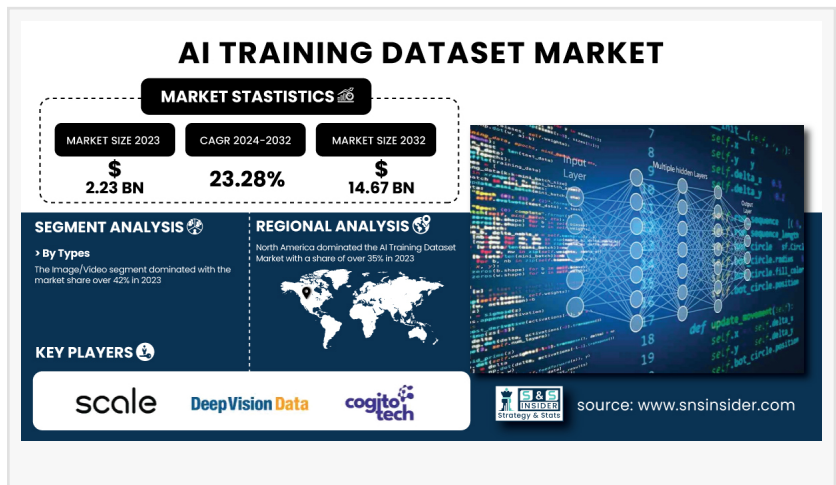


AI Training Dataset Market to Reach USD 14.67 Billion by 2032 | SNS Insider

The AI Training Dataset Market, valued at USD 2.23 Bn in 2023, is projected to reach USD 14.67 Bn by 2032, growing at a 23.28% CAGR from 2024-2032.

AUSTIN, TX, UNITED STATES, February 12, 2025 /EINPresswire.com/ -- According to SNS Insider, the [AI Training Dataset Market](#) was valued at USD 2.23 billion in 2023 and is projected to reach USD 14.67 billion by 2032, growing at a CAGR of 23.28%

during the forecast period of 2024-2032. The growing reliance on AI-driven applications across industries is driving the demand for diverse and high-quality training datasets, fueling market expansion.



Get Sample Copy of Report: <https://www.snsinsider.com/sample-request/3166>

Some of Major Keyplayers:

- Amazon Web Services Inc. (Amazon SageMaker Ground Truth, Labeling Services)
- Scale AI, Inc. (Data Labeling Platform, Sensor Fusion for Autonomous Vehicles)
- Deep Vision Data (Custom AI Training Data Solutions)
- Cogito Tech LLC (Image and Video Annotation, Text Data Labeling)
- Google LLC (Google Cloud AutoML, Dataset Search)
- Lionbridge Technologies, Inc. (AI Training Data Services, Multilingual Data Annotation)
- Alegion (Data Labeling and Annotation Tools, Video Annotation for Autonomous Vehicles)
- Microsoft Corporation (Azure Machine Learning, Custom Vision AI)
- Samasource Inc. (Data Annotation and Validation Services for AI)
- Appen Limited (Image and Speech Data Collection, Crowdsourced Annotation)
- iMerit Technology Services (Image Annotation, NLP Training Data)
- Figure Eight Inc. (Human-in-the-Loop Data Annotation Platform)
- Reality AI (Sensor Data Labeling for Industrial Applications)
- Playment (3D Bounding Boxes, Sensor Fusion Labeling for Autonomous Vehicles)

- Mighty AI (Computer Vision Training Datasets for Autonomous Vehicles)
- Trilldata Technologies (AI Data Engineering and Dataset Preparation)
- Clarifai (AI Model Training, Image and Video Annotation)
- Datasaur (Text Annotation and NLP Training Data)
- Labelbox, Inc. (AI Data Labeling and Collaboration Platform)
- V7 Labs (Image and Video Dataset Preparation, Automated Labeling Tools)

By Type: Image/Video Dominates with the Fastest Growth

The image/video segment dominated the market and accounted for the largest market share owing to a large number of applications of AI in computer vision, facial recognition, self-driving vehicles, and security surveillance. From diagnostics in healthcare through medical imaging, to customer behavior analysis in retail, to traffic monitoring, and surveillance systems in our smart cities, AI-based image and video recognition technologies are being deployed widely. This segment is further propelled by the growing adoption of deep learning algorithms that necessitate large-scale labeled datasets of image and video content.

The image/video is also estimated to have the largest CAGR over the review period, due to the increasing penetration of artificial intelligence PC and perk hardware, augmented reality technologies, deep within virtual penetration, and deepfake detection. Furthermore, the growth of AI-powered media and entertainment applications such as video content recommendation engines and automated video editing tools will also contribute significantly to the market growth.

By Vertical: The IT Sector Leads with Accelerated AI Adoption while Healthcare is Expected to Fastest Growth

The IT sector dominated the market and accounted for a significant revenue share in 2023, Owing to the heavy adoption of AI systems in various applications pertaining to cybersecurity, automation, predictive maintenance, and software development. AI-powered chatbots, IT service management, and automatic code generator tools benefit from good datasets for enhanced speed and precision.

The healthcare segment is expected to register the fastest CAGR during the forecast period as AI models will be more widely used to assist in medical imaging, drug discovery, patient monitoring, and personalized treatment plans. In the case of AI-powered diagnostic tools, having access to millions of correctly labeled CT scans, MRI images, and genomic datasets enables rapid adoption of training datasets in this field.

AI Training Dataset Market Segmentation:

By Type

- Text

- Image/Video
- Audio

By Vertical

- IT
- Automotive
- Government
- Healthcare
- Audio
- Retail & E-commerce
- Others

Enquiry Before Buy this Report: <https://www.snsinsider.com/enquiry/3166>

Regional Analysis: North America Leads, While Asia-Pacific Grows the Fastest

North America dominated the market and accounted for 35% of revenue share in 2023 owing to the extensive presence of AI research & development institutions, tech giants, and government-supported AI development projects. It has a matured AI Ecosystem in the region, where organizations including Google, Microsoft, and Amazon are pouring money into AI training datasets. DS systems helped to record further demand for AI applications in healthcare, financials, and autonomous systems.

The Asia-Pacific region is expected to record the fastest CAGR, owing to increasing investments in technologies in China, Japan, and India. The growth is driven by government initiatives promoting AI adoption, the growth of AI-based startups, and digital transformation projects. These factors are escalating the demand for AI training datasets in this region, in addition to the booming e-commerce sector, AI-enabled manufacturing automation, and the rising adoption of AI in smart cities.

Recent Developments in AI Training Dataset Market (2024)

- January 2024 – Google AI announced the launch of a new multilingual NLP dataset to improve AI language understanding across multiple languages.
- March 2024 – OpenAI introduced an ethically sourced dataset initiative aimed at reducing biases in AI models and enhancing the accuracy of AI-driven decision-making.

Access Complete Report: <https://www.snsinsider.com/reports/ai-training-dataset-market-3166>

Table of Content:

1. Introduction

2. Executive Summary
3. Research Methodology
4. Market Dynamics Impact Analysis
5. Statistical Insights and Trends Reporting
6. Competitive Landscape
7. AI Training Dataset Market Segmentation, By Type
8. AI Training Dataset Market Segmentation, By Vertical
9. Regional Analysis
10. Company Profiles
11. Use Cases and Best Practices
12. Conclusion

About Us:

SNS Insider is one of the leading market research and consulting agencies that dominates the market research industry globally. Our company's aim is to give clients the knowledge they require in order to function in changing circumstances. In order to give you current, accurate market data, consumer insights, and opinions so that you can make decisions with confidence, we employ a variety of techniques, including surveys, video talks, and focus groups around the world.

Akash Anand

SNS Insider | Strategy and Stats

+1 415-230-0044

[email us here](#)

Visit us on social media:

[Facebook](#)

[X](#)

[LinkedIn](#)

[Instagram](#)

[YouTube](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/785330826>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.