



# IRI Voracity

An Insatiable Appetite for Data

- Big Data Integration & Analytics
- File, DB, DV & ETL Tool Migrations
- Data Profiling, Cleansing & Masking
- Powered by CoSort, Hadoop & Eclipse

## An Insatiable Appetite for Data



## Product Summary

# What is Voracity?

IRI Voracity® is a modern software platform for:

- data discovery
- data governance
- data integration
- data analytics
- data migration
- data curation

Voracity is powered by the proven performance of IRI CoSort or interchangeable Hadoop engines.

Voracity jobs are designed in the rich, familiar, and extensible IRI Workbench IDE, built on Eclipse™.

*Voracity is a robust, multi-use-case data management and manipulation solution stack for enterprise:*

## DW & BI Architects

to design and run big ETL or reporting jobs, and speed or leave old ETL tools.

## Big Data Scientists

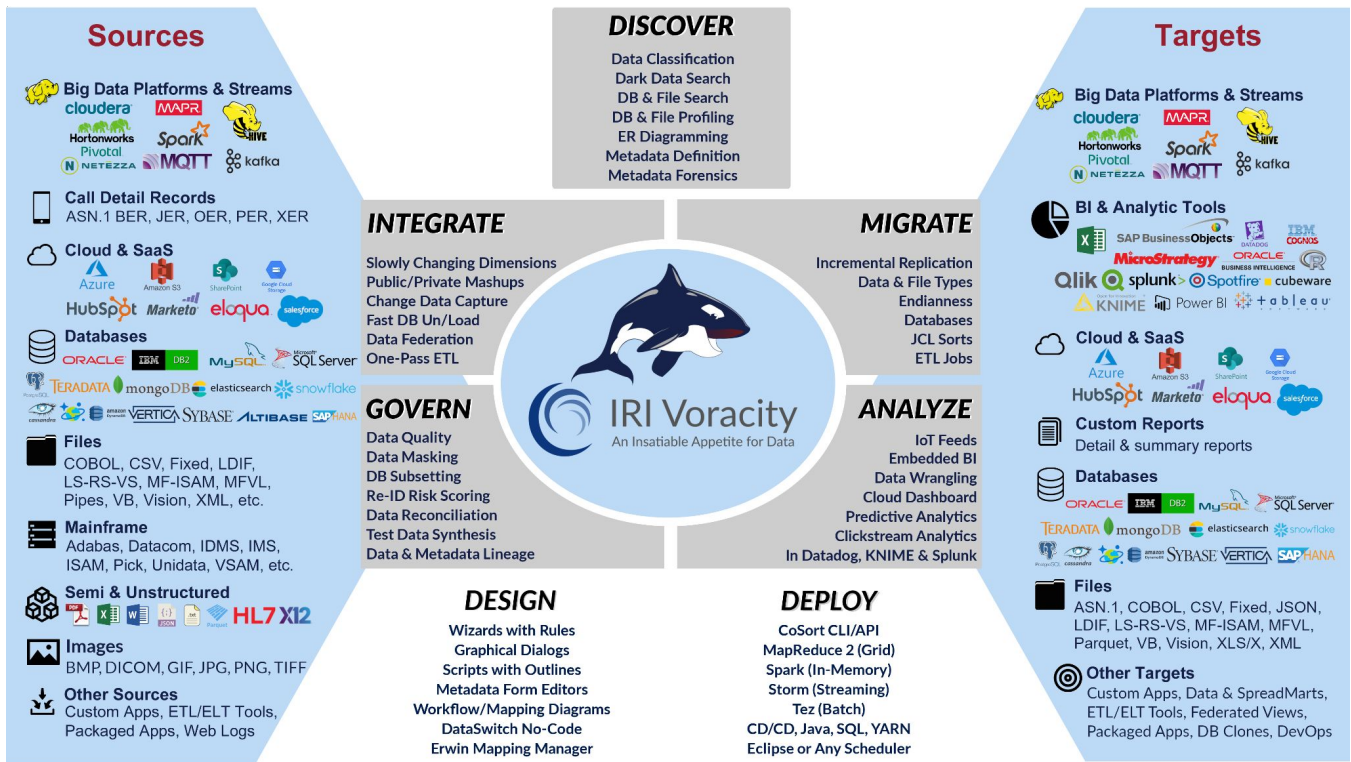
to find, define, and wrangle raw data for experimentation and analytics.

## Database Administrators

to speed unloads, loads, and reorgs, migrate and replicate, or report on changed data.

## Data Stewards

To profile, cleanse, unify, risk-score, PII-mask, track, and prototype data.



## Why Use Voracity?

**Speed**

Voracity overcomes the performance challenges of turning big data into actionable, governed information without taxing databases or budgets. Powered by CoSort or Hadoop engines. Voracity moves, munges, and masks big structured, semi- and unstructured data sources on-premise or in the cloud. Advanced transformation algorithms and multi-task consolidation give Voracity a 10-to-1 performance edge over SQL, BI, and legacy ETL tools.



**Savings**

Voracity bends the cost, complexity, and risk curves away from:

- legacy megavendor stacks
- disjointed Apache projects
- separate data federation, quality, masking, and wrangling platforms

Voracity's one UI supports the outcomes that CIOs need and the collaborations that IT and business users want. CFOs enjoy Voracity's affordable subscription model that doesn't count cores, users, jobs, sources, or rows.



**Simplicity**

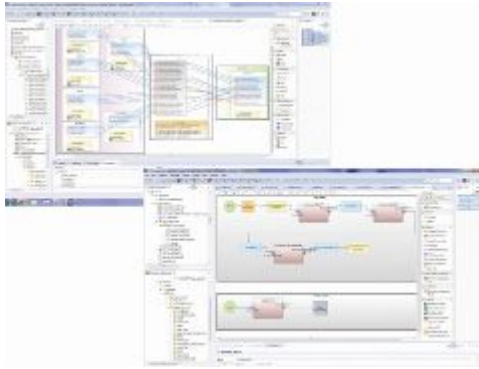
Voracity job control in Eclipse is an ergonomic and intuitive process, as multiple options for both design and deployment reduce learning curves and support functional expansion. Shareable, explicit metadata also eases onboarding and support for new users and new forms of data. API integrations with DataSwitch, and support for erwin and MIMB metadata exchange hubs, facilitate third-party ETL, BI, SQL, and data modeling tool conversions to Voracity.



## Data Discovery

Many Voracity users begin with data searching and tagging via automated pattern and fuzzy-match scanning, statistical reporting, and metadata creation. Voracity features user-friendly wizards for:

- Database profiling, classification, and ER diagramming
- Dark (unstructured) data searching and extraction
- Flat-file statistical reporting, searching, and classification
- Metadata definition, rule building, and team sharing



## Data Integration

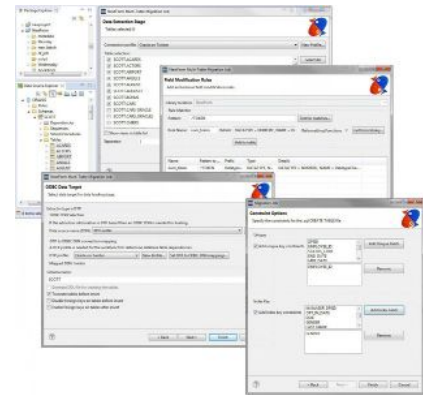
Voracity rapidly moves and manipulates disparate data sources while using simple, open metadata that you can modify in multiple ways. It accelerates or replaces legacy ETL tools with superior technology for:

- Big data warehouses or data lakes
- Data cleansing and standardization
- VLDB unloads, external transforms, reorgs, or loads
- Change data capture and slowly changing dimensions

## Data Migration

Voracity offers multiple data migration, replication, and re-formatting options to convert and re-use data stored in legacy and SQL DBs, index and sequential files. Voracity can also search and mask semi- and unstructured data in DBs, documents, images, logs, cloud apps, etc.

- Remap data types, record layouts, file formats, and schema (e.g. to Data Vault)
- Move column values, layouts, and relationships (constraints) between databases
- Replicate data from one or more sources to one or more table, file, or report targets
- Federate or virtualize data by producing ad hoc views from disparate sources



## Data Governance

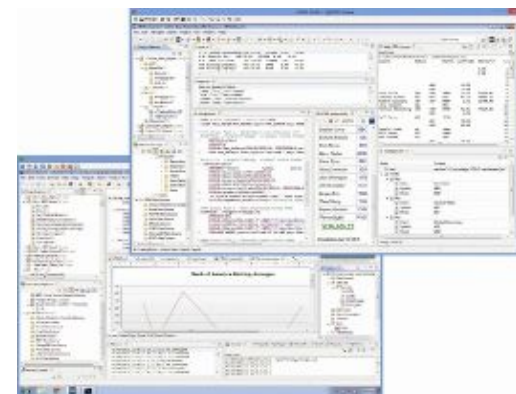
Voracity supports data governance and information stewardship, and integrates these activities into its other data management functions:

- Master data management - reconcile and unify table and file values
- Enterprise metadata management - track, share, and secure metadata
- Static and dynamic data masking - comply with data privacy laws
- Test data creation - for DevOps, DW/BI prototypes, benchmarks, demos, etc.

## Analytics

In the same Eclipse pane of glass, IRI Workbench, Voracity can prepare, present, and provide comprehensive, custom insights from public and private data for descriptive, diagnostic, predictive, or prescriptive analytics via:

- Embedded BI - custom 2D detail and summary reports with calc/stat functions
- BIRT & KNIME integrations - feed results and metadata in memory, in Eclipse
- Data Wrangling - send results to tools like R or Tableau to speed visual insight
- Cloud Dashboards - design custom reports in Splunk, Cubeware, Datadog, etc.



# Data Life Cycle Management

*Rapid, comprehensive enterprise data curation. Get it done with Voracity:*

## Profile & Acquire

Discover, classify, and extract data and metadata as you connect to disparate data sources. Define custom data structures, recast date formats, and build test data.

## Cleanse & Unify

Filter, enrich, scrub, and standardize data in multiple sources. Select, fuzzy-search, and merge reference data into master tables and values.

## Protect & Audit

De-identify data you classify as you filter, transform, report, or hand it off. Encode, encrypt, hash, pseudonymize, delete, redact, tokenize, and log it.

## Process & Provide

Integrate, migrate, govern, and analyze data in the same job and I/O pass using CoSort or Hadoop. Visualize and feed real, or test, data targets in any structured format and silo.

## Express & Predict

Aggregate, cross-calculate, and format data in detail, summary, and trend reports, or hand-off results to your analytic tool or dashboard display.

## Convert & Replicate

Migrate legacy databases, files, and data types, or specify new target record layouts. Make copies or subsets of data in any format or schema.

## Virtualize & Share

Federate, save, or populate multiple targets at once, and connect to them and their metadata in secure repositories like EGit for change tracking and version control.

## Interactive Job Design Features

- Fit-for-purpose job creation wizards
- Help-enabled graphical dialogs and form editors
- Syntax-aware script editors with dynamic outlines
- State-of-the-art visual workflow and mapping diagrams
- erwin Mapping Manager compatibility for 'code-free ETL'

## Automatic ETL Tool Migration

Thanks to metadata integrations with erwin and MIMB, legacy BI/DW and SQL architects can automatically migrate many of their data mapping and reporting jobs to cheaper, faster, and/or simpler Voracity workflows. Hadoop users can run many Voracity tasks in MR2, Spark, etc. without recoding. And through its open workflow API, your apps can read or feed Voracity data.

## About IRI

IRI is a renowned US data management and protection ISV. Founded in 1978, the company has a long history of technical innovation and thought leadership, particularly in high speed data transformation and data-centric security.

IRI expertise across the data life cycle, and the superior price-performance of IRI software, continue to distinguish the company in the IT industry, and maintain the trust of data-driven EIM architects and expert partners worldwide.



© 2022 Innovative Routines International (IRI), Inc. All Rights Reserved. CoSort, Voracity, NextForm, FieldShield, CellShield, DarkShield, and RowGen are trademarks or 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.

2194 Highway A1A  
Melbourne, FL 32937 USA  
1.321.777.8889 \* 1.800.333.SORT



[iri.com/voracity](http://iri.com/voracity)