

MODULE 2 — Types of Data Based on Structure

Data can be classified into three major categories based on structure.

Data Type	Organization	Examples	Flexibility
Structured Data	Fixed rows & columns	Spreadsheets, DB tables, Sales reports	Low
Semi-Structured Data	Flexible tags/markers	JSON, XML, Emails	Medium
Unstructured Data	No predefined schema	Images, Videos, Audio, Social posts	High

2.2 Structured Data

Structured data is highly organized and follows a predefined schema — stored in rows and columns for fast querying.

✓ Advantages	⚠ Limitations
Fast processing	Less flexible
Accurate analysis	Difficult to handle multimedia content
Easy management	

2.3 Semi-Structured Data

Semi-structured data does not follow a rigid table format but contains tags or markers for organization (e.g., JSON, XML, Emails).

✓ Advantages	⚠ Limitations
Scalable and adaptable	Slower complex queries
Human & machine-readable	Possible inconsistencies
Efficient data exchange	
Supports nested data	

2.4 Unstructured Data

Unstructured data has no predefined format or organization — rich in information but difficult to process, requiring advanced AI.

AI Technologies Required

- Natural Language Processing (NLP)
- Computer Vision
- Speech Recognition
- Deep Learning

KEY CONCEPT

Structured data tells AI what is happening. Unstructured data helps AI understand the world like humans.

2.5 AI Use Cases Based on Data Structure

Data Type	AI Applications	Models Used
Structured Data	Fraud detection, Sales forecasting	Regression Models, Decision Trees
Semi-Structured Data	Email classification, Log analysis	Transformers, NLP Models
Unstructured Data	Image recognition, Sentiment analysis	CNNs, LLMs