

## A Comprehensive Guide to Al Data Collection

## Have you ever wondered?

How do devices as simple as a watch accurately predict heart attacks in humans? How is it possible that cars and automobiles have always required a driver to suddenly go driverless on roads? How do chatbots make us believe that we are talking to another human on the other side?

If you observe the answer to every question, it boils down to just one element – DATA. Data lies at the center of all AI-specific operations and processes. It is data that helps machines understand concepts, process inputs, and deliver accurate results.



## Acquiring Al Training Data for your ML Models?



# Factors to consider for an effective Budget



### **Data Quality**

Well, like the very data we feed into our systems, quality has a lot of factors and parameters associated with it as well. If you reach out to AI experts or machine learning veterans, they might share any permutation of high-quality data is anything that is –



## Benefits of an end-to-end Al Training Data service provider

#### **Relevant Datasets**

Vendors understand your market segment, use cases, demographics, to fetch you the right datasets.

#### **Data Diversity**

Vendors source datasets such as images, videos, text, audio files, or all of these from diverse sources across the globe.

#### **Data Curation**

Vendors structure and tag the data with attributes and insights that machines understand.

#### **Data Annotation**

SME's from specific domains take care of annotating crucial pieces of information for greater accuracy.

#### **Data De-identification**

Adhere to HIPAA, GDPR, and other industry compliances to eliminate legal complexities.

#### **Data Bias**

Eliminate bias from datasets, ensuring you have objective results and inferences.

## Choosing the right AI Data Collection Vendor





www.shaip.com