IBM AI course (MOD 3)

Structured, semi-structured, or unstructured data: What are the differences?

Data is raw information. Data might be facts, statistics, opinions, or any kind of content that is recorded in some format. This could include voices, photos, names, and even dance moves!

Data Types:

1. **Structured Data:**

   - Highly organized and typically quantitative.

   - Arranged in rows and columns, resembling a spreadsheet (e.g., Google Sheets or Microsoft Excel).

   - Examples include names, dates, addresses, credit card numbers, and stock information.

2. **Unstructured Data (Dark Data):**

   - Primarily qualitative and lacks inherent organization.

- Cannot be easily processed or analyzed with conventional data tools and methods.

- Examples encompass images, text documents, customer comments, medical records, and song lyrics.

3. **Semi-structured Data:**

   - Acts as a bridge between structured and unstructured data.

   - Does not adhere to a predefined data model.

   - Combines characteristics of both structured and unstructured data.

   - More complex than structured data but easier to store than unstructured data.

   - Utilizes metadata to identify specific data characteristics, enabling organization into records and preset fields.

   - Enhanced cataloging, searching, and analysis capabilities compared to unstructured data.

   - Example: A social media video is unstructured data, but it typically includes text elements (e.g., hashtags) for easy categorization and search, making it semi-structured.

> **Experts estimate that about 80% of all the data in today's world is unstructured. It contains so many variables and changes so quickly that no conventional computer program can learn much from it.**

Analyzing unstructured data

Imagine a computer sifting through billions of unstructured data entries. To make sense of it all, you'd need a program capable of sorting through countless possibilities and a vast list of keywords. Unstructured data holds answers in fields like disease prevention, crime, and finance. AI, inspired by the human brain, can bring order to this chaos, uncovering insights, learning from the data, and making increasingly accurate predictions. This is the transformative Era of AI.