Individual Task-II

MY DAILY DATA INVENTORY

(tracking the types and sources of data I interact with daily)

Introduction:

In the modern digital environment, every Individual continuously generates and interacts with data. From unlocking a smartphone in the morning to watching videos at night, data Is being created, stored, processed, and analyzed. This daily Interaction with digital systems forms our Daily Data Inventory.

The purpose of this group task Is to track the types and sources of data we interact with regularly and classify them into three major categories:

1. Structured: Organized in predefined formats, easy to store in tables.
2. Semi-Structured: Has markers or tags but no fixed schema.
3. Unstructured: Free-form data, including text, images, and multimedia.

Understanding this classification helps us connect theoretical concepts of Big Data with real-life experiences.

Understanding Data Types:

1 .Structured Data:

Structured data is highly organized and stored In a fixed format, usually In tables with rows and columns. It is easy to search, store, and analyze using relational databases.

Characteristics:

- Stored in relational databases
- Follows a predefined schema
- Easily searchable using SQL
- Numeric or categorical In nature

Examples in daily life:

- Bank transaction records

- Student attendance data
- Marks and grades
- Contact lists ⬜ Online order history

2. Semi-Structured Data:

Semi-structured data does not follow a strict tabular format but contains tags, labels, or metadata to organize elements. It is flexible compared to structured data.

Characteristics: ▪
- No fixed schema
- Contains markers like tags or key-value pairs
- Often stored in JSON or XML format
- Used in APIs and web applications

Examples in daily life:
- Emails (To, From, Subject, Body)
- Chat backups (JSON format)
- Website cookies
- Online form submissions
- Location logs

3. Unstructured Data:

Unstructured data has no predefined format and cannot be easily stored In traditional databases. It makes up the majority of global data today.

Characteristics:
- No fixed structure
- Complex to analyze
- Requires AI/ML tools for processing
- Stored in distributed storage systems

Examples in daily life:
- Photos and selfies
- Voice recordings
- Videos
- Social media posts

Daily Data Sources and Classification

Source / App	Data Type	Format	Classification
WhatsApp	Messages, images, videos	Text, multimedia	Semi-structured / Unstructured
Email	Emails, attachments	Text, HTML, PDF	Semi-structured / Unstructured
Social Media	Posts, comments, likes	Text, multimedia, JSON	Semi-structured / Unstructured
Mobile Banking	Transactions, account info	Tables, JSON	Structured
E-commerce	Products, reviews, orders	Tables, JSON, text	Structured / Semi-structured
Web Browsing	History, cookies, pages	Text, logs, HTML	Semi-structured / Unstructured
Fitness Apps	Steps, heart rate, sleep	CSV, JSON	Structured

Analysis:

-Structured: Numeric, tabular, easy to query (e.g., banking, fitness, IoT). -SemiStructured: Flexible, partially organized (e.g., emails, JSON logs, social media metadata).

-Unstructured: Free-form data requiring advanced tools (e.g., images, audio, PDFs).

-Most daily data is semi-structured or unstructured. Structured data is limited to finance, health, and IoT systems.

Observations:
-      Semi-structured and unstructured data dominate digital interactions.
-      Structured data allows easy storage and analysis.
-      Semi-structured data offers flexibility for partially organized information.
-      Unstructured data requires advanced processing like NLP, image, or audio analysis.

Conclusion:

Daily interactions generate a rich mix of structured, semistructured, and unstructured data. Proper classification helps in managing, storing, and analyzing this data effectively. While structured data is easier to handle, semi-structured and unstructured data require advanced processing techniques to convert information into actionable insights. The increasing reliance on apps, websites, and IoT devices has led to an exponential rise in semi-structured and unstructured data. Awareness of data types allows individuals and organizations to adopt appropriate storage, retrieval, and analytical strategies, ensuring efficient data-driven decision-making in daily life.