

# **Srinivas University**

Name : Adithya Acharya

USN : 01SU24CS011

Branch : CSE

Sem : 4<sup>th</sup>

Topic : Module 2 Group Task

---

# Big Data Process Mapping

## Example System: Google Maps

---

### 1. Introduction

Google Maps is a real-world big data system that processes massive amounts of real-time and historical data to provide navigation, traffic updates, route optimization, and location-based services.

It handles **billions of user requests daily**, making it a perfect example of a big data architecture.

This mapping explains the complete data flow:

**Data Sources → Data Collection → Storage → Processing → Output**

---

### 2. Data Sources (Input Layer)

Google Maps collects data from multiple sources:

#### 2.1 User Devices

- GPS location from smartphones
- Speed and direction of movement
- Search queries (e.g., “restaurants near me”)
- User reviews and ratings

#### 2.2 Satellites & Street View

- Satellite imagery
- Road images
- Geographic mapping data

#### 2.3 Traffic Sensors

- Smart traffic lights
- Highway cameras
- IoT road sensors

## 2.4 Third-Party Data

- Government road data
- Construction updates
- Business listings

These sources generate **structured, semi-structured, and unstructured data** continuously.

---

## 3. Data Collection & Transmission

1. Data is collected through:
  - Mobile apps
  - Web browsers
  - APIs
2. Data is transmitted via:
  - Internet networks
  - Secure cloud connections
3. Real-time streaming data is sent to cloud servers.

This stage involves **high-speed ingestion systems** capable of handling terabytes of data per second.

---

## 4. Data Storage (Big Data Storage Layer)

Because of the massive volume, data is stored in distributed systems.

### Storage Technologies Used:

- Distributed file systems
- Cloud storage

- NoSQL databases
- Data warehouses

Key characteristics:

- Scalability
- Fault tolerance
- Data replication
- High availability

Historical traffic data is stored for pattern analysis, while real-time data is temporarily stored for quick processing.

---

## **5. Data Processing (Analytics & Computation Layer)**

This is the most important stage.

### **5.1 Real-Time Processing**

- Detects traffic congestion
- Estimates travel time
- Suggests alternate routes

### **5.2 Batch Processing**

- Analyzes historical traffic trends
- Improves route prediction models
- Updates maps periodically

### **5.3 Machine Learning Algorithms**

- Predict traffic based on patterns
- Identify accidents or roadblocks
- Improve location recommendations

Processing frameworks divide tasks across thousands of servers.

---

## 6. Output Layer (User Services)

After processing, the results are delivered to users:

- Optimized routes
- Traffic updates
- Estimated arrival time (ETA)
- Nearby places suggestions
- Live navigation guidance

The output is displayed through:

- Mobile apps
- Web interface
- APIs for third-party apps

All of this happens in **seconds**.

## 7. Complete Data Flow Diagram



---

## **8. Big Data Characteristics in Google Maps**

This system clearly shows the 5 V's of Big Data:

1. Volume – Massive amounts of location data
  2. Velocity – Real-time streaming data
  3. Variety – GPS, images, text, sensor data
  4. Veracity – Accuracy and reliability checks
  5. Value – Useful navigation and insights
- 

## **9. Conclusion**

Google Maps is a strong example of a real-world big data system. It integrates:

- Massive data collection
- Distributed cloud storage
- Advanced analytics
- Machine learning models
- Real-time user interaction

This process mapping shows how raw data from millions of devices is transformed into meaningful navigation services within seconds.