ASSIGNMENT

INDIVIDUAL TASK-2

PRIVACY AND POLICY OF BIG DATA

Introduction to Big Data and Privacy

Big Data refers to extremely large and complex datasets that are generated from digital activities such as social media usage, online transactions, mobile devices, IoT sensors, healthcare systems, financial services, and government databases. Big data is commonly characterized by the **5 V's**:

- **Volume** – Massive amounts of data
- **Velocity** – Rapid generation and processing
- **Variety** – Different types (text, video, audio, sensor data)
- **Veracity** – Accuracy and reliability
- **Value** – Insights extracted for decision-making

Organizations such as Google, Amazon, Meta Platforms, and Microsoft collect and analyze massive datasets to improve services, personalize experiences, and generate revenue.

Understanding Privacy in the Era of Big Data:-

Data Privacy:

Data privacy refers to the protection of personal information from unauthorized access, misuse, or disclosure. It ensures that individuals have control over how their information is collected, stored, and shared.

Personal data may include:

- Name and contact details
- Biometric data

- Location data
- Financial information
- Health records
- Online behavior patterns
-

Why Privacy Matters:

Privacy is important because:

- It protects individual freedom and dignity

- It prevents identity theft and fraud
- It reduces discrimination risks
- It builds trust between users and organizations

Big data analytics can reveal sensitive patterns about individuals, including political opinions, medical conditions, or financial status—even if the original data appears harmless.

Major Privacy Risks in Big Data

Data Breaches:

Large datasets are attractive targets for hackers. Cyberattacks on organizations can expose millions of users' personal data.

Example: Data breaches in companies such as Equifax exposed sensitive financial information of millions of people.

Re-identification of Anonymous Data:

Even anonymized datasets can sometimes be re-identified by combining multiple datasets. Big data increases this risk because:

- More data points are available
- Advanced analytics can detect patterns
- AI systems can reconstruct identities

Surveillance and Tracking

Governments and corporations can use big data for monitoring citizens' activities. Technologies such as facial recognition and location tracking may lead to mass surveillance.

Profiling and Discrimination

Big data algorithms can create profiles based on behavior, credit history, shopping patterns, or social activity. This may result in:

- Biased hiring decisions
- Loan denial
- Insurance discrimination
- Targeted political manipulation

General Data Protection Regulation (GDPR)

The **General Data Protection Regulation (GDPR)** is one of the strictest privacy laws in the world. It applies to organizations handling data of EU citizens.

Key principles:

- Lawful and transparent processing
- Data minimization
- Purpose limitation
- Accuracy
- Storage limitation
- Integrity and confidentiality

California Consumer Privacy Act (CCPA)

The **California Consumer Privacy Act (CCPA)** gives California residents rights over their personal data.

Consumers can:

- Know what data is collected
- Request deletion
- Opt out of data selling

Challenges in Big Data Privacy

1. Rapid technological change
2. Cross-border data transfers
3. Cloud computing complexities
4. Enforcement difficulties
5. Lack of user awareness

Global companies operate across jurisdictions, making regulation enforcement complex.