# Secure PII at its "Startpoints" … Data-Centrically

**When it comes to** personally identifiable information (PII) in points along our networks, people usually think first about "endpoint security" tools and techniques. They have been the most obvious and advertised way to protect the PII we store in our mobile phones (or provide through apps), as well as the laptops, desktop PCs, servers, and networks/switches through which they connect. We also think about storage devices like thumb and hard drives, or the folders, files, and entire databases that can be encrypted within them.

Fortunately, our industry also thinks about securing this data as it enters and moves around transaction and analytic systems. IRI, The CoSort Company, has begun to more broadly refer to the data-centric protection of these original PII values as *startpoint security*. Within that new definition, featured below, are many targeted ways to find, secure, and account for the PII in the databases and files our applications use. After all, it is in those repositories where PII is created, stored/queried, processed, and/or moved along the endpoints.

Indeed, IRI has focused historically on the granular protection of PII in data silos at rest (static data masking), and in transit (dynamic data masking). Securing PII directly in these sources or startpoints instead of just its endpoints continues to provide several benefits, including:

- **Efficiency—** It's much quicker (and less resource-intensive) to encrypt or apply other de-identification functions to discrete values than to everything else around them.
- **Usability—**By masking only what's sensitive, other data around it is still accessible.Those secured values can also move safely between databases, applications, and platforms (on premise or in the cloud).
- **Breach nullification—**Any misappropriated data is already de-identified.
- **Accountability—**Data lineage and audit logs pointing to specific element protections are a better way to verify compliance with privacy laws applicable to specific PII (identifiers).
- **Security—**Some data masking techniques cannot be reversed, and many applied at once are harder to reverse than a single technique. Think as well about the difference in vulnerability here, too; i.e., if the encryption key is compromised, an entire network could be exposed instead of a single column of data.
- **Reversibility—**If the masked PII needs to be reversed, that is possible if a function like encryption or lookup-pseudonymization was used.
- **Testing—**Masked production data can also be used for prototyping DB and ETL operations, platform benchmarking and Devops.

The actual de-identification of the data is nonetheless only one piece of a larger approach.

## IRI DEFINES NINE ASPECTS OF STARTPOINT SECURITY

At first, it may seem that we're just coining another buzzword, or a euphemism for data masking. However, data masking is just an included part of startpoint security. Under IRI's definition, startpoint security also takes into consideration eight more related pieces:

1. **Permission & Disclosure—** authorizing you to store submitted PII via user agreement
2. **IAM & RBAC—**managing access to data sources, (un)masking jobs, and programs
3. **Discovery & Classification—** searching and cataloging PII to find and mask it consistently
4. **Data & Metadata Lineage—**saving and analyzing changes to data and masking jobs
5. **Latency—**architecting and configuring static or dynamic data masking jobs
6. **Risk Scoring—**determining the statistical likelihood of re-identification (for HIPAA)
7. **Audit Logs—**seeing or querying who did what, and who saw what, when, and where
8. **Assessment & Insurance—** conducting expert procedural, statistical, and legal reviews

Many of these additional considerations are not exclusive to startpoint security, but data classification, lineage, latency, and re-ID risk scoring are certainly more relevant in the data-centric realm than the are to endpoint security. And each one of them could take up an entire article this size just to introduce them in more detail.

## STANDALONE OR INTEGRATED?

Data masking products and practices are typically deployed in response to a breach or to help comply with a particular privacy law. Most data masking tools usually provide just one or two of the capabilities listed above, and analysts recognize such vendors as narrowly.

IRI has endeavored to offer more in this space by addressing these many well-defined concerns, and combining them in its standalone "FieldShield" product. IRI also includes FieldShield in its larger Voracity data management platform where data discovery, integration, migration, governance, and analytics co-exist in one pane of glass, are powered through a one-pass manipulation engine (CoSort), and supported by one metadata infrastructure. ■

IRI DATA PROTECTOR
www.iri.com/products/iri-data-protector