



IRI DarkShield
Unstructured Data Search & Security

- Classify PII and Discover It Everywhere
- Extract Values and Report on Metadata
- Mask Simultaneously, Later, or in Batch
- Support TDM and Compliance Goals

Discover, Deliver, and Delete PII in Dark Data Sources



Product Summary

The Dark Data Conundrum

The Challenge

Gartner defines **dark data** as the "information assets organizations collect, process and store during regular business activities, but generally fail to use for other purposes" like analytics or digital business ... Organizations often retain dark data for compliance or archive retrieval. Its storage and security "typically incurs more expense (and sometimes greater risk) than value."

Source: Gartner IT Glossary

Knowledge of business and regulatory requirements

Classification of sensitive data and its authorized recipients

Implementation of policies and techniques that support these requirements and protect PII

PII search, remediation, and reporting techniques are particularly challenging in dark data environments due to volume, variety, and unstructured nature.

Enter DarkShield V6

IRI DarkShield supports the risk and controls framework in enterprise IT environments by classifying, finding, extracting, masking, and reporting on PII and other data 'hidden' in structured, semi-structured, and unstructured sources on-premise or in the cloud.

DarkShield rapidly scans and masks data in RDB C/BLOB columns and NoSQL databases, plus files in EDI, text, Parquet, SQL, image, audio, PDF and MS Office formats ... in LAN folders or Azure, GCP, Amazon S3 and SharePoint Online cloud stores. API users can source data even further afield.



DarkShield discovers PII in default or user-defined data classes that match:

- 1) values stored in a lookup table (either exact or fuzzy matches)
- 2) named 'path' or column filters for JSON, XML, CSV, RDB, and Excel sources
- 3) AI-models provided for relational database value discovery
- 4) stored or new Java Regex patterns, or segments in HL7, X21 and FHIR EDI files
- 5) drawn (bounding box) regions in images, or handwritten signatures in documents
- 6) machine-learned named entities in OpenNLP, PyTorch or Tensorflow NER models
- 7) OCR, OCR-A and MICR font recognition in images

Whenever DarkShield finds a match, it applies the masking function you assigned to a "rule matcher" associated with data meeting specific search criteria. DarkShield can optionally detect and redact signatures in documents, and in V7 will detect, recognize and protect handwritten PII and faces.

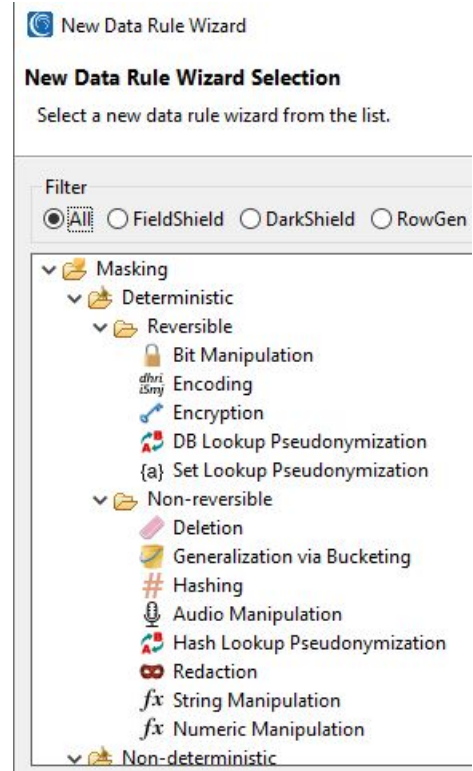
Searching and masking operations can be combined or performed separately, from a GUI wizard, batch job (CLI), API call, or web service. DarkShield can also extract and leverage search results and attendant metadata in HTML5 charts, delimited and JSON logs, DSARs, textual ETL, and SIEM/SOC tools.

Masking Functions

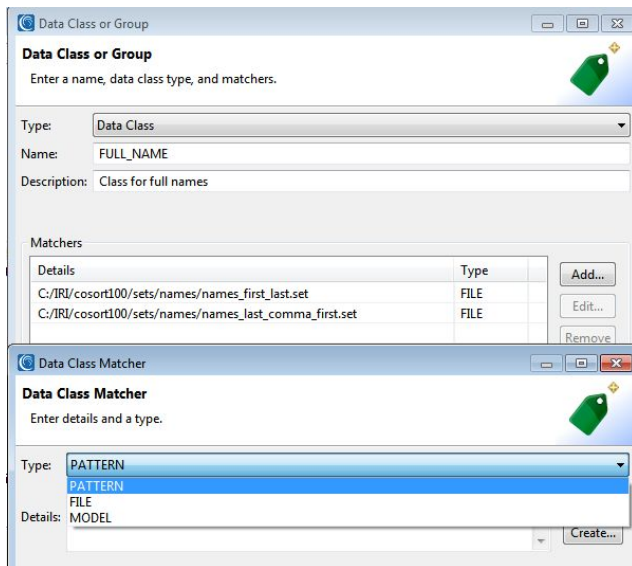
DarkShield users can employ the same deterministic and non-deterministic [data masking functions](#) that FieldShield users have to secure PII in structured sources for years.

Frequently applied out-of-the box methods include:

- Format-preserving (or not) encryption
- Lookup pseudonymization
- Redaction / obfuscation
- String manipulation
- Deletion / erasure
- Scrambling
- Fabrication
- Encoding
- Hashing
- Blurring



The masking rule you match to each data class should depend on the desired results for the ciphertext; i.e., whether it could be reversed, how it should appear (conform to format constraints), and if each value must be unique. DarkShield can replace existing or create new files in current or cloned folder structures.



Business Benefits

Only DarkShield supports the combination and automation of the difficult but necessary processes of [data classification](#), discovery, extraction, masking, *and* audit reporting across multiple data sources and silos *all at once*.

Multiple search methods and threads are deployed in conjunction with multiple data masking functions. All this optimization and consolidation speeds compliance efforts, application deployment, systems testing, and document management activity.

DarkShield also:

- helps [GDPR](#) data collectors and processors comply with data portability, rectification, and right-to-be-forgotten provisions.
- operates in the same free Eclipse™ job design and metadata environment, called [IRI Workbench](#), with FieldShield and other data governance and management functions.
- is affordable for perpetual use standalone, in a bundle with other [IRI Data Protector](#) suite products, or for “free” inside IRI Voracity total data management platform subscriptions.

USER INTERFACES

- OpenAPI** (Swagger Web Page)
- IRI Workbench IDE** (Graphical Job Wizards)
- Command Line** (Curl Commands or DarkShield CLI App)

TASK SPECS

Search Contexts (Data Classes)

Search Matchers

- Location Matchers
- Data (Content) Matchers

Masking Contexts

Data Rules

- Masking Functions
- Generation Functions

TRANSFER PROTOCOL

TLS https://

ENGINES

REST API

- Base (text) RPC
- Files RPC
- RDB RPC
- NoSQL RPC

AUDIT LOGS

- JSON
- CSV
- Dashboards
- Log4j2

DarkShield Aggregate Results Report

The hidden below are the top 100 unmasked records by the number of PII matches to DarkShield data class search criteria. Hover over each bubble to reveal the data source containing PII and the number of matches found.

Aggregate Data Source Report

DATUM NAME	DATA CLASS	NUMBER OF MATCHES	MASKED STATUS
file://P:/SampleData/xlsx/excel Tests/Performance Tests/bigData500k.xlsx	First_Name	1056881	UNMASKED
file://P:/SampleData/xlsx/excel Tests/Performance Tests/bigData500k.xlsx	Last_Name	824483	UNMASKED
file://P:/SampleData/json/rows.json	Phone_Number	544926	UNMASKED

Supported Approaches and Data Sources

DarkShield API users can also find and mask PII in other sources through “glue code” in any language.



Compatible Platforms and Applications

- Amazon S3
- Azure Blob
- GCP
- OneDrive
- SharePoint
- LAN drives (SMB)

IRI Workbench Integration

- IRI FieldShield - DB and flat-file masking
- IRI CellShield - Excel spreadsheet masking
- IRI CoSort - Data transformation and reporting
- IRI Voracity - Big data integration, test data, etc.

2194 Highway A1A

Melbourne, FL 32937 USA

1.321.777.8889 * 1.800.333.SORT



iri.com/darkshield