

IRI Data Protector Suite











IRI, The CoSort Company

Vendor Background

- Specializing in data management and data protection
- Innovator in data manipulation functionality and speed
- 7/8 products share 1 metadata and Eclipse GUI
- Privately owned since 1978
- Headquartered 1 hour southeast of Orlando, FL
- Resellers in more than 40 international cities



Selected IRI Data Masking Customers

Most IRI customers process and protect PII in databases, flat-files, or "big data" repositories, including Hadoop or NoSQL DBs. Others are building or testing applications. IRI's data masking consultants and clients are mostly in healthcare and BFSI.











































IRI Data Manager Suite



Speed or replace batch, BI, ETL, sort, and SQL programs

- Filter, join, aggregate, pivot, cleanse, lookup, calc, etc.
- Map, migrate, federate, and replicate data from 150 sources
- · Segment data, capture changes, report details / summaries
- Analyze changing dimensions, support complex transforms



Speed RDBMS unloads for archival, migration, reorg, and ETL

- Extract tables to flat files in parallel using SQL queries
- Convert and re-format to change data types and layouts
- · Create the data definitions for IRI software and DB loads
- Pipe to CoSort and DB loaders for faster reorg and ETL



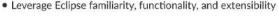
Unlock data and move between apps, DBs, and platforms

- · Convert, federate, remap, and replicate legacy data
- Migrate data between databases and create new tables
- Change file formats, data types, and endian conditions
- · Find, extract, and structure data in dark data documents



Consolidate tools and tasks to process, protect, prototype, present

- Discover, define, and manage data in legacy and new sources
- Combine data integration, migration, governance, and analytics
- · Exploit CoSort and Hadoop engines for optimum throughput



© 2018, Innovative Routines International (IRI), Inc. All Rights Reserved. CoSort, Voracity, NextForm, FieldShield, CellShield, and RowGen are registered trademarks of IRI. FACT is a trademark of DataStreams Corp. Chakra Max is a trademark of WareValley Co. Ltd. Other product, brand, or company names may be (registered) trademarks of their respective holders.

IRI Data Protector Suite



Comply with privacy laws, nullify breaches, govern data

- · Search, profile, and classify sensitive data in DBs and files
- Encrypt, hash, redact, pseudonymize, randomize, tokenize
- Apply cross-table rules to save time and referential integrity
- Create an XML audit log of each job to verify compliance



Profile and de-identify PAN/PHI/PII in Excel spreadsheets

- Define or use patterns to search for sensitive data
- · Locate, report, and open all found ranges in the LAN
- Click to encrypt, mask, or pseudonymize data directly
- Auto-log protections to verify privacy law compliance



Prototype DBs and ETL, stress-test, outsource, benchmark

- Use real data models and formats, not production data
- Combine generation and selection, create new formats
- Preserve referential integrity and frequency distributions
- · Feed test DBs, files, and custom reports simultaneously



Define, monitor, block, and audit DB activity using RBAC

- High-volume, data-centric audit and protection (DCAP)
- · Monitor, block, alert, and log users in real-time
- · Low-impact on DB performance and availability
- · Scan for, and dynamically mask, sensitive data





The CoSort Company

Sources



cloudera























mongo DB SQL Server SYBASE DATASTAX: ALTIBASE

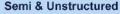


COBOL, CSV, LDIF, LS-RS-VS, MFVL. Text. VB. Vision, XML

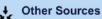


Adabas, Datacom, IDMS, IMS, ISAM, Pick, Unidata, VSAM, etc.









Custom Apps, ETL/ELT Tools, Packaged Apps, Web Logs

DISCOVER

Data Classification Dark Data Search DB & File Profiling **ER Diagramming** Metadata Definition Metadata Forensics Multi-Method Search

IRI Voracity

An Insatiable Appetite for Data

INTEGRATE

Slowly Changing Dimensions Public/Private Mashups Change Data Capture Fast DB Un/Load Data Federation One-Pass ETL

GOVERN

Data Quality Data Masking **Data Lineage** Data Reconciliation **Test Data Generation** Metadata Management

MIGRATE

Incremental Replication Data & File Types Endianness **Databases** JCL Sorts **ETL Jobs**

ANALYZE

Embedded BI Cloud Dashboard **Data Preparation Predictive Analytics BIRT & Splunk Feeds** Clickstream Analytics

DESIGN

Wizards with Rules **Graphical Dialogs** Scripts with Outlines **Workflow Palettes** Mapping Diagrams Metadata Form Editors **ADS Mapping Manager**

DEPLOY

CoSort CLI/API (SMP) MapReduce (Grid) Spark (In-Memory) Storm (Streaming) Tez (Batch) Java, Kafka, SQL Eclipse & Other Job Launchers

Targets













Hortonworks























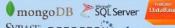












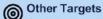








COBOL, CSV, LDIF, LS-RS-VS. MFVL, Text, VB, Vision, XML



Custom Apps, Data & SpreadMarts, ETL/ELT Tools, Federated Views, Packaged Apps, Test Suites

































IRI Data Protector Suite

What FieldShield Does

- Connect and interact with multiple source and target repositories, on-prem or cloud
- Discover and classify sensitive data in DB, flat-file, and dark-data sources
- Secure fields with PII, PHI, etc. via 12 built-in masking function categories
- Address multiple protections and recipients in one job script, one I/O
- Apply protection rules across tables and preserve referential integrity
- Support conditional security, i.e. based on patterns, values, or ranges
- Specify protections and layouts in Eclipse GUI and portable 4GL job scripts
- Integrate with DB apps via ODBC and .NET and Java SDK for dynamic data masking
- Retain data realism (e.g. FPE and pseudonyms) for testing and outsourcing
- Operate concurrently in big data EGL, migration, sub-setting, and BI/analytic jobs
- Log job and system runtime detail to an XML audit file to verify compliance
- Within Voracity, support streaming input and Hadoop execution paradigms



FieldShield Data Sources (Native)

Ac	cucobol Vision	Delimited	MaxDB	SQL Server
Alt	tibase (FACT)	Derby (WB)	Mongo (WB)	SQLite
A	ASN.1 TAP3	ESDS	MF-ISAM	Sybase ASA/E & IQ
В	IRT DB (WB)	Excel (WB)	WF Var. Length	Tibero (WB)
ВІ	RT Hive (WB)	ELF web logs	MySQL	Teradata (WB)
BIF	RT JDBC (WB)	Fixed	Oracle	Text
BIF	RT POJO (WB)	Heap / print	Outlook (WB)	UTF-8 & 16
	C-ISAM	HSQLDB (WB)	PDF (WB)	Variable Block
C	CLF web logs	IDX 3, 4 & 8	PostgreSQL	Variable Sequential
	CSV	Informix	Powerpoint (WB)	VSAM MVS (UniKix)
	DB2 (UDB)	Ingres	Record Sequential	Web Services (WB)
DB2	? for i5/OS (WB)	LDIF	RTF (WB)	Word (WB)
DB2	2 for z/OS (WB)	Line Sequential	SQL Anywhere	XML



FieldShield Data Sources (Legacy)

Access	D3	GA-Power 95, R91	K-ISAM	Pathway	RMS
Adabas	Datacom	Gemstone	Knowledgeman	PDS	Reality/X
Advanced Pick	Dataflex	GENESIS	KSDS	PervasiveSQL	RRDS
ALLBASE	Db4o	Gigabase	Lotus	Pick/Pick64+	SAP HANA
Alpha5	dBase	H2	Manman	PI-Open	Sequoia
Amazon RDS	Desktop Adapter	IDMS	Mentor / pro	Powerflex	Sharebase
Azure	DL/1	IDS	MO	Powerhouse	Supra
BizTalk	DSM	Image	Model 204	Progress	Terracotta
Cache	Enscribe	IMS	Mumps	QueryObject	Total
Clipper	Enterprise Adapter	Interbase	MyBase	rBase	Ultimate
Codasyl	FileMaker	Intersystems	Netezza	R83	UltPlus
CorVision	Firebird	ISM	NonStop SQL	Rdb	Unidata
ConceptBase	Focus	Jasmine	ObjectStore	REALITY	Universe
D-ISAM	FoxPro	JBase	Paradox	Red Brick	VSAM VSE



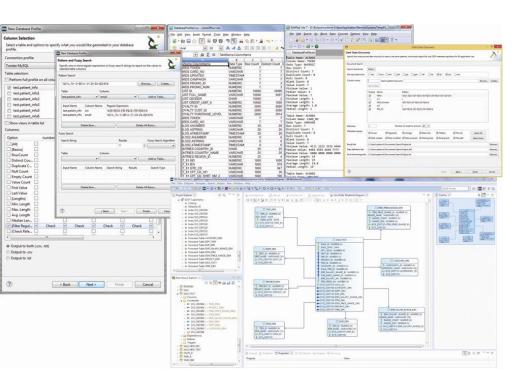
FieldShield Data Sources (Modern)

Amazon EMR Hive	FinancialForce	Marketo	Pivotal Greenplum
Apache Cassandra	Force.com apps	MongoDB	Pivotal HD Hive
Apache Hadoop Hive	Hortonworks Hive	MS Dynamics CRM	Salesforce.com
Cloudera CDH Hive	Hubspot	MS SQL Azure	ServiceMAX
Cloudera Impala	Lightning Connect	Oracle Eloqua	Spark SQL
Database.com	MapR Hive	Oracle Service Cloud	Veeva CRM



Sensitive Data Discovery - Multiple Wizards

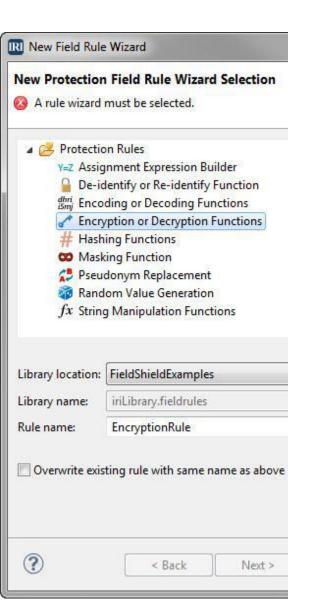
FieldShield (and all IRI software) includes PII discovery capability, which includes cross-source data identification and classification, string (literal or dictionary), pattern, and fuzzy-logic searches, statistical reporting, and automatic metadata creation. Fit-for-purpose GUI wizards deliver:



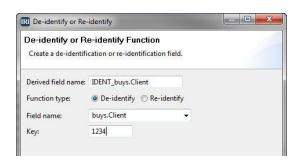
- DB and file classification, with rule matcher libraries
- DB profiling and E-R diagramming
- Dark data discovery and structuring, with metadata reporting
- Flat-file statistical and value searching
- Structured data metadata discovery and definition



Static Data Masking Functions (1-3 of 12)

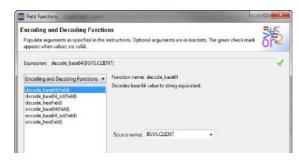


De- & Re-Identification



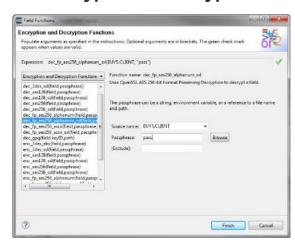
- For ASCII data
- Less secure
- Reversible

Encoding & Decoding



- Converts binary to ASCII
- Supports base64 & hex
- Reversible

Encryption & Decryption

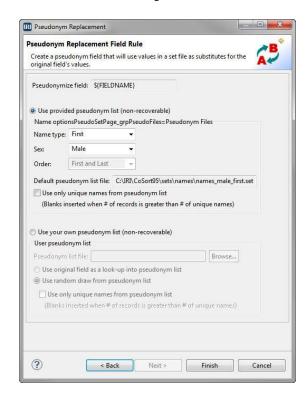


- 3DES EBC & SSL
- AES-128 & -256 CBC
- AES-256 Format-Preserving
- GPG (PGP-compatible)
- FIPS-compliant OpenSSL
- Custom



Static Data Masking Functions (4-6 of 12)

Pseudonymization



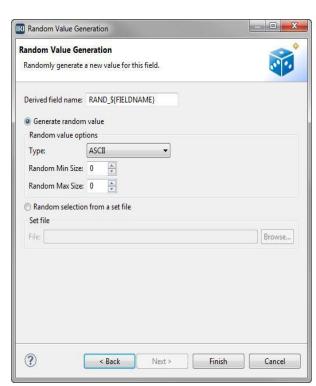
- Provides realistic names
- Reversible lookup values
- Non-reversible selection

Character Masking



- Partial/full-field masking
- Conditional omission
- Non-reversible

Randomization

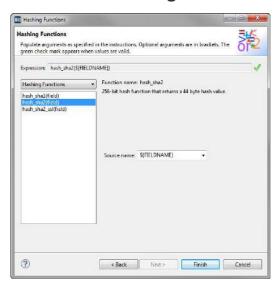


- Random data generation
- Random data selection
- Non-reversible



Static Data Masking Functions (7-12 of 12)

Hashing

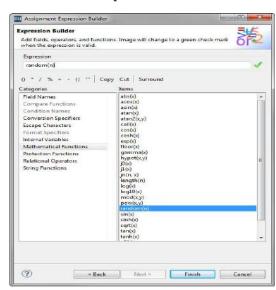


- SHA-1 & 2 cryptographic
- Returns hash of fieldstring
- Use for integrity checking

#10 Hashing

Target layout declaration with or without selection logic

Expressions

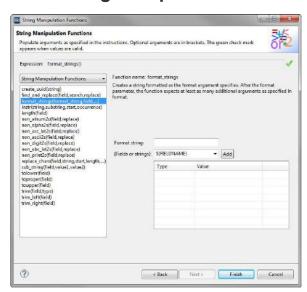


- Mathematical operations
- PCRE logic
- Custom blurring

#11 Tokenization

DB-value substitute for PCI DSS

String Manipulations



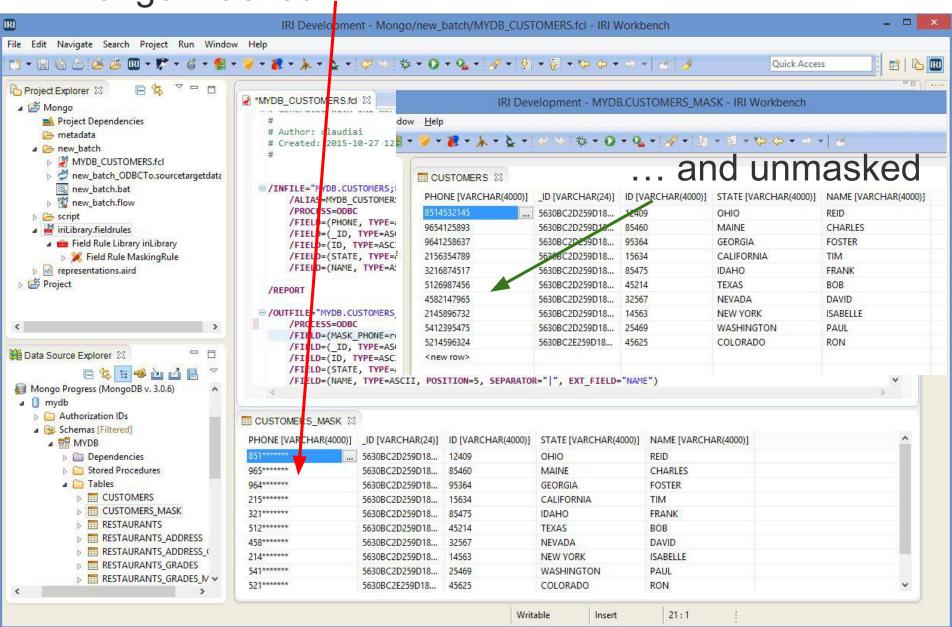
- Find, replace, and add
- Reposition and trim
- Use INSTR information

#12 Custom Function

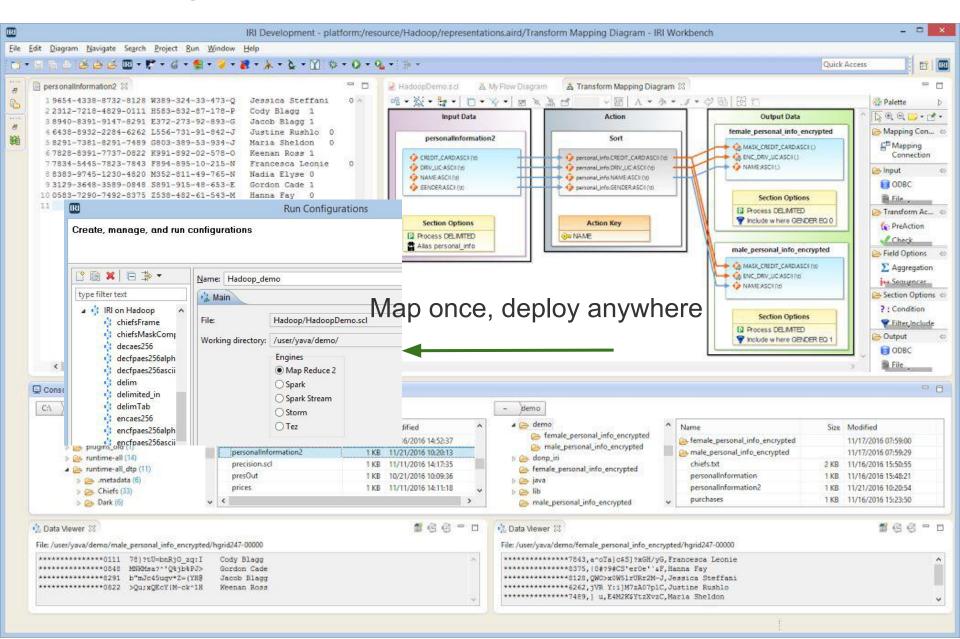
User's field-level call



Mongo Masked



Masking et al in Hadoop, Too



IRI Dynamic Data Masking Options

Method	Operation
ODBC Select / Update	Apply protections with precision to any given column value(s) in qualifying row(s)
DB App Invocation	Use .NET or Java SDK library functions or system-call job scripts on the fly
In-Situ Redaction	User and SQL-specific full and partial column masking on query (Chakra Max)
Custom I/O Procedures	Drive real-time application data directly to/from FieldShield jobs in memory
Real-Time Processing	Hadoop Spark and Storm processing of dynamic input streams in Voracity

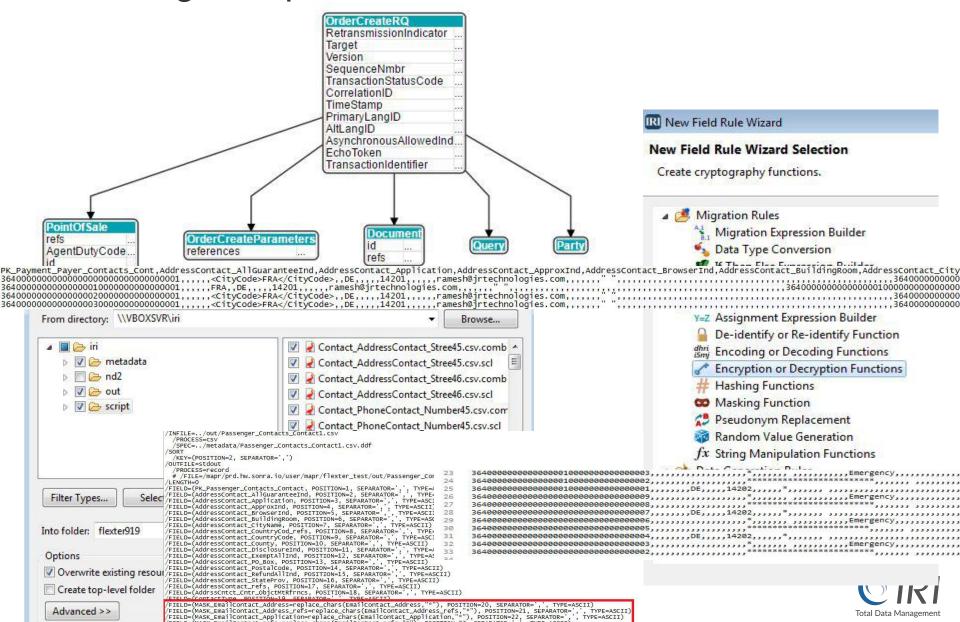
Encryption Key Management Options

- 1. Passphrase (key string) embedded in script
- 2. String as environment variable
- 3. String in (securable) key file
- 4. Multi-factor authentication via Towsend Security Alliance Key Manager

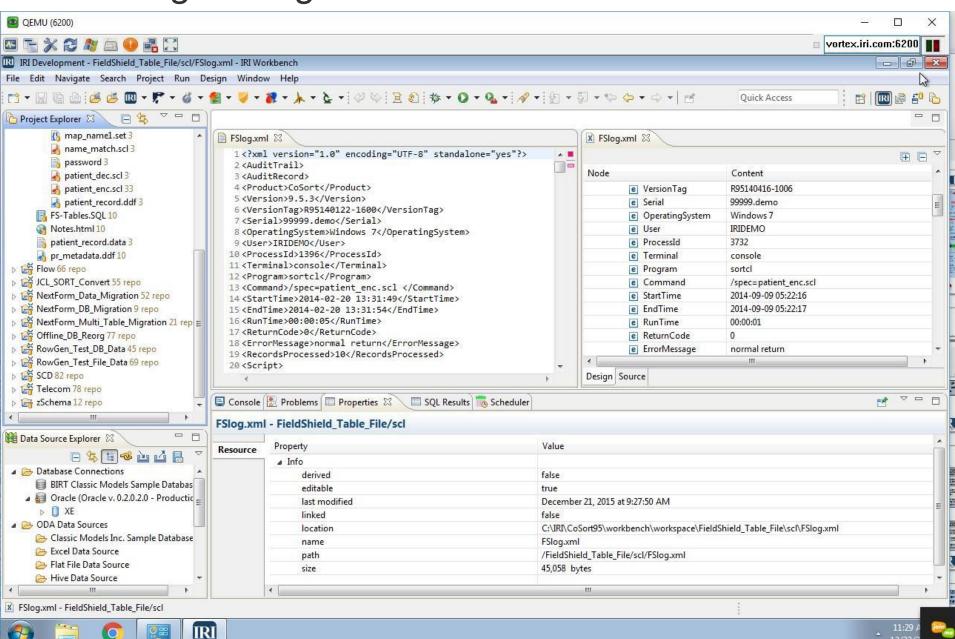


Advanced >>

Masking Complex XML (via Sonra Flexter pre-parsing, JSON next!)



Audit Log Management



User Profiles

 Vertical industries and governmental agencies storing, processing, or outsourcing applications with sensitive data, such as:

Banks

Health Care

Census / Tax

Insurance

Defense

Schools

- Application, DB, and DW users handling sensitive data
- CISOs, compliance teams, consultants, IT managers, and solution architects



Use Cases

Tesco Bank/RBS UK

- Decrypt and re-encrypt fields in credit card migration and test files
- Generate and manage encryption and user ID keys
- Other projects protect 38,265 records per minute on Windows

Accenture Singapore

- Design and run encryption and masking jobs on Linux servers
- Secure PHI for the Ministry of Health Holdings (MOHH)'s Oracle DB
- Row sequencing and job audits

Medicx Media Solutions USA

- Encryption and hashing functions to PII and PHI in geo-medical consumer health databases
- Exceeds HIPAA requirements in provisioning mScores[™] data to digital and direct marketers



Key Differentiators

Developer Support

- Version controls
- Master data definition
- Secure key management
- Project management (teaming)
- SDK supports .NET and Java calls
- Source data and metadata discovery
- XML job audit logs (for verifying privacy law compliance)

Price Performance

- The data-centric security tool with:
 - → The most sources
 - → The most protection functions
 - → The most target file formats
- Fastest standalone protection software

One-Stop-Shop

- Extract, Transform, and Load (ETL)
- Data and DB migration
- Test data generation
- Advanced BI
- Reporting

Ease-of-Use

- Familiar Eclipse GUI
- Self-documenting 4GL syntax
- Easy management and modification of jobs/metadata



Competitive Advantages

vs. IBM

- FieldShield scripts simpler than Optim interoperability model and Javascript options
- Seamless integration with more sources
- More functions
- Lower cost

vs. CA (Grid-Tools)

- FieldShield's CoSort engine is faster than
 Grid-Tool's Fast Data Masking
- Tight integration with data profiling, ETL,
 data quality, and BI operations
- Multi-target/format options
- Lower cost

vs. Informatica

- FieldShield DDM inclusive with product (compared to Informatica's upgrade)
- More protection functions
- Integration with Eclipse and Excel
- Access to 4GL scripts
- Lower cost

vs. GRT

- FieldShield has more masking and encryption functions
- Hash, decode, and pseudonymize functions
- Faster and more extensible in the IRI
 Workbench IDE
- Lower cost







IRI Data Protector Suite

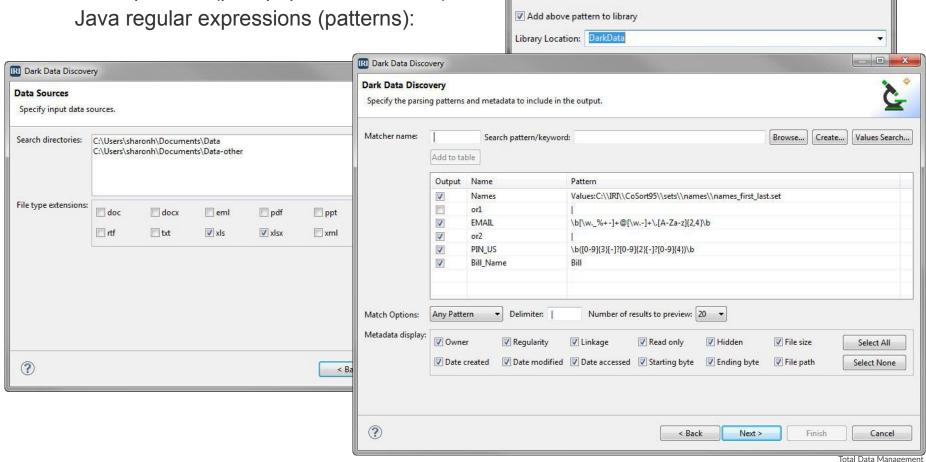
What CellShield Does

- Discover, report, and mask PII and perform audit actions in Excel 2010 & 2013
- Search & secure spreadsheets (and other dark data) throughout a LAN
- Provide common and allow new search pattern definitions for PII formats
- Search for strings in a dictionary, and find/fix PII floating in cells
- Support reuse and sharing of patterns in project or cloud repositories
- Generate a report of all patterns found and open it for action in a worksheet
- Open applicable worksheets and highlight the located ranges for protection
- Click to encrypt, mask, or pseudonymize with supplied functions and options
- Reveal data with the decryption key, or if reversible pseudonym was used
- Overlay results directly into the affected cells or in another worksheet
- Move between or bulk-remediate all identified worksheets and ranges
- Auto-insert protection details into an un-editable audit column in the report



CellShield PII Discovery

The dark data profiling wizard in the IRI Workbench searches network-wide for sensitive data in spreadsheets based on user-specified (plus popular and saved) Java regular expressions (patterns):



New Pattern

Name:

Pattern:

Description:

Create a regular expression pattern

EMAIL

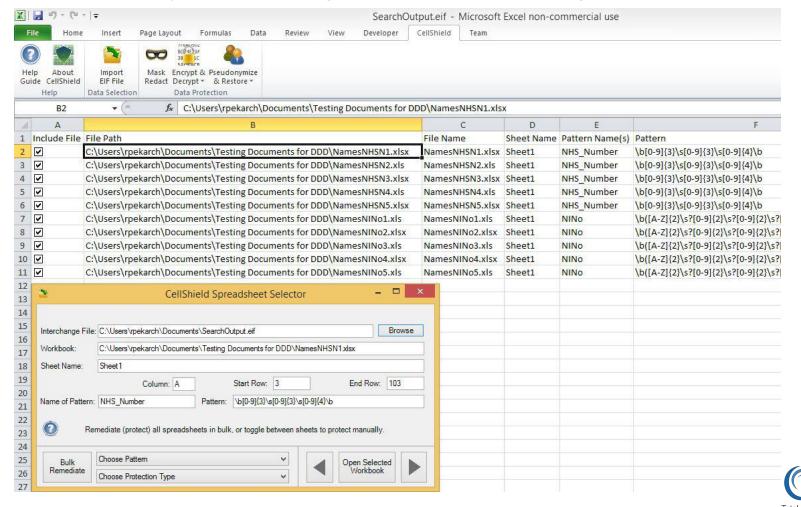
Email

 $b[\w._%+-]+@[\w.-]+\.[A-Za-z]{2,4}\b$

Browse Common Patterns...

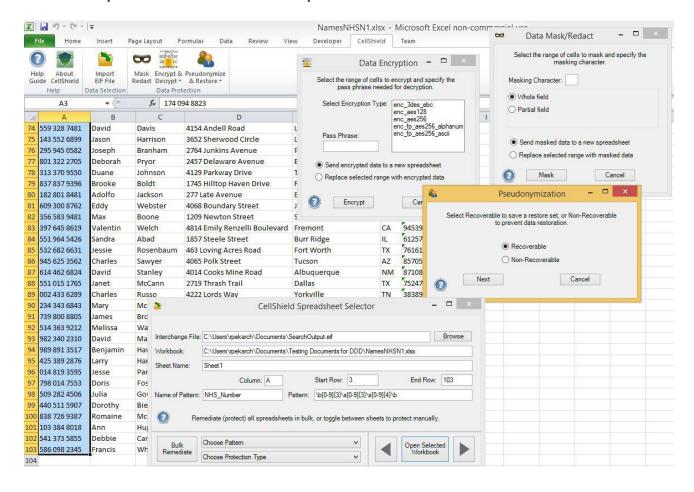
CellShield Reporting

The report produced by the profiling wizard opens in a dynamic worksheet supported by an action dialog for protection and auditing activities



CellShield Protection

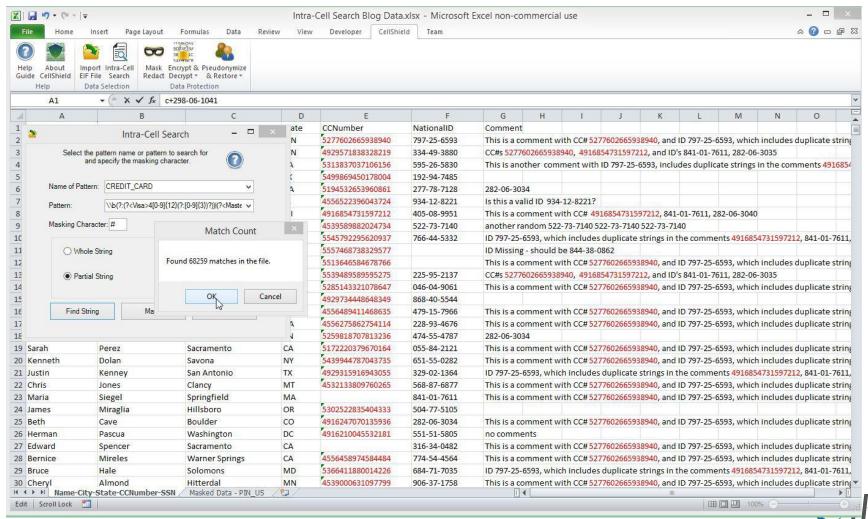
Perform point-and-click encryption and decryption, masking (full or partial cell), or pseudonymization (reversible and nonreversible) of the applicable ranges within the spreadsheets in the report:





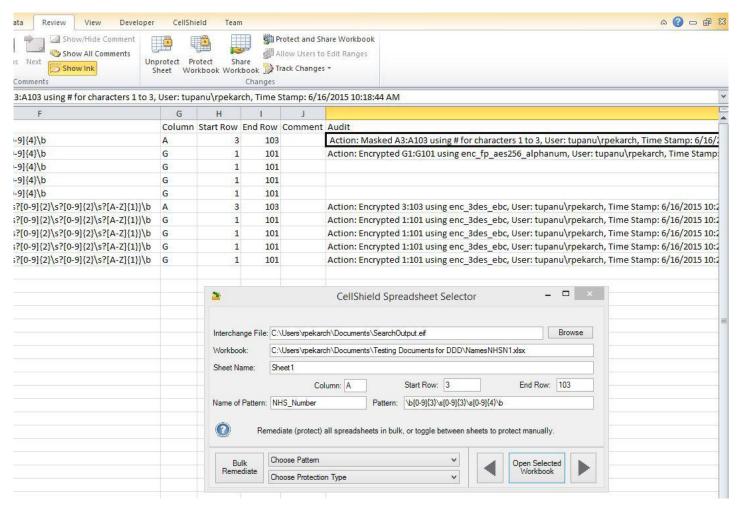
CellShield Intra-Cell Search & Mask

Feature finds and fixes floating PII, ad hoc, or en masse



CellShield Auditing

An uneditable log entry for the protection applied to each pattern identified in the report is automatically appended on each action:









IRI Data Protector Suite

What RowGen Does

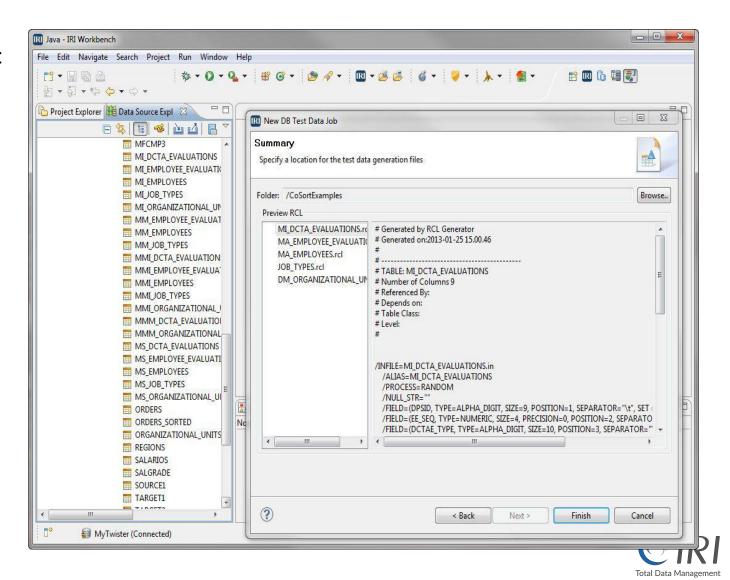
- Create synthetic but realistic random and random-real test data simultaneously
- Improve DB prototypes, application quality, benchmarking, and outsourced operations
- Use standard DB DDL, production file, and custom metadata to define layouts
- Preserve structural and referential integrity of real EDW DBs for testing
- Produce data in any types, structure, volumes, value ranges, and if condition
- Synthesize composite data values and custom (master) data formats
- Generate computationally valid and invalid NID (Codice Fiscale, etc.) SSNs, and CCNs
- Set and graph test data value distributions (linear, normal, random, etc.)
- Apply common attribute rules (like lookups) rules for pattern-matched field names
- Filter, transform, and pre-sort test data while it's being generated
- Write loader metadata and perform direct path loads for test DB populations
- Build test flat-file and custom/structured detail and summary report targets
- Subset and mask databases automatically for test purposes
- Provide SDK functions for generating test data in Java apps and Hadoop



Use Existing Data Models and Metadata

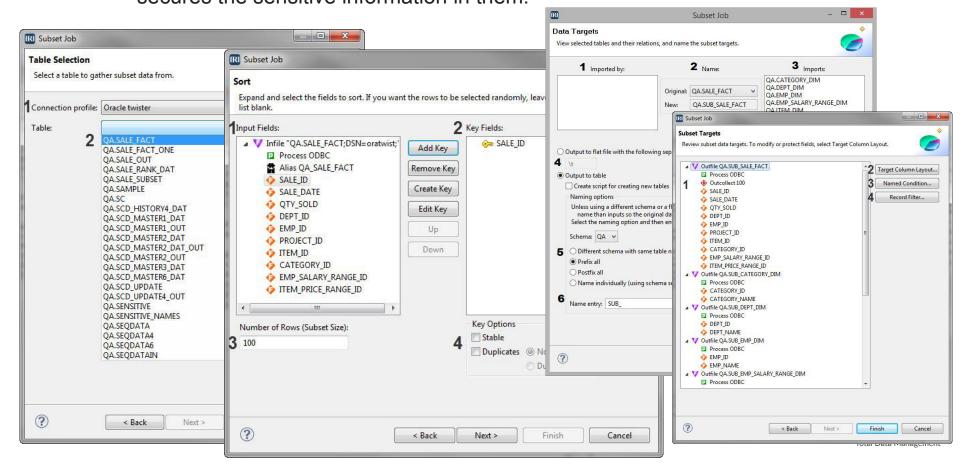
Build Test Data for:

- CoSort
- DataStage
- o DB2
- Hadoop
- Informatica
- NoSQL DBs
- Oracle
- SQL Server
- Sybase
- Teradata
- CSV
- XML
- LDIF
- COBOL



DB Subsetting with the Masking Option

RowGen's subsetting and test data generation wizards facilitate DB and EDW prototyping. Smaller, referentially-correct copies of larger table extracts ensure production data is safe and test data is realistic. Masking secures the sensitive information in them.



User Profiles

Anyone doing DB testing, app development, stress-testing, or benchmarking, including:

- Developers (programmers)
- DBAs and DW (ETL) architects
- Analysts and consultants

Use Cases

Bank of Montreal

Generates safe, realistic 20GB Oracles tables with RI for query testing

MasterCard Peru

Synthesizes PAN and PII in files to support OLTP and app testing

Transitive UK

Simultaneously creates and transforms data to test cross-OS virtualization



Key Differentiators

- Big data generation and population performance
 (embedded CoSort pre-sorting engine speeds bulk loads)
- 2. Synthetic data that's broader and safer than real data
- Concurrent test data manipulation and reporting
- Familiar Eclipse IDE and simple, portable, modifiable test data generation and auto-built DB loader scripts
- Metadata compatibility with IRI DM
 (and third-party ETL, BI, etc. platforms via AMM/MIMB)







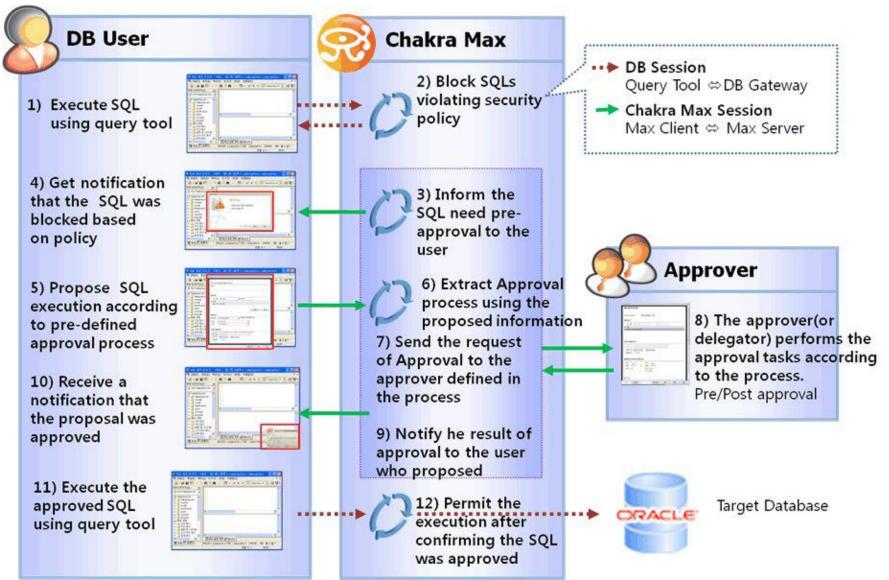
IRI Data Protector Suite

What Chakra Max Does

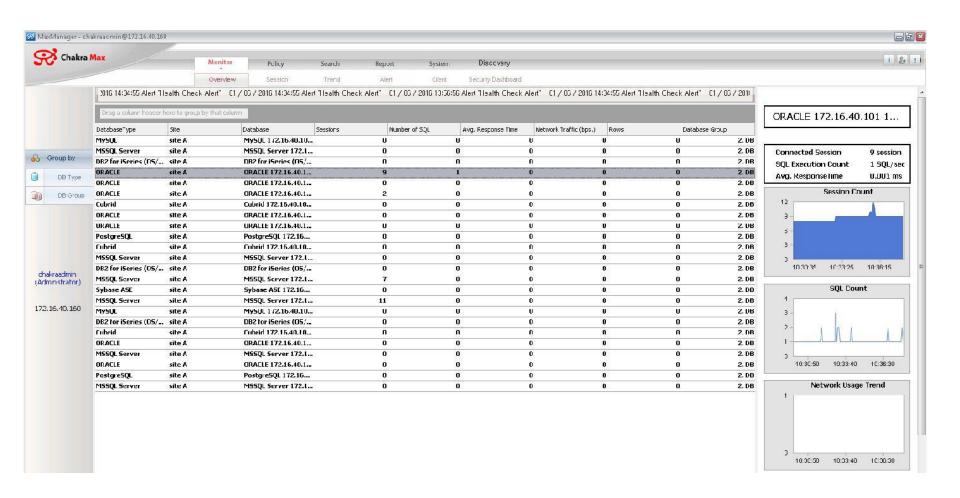
- High-volume, data-centric audit and protection (DCAP) for 20 different databases
- Monitor, block, alert, and log users and traffic in real-time without impacting
 DB performance or availability
- Control DB access and permissible SQL executions by user and rule
- Mask data dynamically with full or partial field value redaction
- Apply policies across multiple DB instances at once
- PII access logging
- ISO 27001-compliant activity reporting and log analysis
- Uses a 3-tier Web Application Server (WAS) agent to collect packets between the client and DB, and it monitors and controls them



Chakra Max Overview

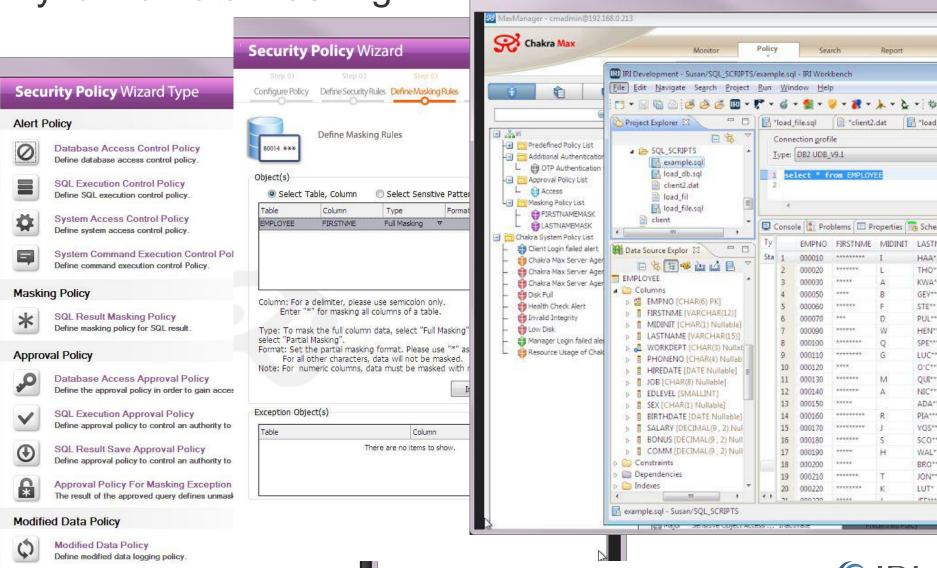


Real-Time Monitoring & Auditing





Dynamic Data Masking





LASTI

HAA*

THO*

KWA*

GEY*

STET

PUL**

HEN*

SPE

LUC**

QUI**

NIC**

ADA*

YOS"

SCO**

WAL*

BROX

JON™

LUT*

User Profiles

- Anyone with databases containing PAN, PII, PHI, or other sensitive information
- Companies needing a standard multi-source DCAP solution
- Businesses requiring central and differential control over DB access
- Anyone needing real-time DB monitoring and alerting

Use Cases

Hankook Performance Tires

Chose Chakra Max's for scalability and low-impact on performance

LG U+

Secures access to PII in customer DBs across mobile services network

University of Maryland Medical Center

 Uses DDM and secure audit log facilities for PHI protection and compliance verification



Key Differentiators

Performance

- Most stable and best-performing
 DCAP solution for high-traffic
 volume environments
- Firewall thousands of DBs at once
- 100k SQL/second monitor speed
- 10k-25k SQL/second audit speed
- Low-impact

Interoperability

- Support for 20 different DBs on Linux,
 Unix, and Windows platforms
- Extensible with IRI Voracity
 - Discount for bundled purchase
 - ➡ Enables all-in-one data discovery, integration, migration, governance, and analytics

Capability

- User-level access and execution privileges
- DDM without changing original DB content
- Monitors all login attempts, transaction commands, and results sets
- Real-time logging
- Automated reports
- 40 formats for printing
- 14 formats for saving

Compliance

- Internal and data privacy law compliance
- Powerful and flexible auditing facilities for both real-time alerts and logging and forensic investigation of the logs
- Complete audit history
- Automatic backups
- Post-deletion restore capabilities





Learn and Share



IRI.com

IRI blog

LinkedIn Data Masking & Protection Group LinkedIn Test Data Management Group













