

# Retail Chain's Data Management Challenges

**Course Title:** CS504049 – Business Intelligence (BI)

**Instructor:** Phuc H. Duong

**Location:** Tan Phong Campus

---

## Background

A large retail chain with over 500 stores across multiple regions is facing difficulties managing and analyzing the data generated from its day-to-day operations. The company wants to improve its decision-making by leveraging its data, but it struggles to make sense of the different types of data coming from various sources.

## Types of Data in Use

1. **Structured Data:** The retail chain collects vast amounts of **structured data** daily. This includes:
  - **Point-of-Sale (POS) transactions:** Information about the products sold, quantities, prices, and timestamps.
  - **Inventory management:** Data on stock levels, restocking dates, and product categories.
  - **Customer information:** Data on customer loyalty programs, including names, contact details, and purchase histories.

This data is stored in well-defined rows and columns in databases, making it easy to query, analyze, and visualize. The company uses this data to monitor sales trends, inventory levels, and customer preferences.

2. **Unstructured Data:** In addition to structured data, the retail chain also generates significant amounts of **unstructured data** from sources like:
  - **Customer reviews and feedback:** Textual reviews left by customers on the company's website and third-party platforms.
  - **Social media interactions:** Comments, likes, and shares related to the company's products or campaigns on social media platforms like Facebook, Instagram, and Twitter.
  - **Email and chat logs:** Conversations between customer support teams and customers through email and chat.

This type of data does not have a predefined structure and is stored in formats like text, images, and audio, making it more challenging to organize and analyze without advanced tools like Natural Language Processing (NLP) or machine learning algorithms.

3. **Big Data:** The retail chain is generating **big data** due to the sheer volume, velocity, variety, and veracity of data it collects. With hundreds of stores, millions of transactions, and countless customer

interactions across digital platforms, the data grows rapidly and in different formats. The company also tracks real-time data like foot traffic in stores using IoT sensors and weather data to optimize inventory and staffing levels.

### Characteristics of Big Data:

- Volume: The company processes millions of transactions and customer interactions monthly, making it difficult to handle with traditional databases.
- Variety: The data comes in various forms – structured sales data, unstructured social media posts, images from in-store cameras, etc.
- Velocity: The data is generated in real-time, especially through online interactions and IoT devices, requiring the company to analyze it quickly to make timely decisions.
- Veracity: The accuracy and quality of data.

### Challenge

The retail chain is struggling to combine and analyze all this data effectively. While structured data from its transactional systems can be easily analyzed using standard BI tools, unstructured data like social media posts and customer feedback remains underutilized. Additionally, the sheer scale of the data presents a challenge to traditional data processing systems.

### Solution

The company decides to implement a **big data analytics platform** that can handle both structured and unstructured data. The platform uses distributed computing to process large datasets and machine learning algorithms to extract insights from unstructured data.

- **For structured data:** The company uses traditional BI dashboards to track sales performance, inventory levels, and customer loyalty.
- **For unstructured data:** The company implements NLP algorithms to analyze customer reviews and social media posts. This allows them to identify customer sentiment and emerging product trends.
- **For big data:** The company uses real-time analytics to monitor foot traffic and optimize store operations, ensuring that staffing and inventory are aligned with demand.

### Results

After implementing the big data platform, the company was able to:

- Improve decision-making by combining insights from both structured and unstructured data.
- Respond more quickly to customer feedback by analyzing reviews and social media interactions.
- Optimize inventory management and staffing levels using real-time data from IoT sensors and external factors like weather.

## Questions

1. What are the key differences between structured and unstructured data in this case?
2. How did the use of big data analytics help the company address its challenges?
3. What tools and techniques could the company use to further improve its data management practices?

-----