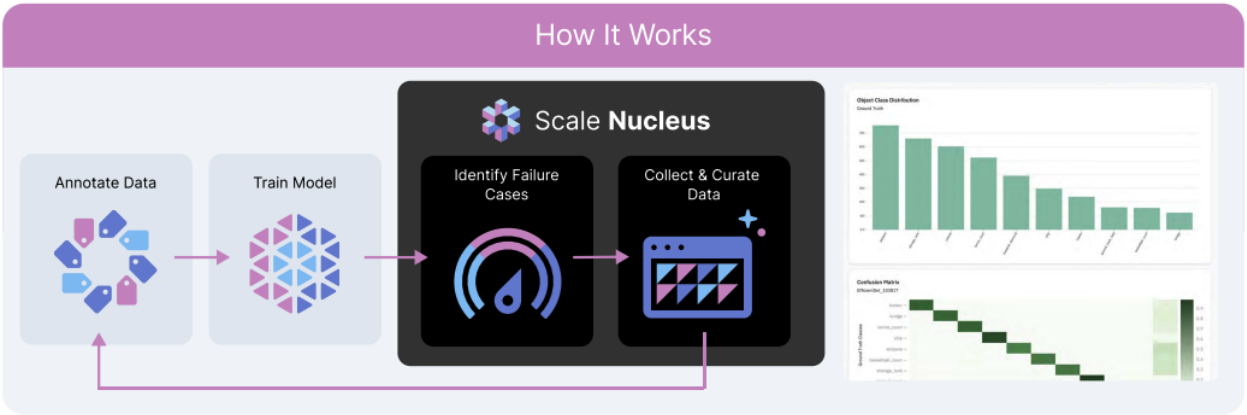


## 2.2 Scale Nucleus:

Nucleus is an AI data management tool which makes it easy to identify model and data issues such as missing labels, inaccurate labels, false positives, rare edge cases and much more. With Scale Nucleus you can also now quickly test, validate, and debug model performance for faster model iteration.

### 2.2.1 An Integrated AI Data Hub

The Nucleus Command Center can provide our partners with a centralized hub from which AI-related actions and analysis occurs. From the Command Center, users can access raw and



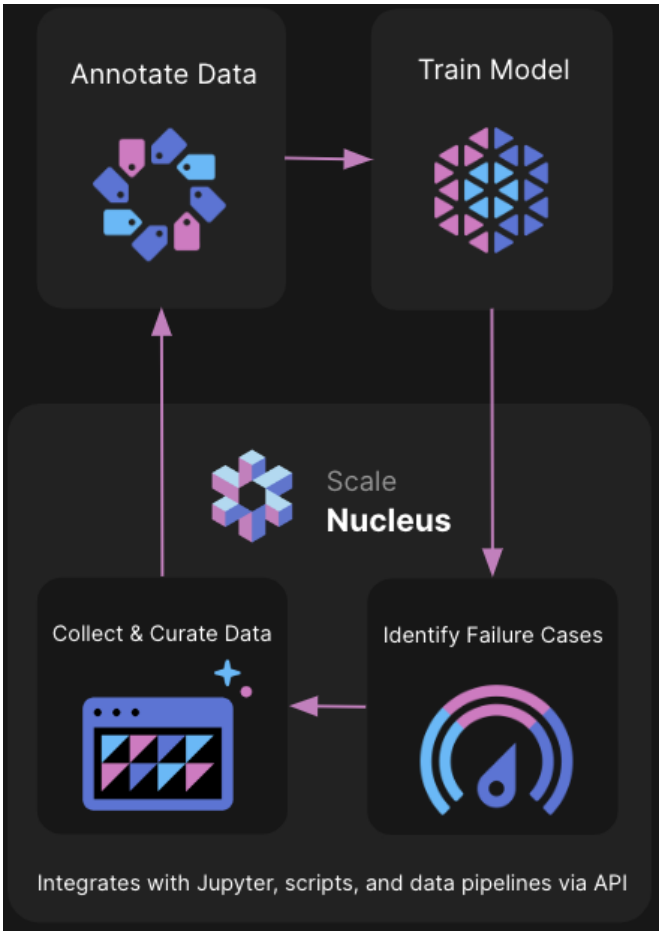
annotated data, discover datasets, set up annotation taskings, evaluate models in failure modes, visualize model performance, and more.

### 2.2.2 ML-based Data Curation and Smart Selection

Nucleus includes tooling that fuses human intuition and ML to select data that will provide maximum performance depending on the chosen model and assigned task. This means that users can efficiently select and assign data for labeling based on machine-assisted identification within minutes instead of days**.**

### 2.2.3 ML Model T&E and Debugging

Through the data and model integration capabilities within Nucleus, users can test, evaluate, and debug ML models directly alongside annotated data to quickly identify prediction errors. Nucleus includes a rich suite of built-in debugging tools and error metrics to surface qualitative and quantitative model failures, thereby simplifying the evaluation process and reinforcing model performance.




### 2.2.4 ML Data Enrichment

Using existing models and infrastructure within Nucleus, users can enrich their ML models and leverage years of ML model training based on petabytes of high-quality annotated data.

### 2.2.5 Efficient, Accurate, Scalable Data Annotation.

Nucleus can integrate Scale's best-in-class labeling network in addition to labeling work completed through third-party sources. Through the Scale Customer Management Portal, Scale's partners users can actively monitor data labeling tasks and quality. Scale also supports automated labeling quality feedback to ensure quality issues are detected early and corrected quickly.