

- Prototype DB and ETL Ops
- Improve Application Quality
- Comply with Data Privacy Laws
- Benchmark Hardware and Software

Test Data Generation and Management



What Does RowGen Do?

IRI RowGen randomly generates or selects data to synthesize realistic data for:

- Structurally and referentially correct RDB prototypes
- Custom detail and summary file or report targets
- DW, ETL/ELT operations, Data Vaults
- Custom applications and procedures

RowGen reads the metadata of production databases, applications, and file layouts to guarantee realistic data structures, values, and primary/foreign key relationships. While producing the data, RowGen can also transform and format it. RowGen uses a powerful Eclipse GUI and open, self-documenting job scripts that support the re-use and sharing of test data layouts and applications.

What are the Benefits of RowGen?

RowGen reduces the risk, effort, guesswork, and costs involved in complex test data generation. Designed with developers, DBAs, ETL architects, and compliance teams in mind, RowGen provides an easier way to generate test data sets that:

- Are safe to outsource
- Comply with privacy laws
- Preserve referential integrity
- Conform to your business rules
- Do not tax database resources
- Do not require complex programs
- Cut development and testing time/costs
- Undergo detail and summary transforms
- Scale across an enterprise data warehouse
- Leverage existing data models and file layouts
- Improve quality control via data realism and ranging
- Populate multiple targets simultaneously, including reports

Because RowGen job scripts are based on, and compatible with, CoSort SortCL program metadata, they use the same data manipulation and formatting parameters to customize the look and feel of the test targets. So naturally, RowGen jobs that generate test data can easily switch to CoSort for manipulating real data in the same format once testing shifts to production. RowGen also uses the same IRI Workbench GUI, built on Eclipse™, which includes multiple data discovery tools.

Selected RowGen Uses

Database Population

Rapidly generate and load millions of rows of ordered test data into multiple tables at once, while adhering to their DDL-defined structures and relationships. Populating tables with pre-sorted data also speeds very large load, unload, and query prototypes.

Application Development

Synthesize, transform, and format data in the GUI to meet the volume, range, and layout requirements of any program or prototype. Or, use the Java software development kit (SDK) to dynamically generate test data in Hadoop. Simulate, stress test, and be less dependent on other developers or confidential data.

Format Sharing

Build test targets ranging from simple, single-field flat files to custom-formatted, detail and summary reports. Outsource development and demonstrate programs with realistic data while keeping private data safe. Use computationally valid generators for: US and Korean Social Security numbers, Italy fiscal codes, Netherlands BSN, Singapore NRIC, and credit cards.

Benchmarking

Leverage fast I/O techniques in RowGen to create massive test tables and files at once. Specify any desired layouts, value ranges, and distributions. Rapidly build scalable data sets for complete and consistent testing of different software and hardware products or configurations.

RowGen Operations

RowGen rapidly builds multiple test data targets - all in different formats, and all at once. Use RowGen to create test data for Oracle, Microsoft SQL Server, Teradata, IBM DB2 UDB, Sybase, Informatica, DataStage, and other platforms. Quickly populate test tables and files for an entire enterprise data warehouse. Write RowGen Control Language (RCL) job scripts by hand, or create them automatically in the IRI Workbench GUI for RowGen.

RowGen in the IRI Workbench

RowGen users get a free Eclipse GUI to discover and classify data in multiple sources, and then create, run, and manage their test data generation jobs. The GUI contains several job wizards and functional dialogs, plus a syntax-aware script editor and dynamic outline for RCL jobs and metadata. For test database creation, RowGen has a fit-for-purpose job wizard that automates:

- *Parsing* - RowGen selects the schema and tables you want to populate. RowGen translates the database table descriptions and integrity constraints into RCL scripts that specify the source structure, dependent sets, and data creation.
- *Generation* - RowGen builds and runs the RCL scripts to create one test file per table that can be bulk loaded, and/or saved for future use and modification.
- *Population* - RowGen bulk loads the target tables in the order necessary to maintain all referential integrity. RowGen also pre-sorts the data in primary key order, which dramatically improves bulk load operations.

Another RowGen wizard creates ad hoc test data targets in standalone flat-file, report, or table formats. RowGen also supports: all and valid (joined) pairs, composite data types and values, standard and weighted distributions, and the inclusion of randomized real data to enhance the appearance of test data without compromising security.

Field Name: NewField

Set file
Set: PERMUTE "C:/IRI/sets/people.set" DEFAULT = "en" Define...

Formatting Attributes
Data Type: ALPHA_DIGIT
Position: 1
Size: 0
Separator: %ASCII
Precision: 0
Alignment: NONE_ALIGN

Frame:
Xdef:
Mill:
Fill:
Null Low:
Null High:
Endian:
Row Id: Set Row ID

Initial value Step value Max value
4 2 4 Define...

Null Percent: 3
Random Min Size: 25
Random Max Size: 150

OK Cancel

Options
Parameters for generating test data

Rows to generate default: 100

Table selected	Use default	Rows to generate
BONUS	<input checked="" type="checkbox"/> true	Default
DEPT	<input checked="" type="checkbox"/> true	Default
DM_EMPLOYEES	<input checked="" type="checkbox"/> true	Default

Delimiter: t
Sequencer:

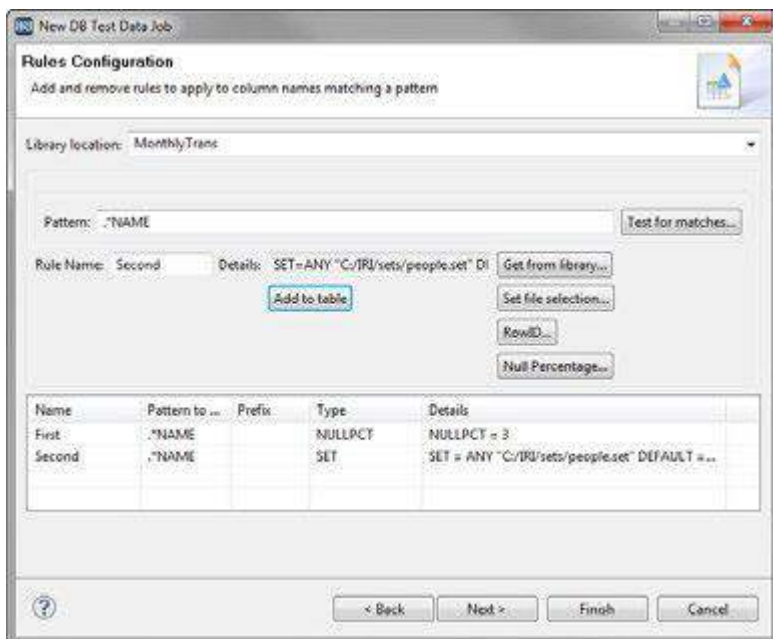
Create loader control files for loading test data into the database

Database type: ORACLE

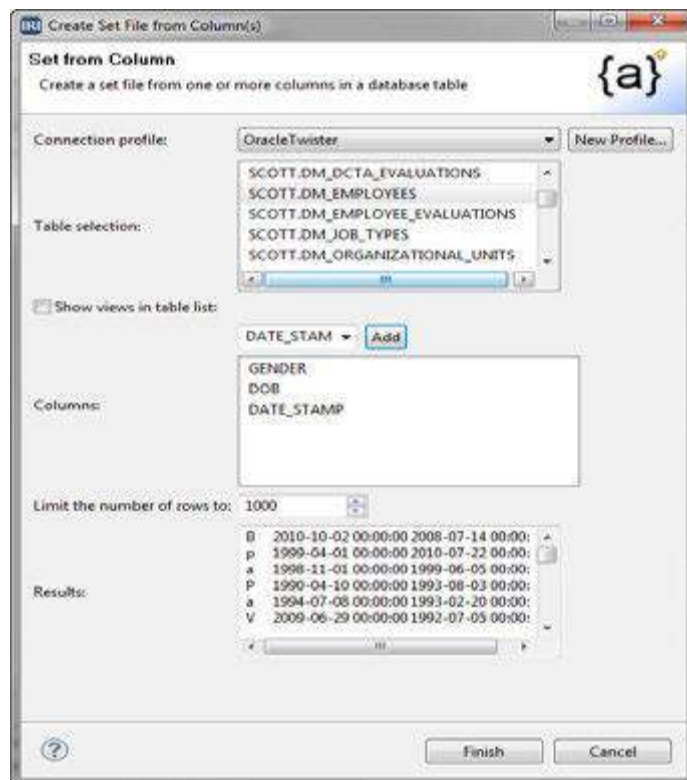
< Back Next > Finish Cancel

The screenshots left and above illustrate some of the ways RowGen users in the IRI Workbench can specify and modify test data generation. The easy-to-use test DB wizard automates RowGen script creation by directly reading DDL information from any JDBC-connected database.

RowGen also allows you to modify the parameters of generated data using a point-and-click interface, or through a syntax-aware job script editor.



Applying a column selection rule in the IRI Workbench GUI for RowGen.



Test-Related Functions

- Random Data Generation and Selection
- Parallel Sorting (speeds loads & queries)
- Conditional Filtering (for value ranging)
- Multi-Targeting (esp. for table relations)
- Calculations, Aggregations & Sequencing
- Frequency Distribution
- Custom Flat-File & Report Formatting
- Logging (in XML format, for compliance)

Test File Formats

- ACUCOBOL-GT Vision
- Fixed- or Variable-Position Flat Text
- Line, Record, or Variable Sequential
- LDAP Interchange Format (LDIF)
- Micro Focus Variable Length & ISAM
- Mainframe Variable Blocked Format
- MS Comma-Separated Values (CSV)
- W3C Extended (Web) Logs
- W3C XML

Test Data Types

- ASCII & EBCDIC Characters
- Numeric, Whole, Currency, IP Address
- Alpha & EBCDIC Digits
- RM and MF COBOL Numerics
- Other Binary Numerics
- US, European, ISO, Japanese Timestamps
- Unicode & Multi-byte Characters

Supported Metadata

- 3NF DDL & compliant data vaults
- COBOL copybooks
- CoSort (SortCL) data definition files (.ddf)
- LDIF & flat XML files
- Microsoft CSV & W3C ELF headers
- Oracle SQL*Loader control files
- DTD/XSD, plus BI/ETL tools via MIMB

Supported Platforms

- AIX (IBM i, p & z Series)
- HP-UX (RISC and Itanium)
- Linux (x86, Itanium, IBM i, p & z Series)
- Solaris (SPARC and Intel)
- Other UNIX (e.g. BSD, Mac OS X, Tru64)
- Windows (XP, 2000-2016, 7, 8, 10)

Compatible Products

- AnalytiX DS - Mapping Manager
- Data Vault - Data Modeling
- IRI CoSort - Data Integration & Transformation
- IRI FACT - Fast Extract for Oracle, DB2, et al.
- IRI FieldShield - Data Masking
- IRI NextForm - Data and Database Conversion
- IRI Voracity - Total Data Management
- MIMB - Meta Integration Model Bridge



© 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. (CoSort Korea). Other product, brand, or company names may be (registered) trademarks of their respective holders.