How to ensure privacy protection?

Anonymization of data at
rest and in transit

As data volume and transactions continue to grow, every organization is concerned about data management and security. In the digital world, data is moving from on-premise to cloud at a faster pace to reduce operational cost and to increase scalability. The biggest concern is around personal data protection. There are regulations that exist for data protection like GDPR, PCI, HIPAA, etc. In addition to these, we need to protect/mask personal data in non-validated environments like training, development and testing. There are multiple tools and technologies available in the market for data masking/anonymization to avoid misuse and to be statutory compliant.

- The top implementation concerns or questions are: How to perform data anonymization/masking, and what are the tools and technologies that are available and apt?

- How to maintain data integration for the attributes that require masking?

- Will there be any impact on data availability SLAs?

- How to restrict data for certain users?

Effective data protection process:

A data protection process encrypts the data and removes personally identifiable information from data sets, so that the actual people whom the data reflects remain masked. Every organization sets its own methods to mask the data, based on the data types and data sources. Below are the best practices to follow for data in transit and data at rest.

When data in transit:

For data which is in movement from point to point either on premise or cloud, the original is replaced with dummy values. This is also called as dynamic data masking so that the end user can't view the original data. Follow the process below for data in transit:

- Follow and monitor data governance process to access incoming data

- Select the appropriate tools and technologies which can provide persistent masking values for same or repetitive incoming personally identifiable information (PII) values while data is in move (Process depicted in Figure 1)

- Ensure persistent masking values for unique incoming PII values. That will take care of the data integrity between the entities/objects. For e.g. 455033112 will always be masked as 566144223

- Apply proper design principles, and optimize the masking process during ETL (Extract, Transfer and Load) to achieve the SLAs
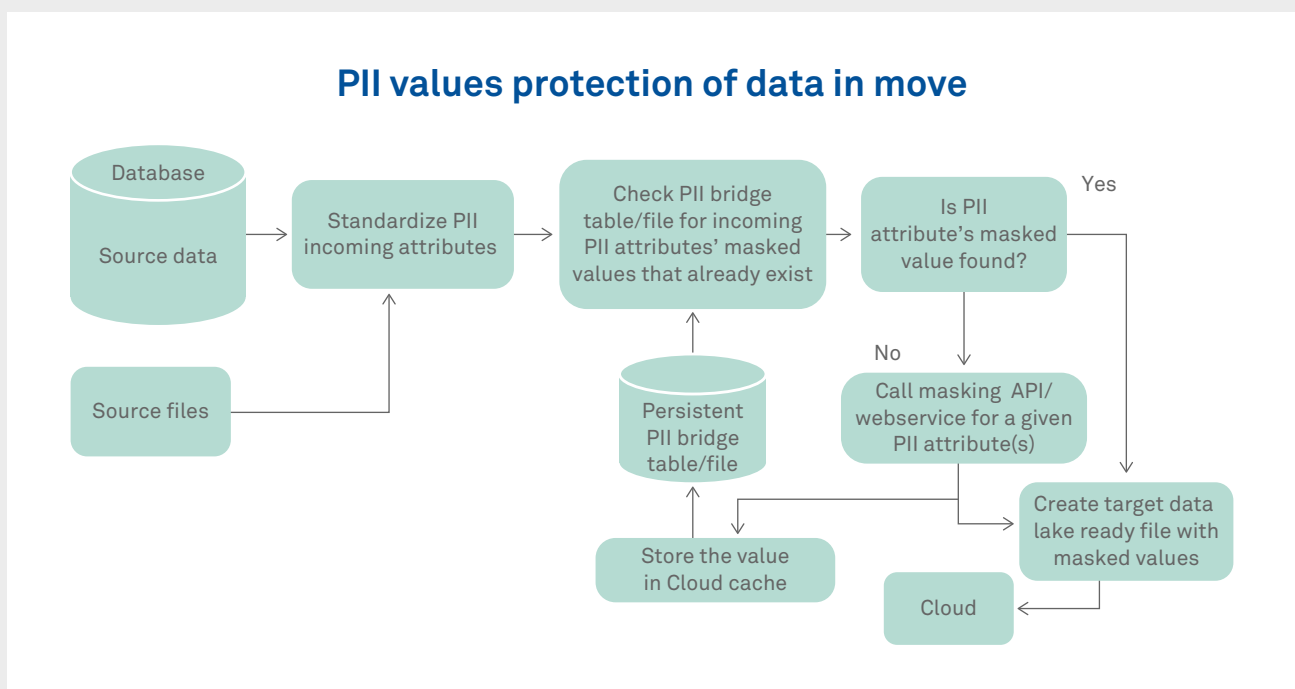
PII values protection of data in move

Figure 1: High level data-masking ETL flow of data in move

When data in rest:

PII data which are in rest either on-premise or cloud which are mostly used for analytics or training purpose, can be protected through proper data access governance process or by masking the data when user sends any requests through Reports or SQLs. Follow the process below when data is in rest:

- Apply and monitor data governance processes to access rest data from the end users either through reporting or direct access through SQLs (Process depicted in Figure 2)

- There are reporting tools which have active directory data access privileges and masking features. Adopt these to protect the PII data
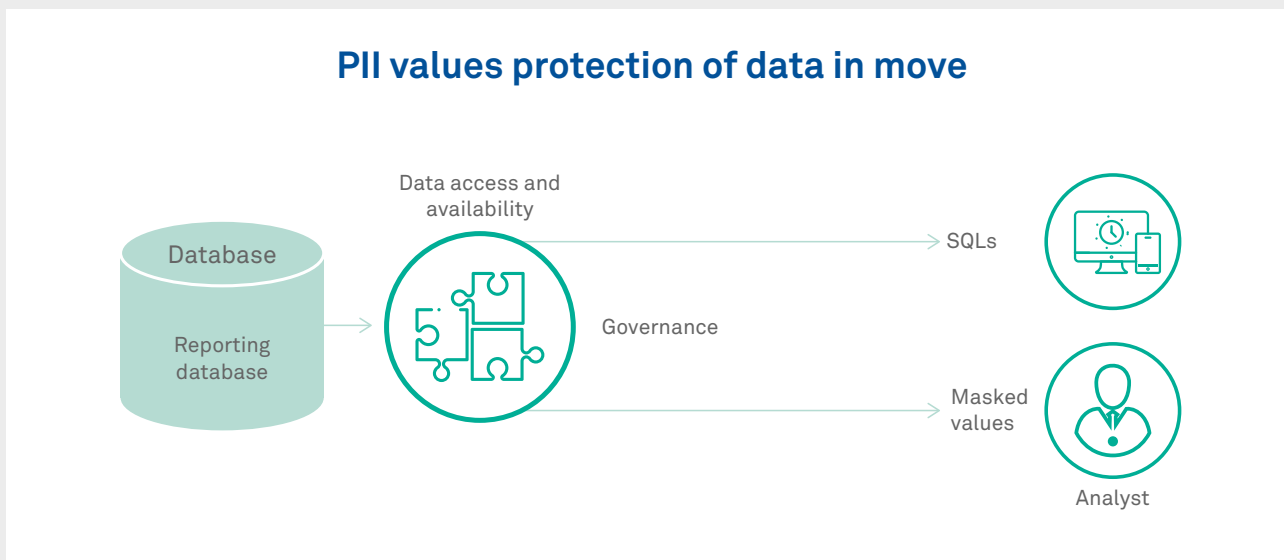
PII values protection of data in move

Figure 2: High level data masking of data in rest

Whether the data is in transit or at rest, a well-defined data protection process must be set by the organization to make sure unauthorized users do not access the data. Data anonymization makes your data statutory-compliant and gives users confidence in their data.

About the author

Purushottam Joshi,
Senior Architect
Data, Analytics & AI, Wipro

Purushottam has over 20+ years of data warehouse and ETL experience. He is currently focused on open source integration technologies and has successfully executed large engagements for global companies. He is a TOGAF certified Enterprise Architect. He is also certified in different database and ETL technologies and supports the practice in managing both, cloud and on premise native ETL tools.

● **Wipro Limited**
Doddakannelli, Sarjapur Road,
Bangalore-560 035, India

Tel: +91 (80) 2844 0011
Fax: +91 (80) 2844 0256
wipro.com

Wipro Limited (NYSE: WIT, BSE: 507685, NSE: WIPRO) is a leading global information technology, consulting and business process services company. We harness the power of cognitive computing, hyper-automation, robotics, cloud, analytics and emerging technologies to help our clients adapt to the digital world and make them successful. A company recognized globally for its comprehensive portfolio of services, strong commitment to sustainability and good corporate citizenship, we have over 175,000 dedicated employees serving clients across six continents. Together, we discover ideas and connect the dots to build a better and a bold new future.

For more information, please write to us at **info@wipro.com**