



- Data Integration & Analytics
- Data & ETL Tool Migration
- Data Discovery & Governance
- Built on CoSort®, Hadoop™ & Eclipse™

## An Insatiable Appetite for Data



**Product Summary**

**IRI**

The CoSort Company

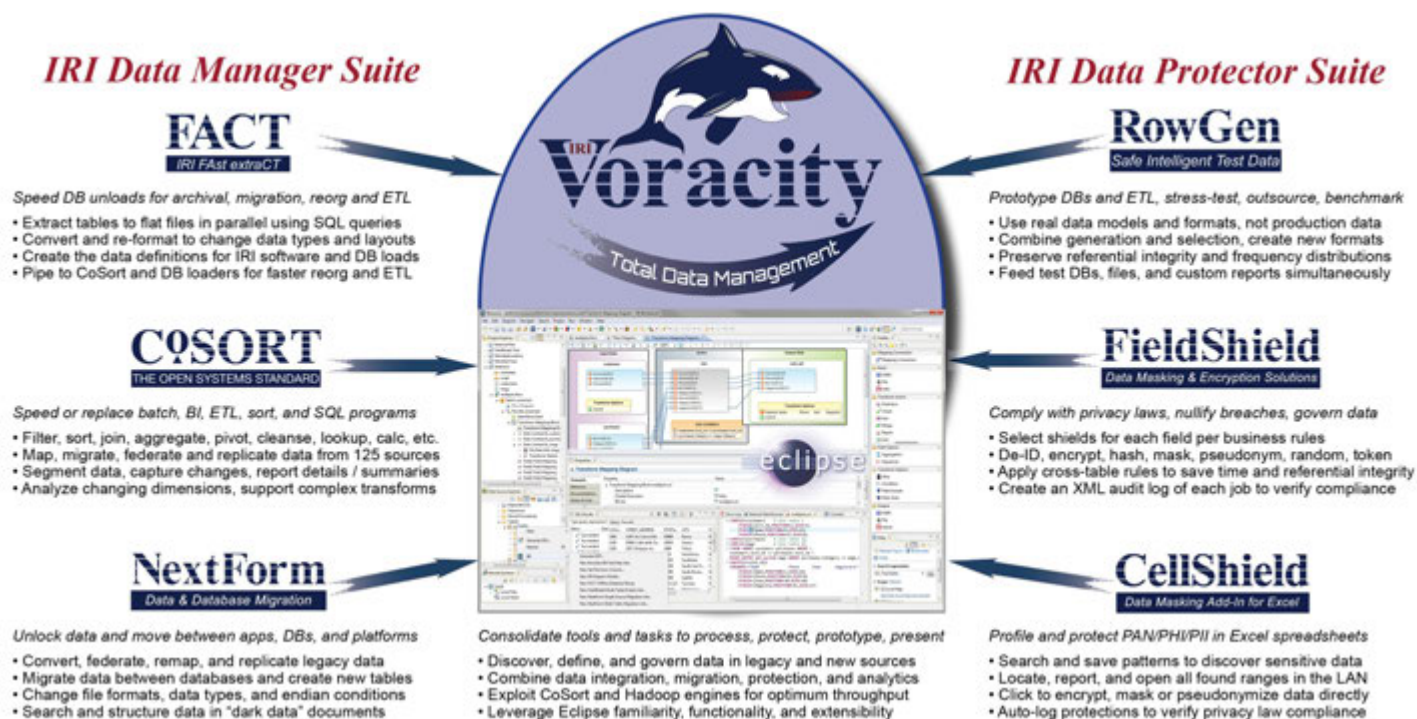
# What is Voracity?

IRI Voracity® is a robust enterprise data management and manipulation platform built

for: who need to:

DW Architects	design, preview, and run ETL jobs, or speed and migrate other tools
Data Scientists	discover, define, and prepare disparate data for analytics
Data Stewards	cleanse, unify, mask, test, trace, and share data
DBAs	speed unloads, loads, queries, and reorgs, or report on changed data
CIOs	automate migration from legacy ETL tools, and consolidate vendors

Voracity performs, combines and speeds: data discovery, integration, migration, governance, analytics, and curation ... managing the enterprise data lifecycle. Voracity is powered by the proven performance of the IRI CoSort engine, or multiple Hadoop engines. Voracity is front-ended in the IRI Workbench GUI, a single Eclipse plug-in with total extensibility.



## Why Voracity?

### Speed

Voracity addresses the volume, variety, velocity, and veracity challenges in turning big data into actionable information. It leverages the proven performance of CoSort -- or seamless Hadoop MapReduce, Spark, Storm, and Tez options -- for moving, manipulating, masking and munging structured, semi-structured, and unstructured data, sourced on premise or in the cloud. CoSort's superior algorithms and task combinations run jobs 6x faster than megavendor ETL tools.

### Savings

Voracity bends the cost, complexity, and risk curves away from legacy ETL packages, disjointed Apache projects, and separate BI, data federation, quality, masking, and MDM platforms. By designing and running everything in one "pane of glass" -- and sometimes in one I/O pass -- Voracity delivers the outcomes that business needs and the consolidation that IT wants. And CFOs prefer its affordable subscription model that doesn't count cores, users, jobs, or rows.

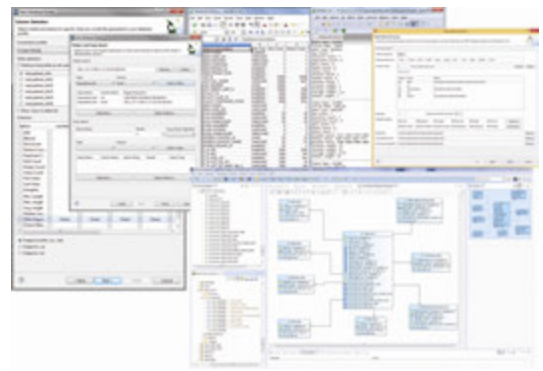
### Simplicity

The intuitiveness of, and ergonomic choices for, Voracity job design and deployment in Eclipse -- plus traceable, self-documenting metadata -- eases onboarding and maintenance. AnalytiX DS compatibility automates conversion of legacy ETL jobs. BIRT integration, embedded BI, data franchising, and self-service, shard-powered big data analytics (pending) mean IT and BI users can speed digital business insights collaboratively within a proven, managed data infrastructure.

## Data Discovery

Many Voracity journeys begin with data identification and classification via pattern and fuzzy-logic searches, statistical reporting, and metadata creation, in user-friendly wizards for:

- Database profiling and ER diagramming
- Dark data searching on extraction
- Flat-file statistical reporting, searching, and classification
- Structured data metadata discovery and definition



## Data Integration

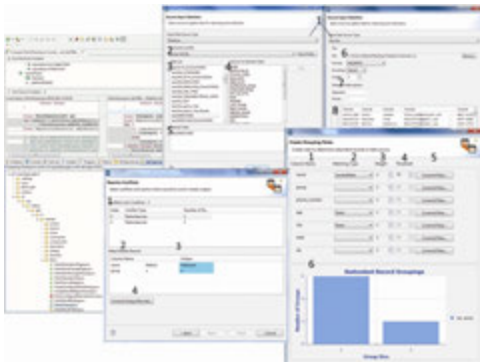
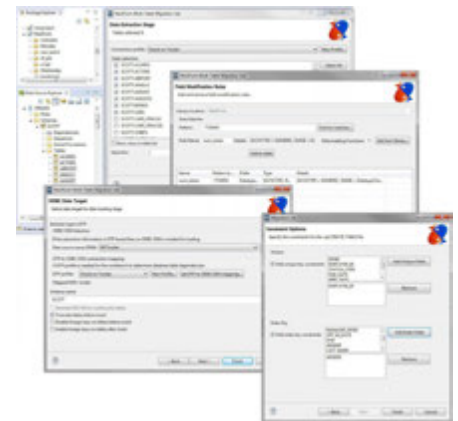
Voracity moves and manipulates disparate data rapidly. It uses simple metadata that all IRI products and AnalytiX DS share. It accelerates or replaces legacy ETL tools with superior technology for:

- Big data warehouse ETL or ELT operations
- Data quality and standardization
- VLDB unloads, external transforms, reorgs, or loads
- Slowly changing dimensions and change data capture

## Data Migration

Voracity offers data migration and reformatting options to convert and re-use data stored in legacy and modern databases, index and sequential files, and dark data documents:

- Remap data types, record layouts, file formats, and endianness
- 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, projects, or folds those activities into its other data management activities. Included capabilities are:

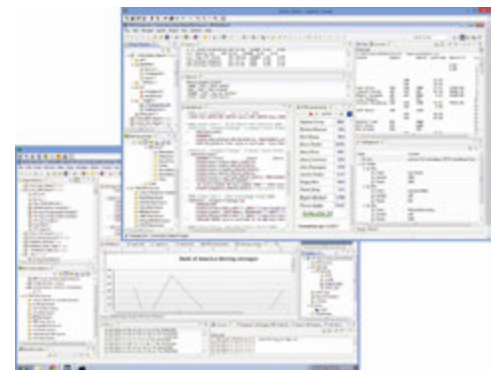
- Master data management - unify and bucket table and file values
- Enterprise metadata management - track, share, and secure metadata assets
- Static and dynamic data masking - comply with data privacy laws
- Test data synthesis - prototype DB/ETL operations and stress-test applications

## Analytics

In the same Eclipse IDE, Voracity prepares, presents, and provides comprehensive, custom insights from public and private data for descriptive, diagnostic, predictive or prescriptive analytics:

- Embedded BI - custom detail and summary reports with calc/stat functions
- BIRT Integration - feed results and metadata into BIRT at reporting time
- Data Preparation - feed results to R et al. to speed time-to-visualization
- Self-Service BI (pending) - shard-powered, immediate visual analytics

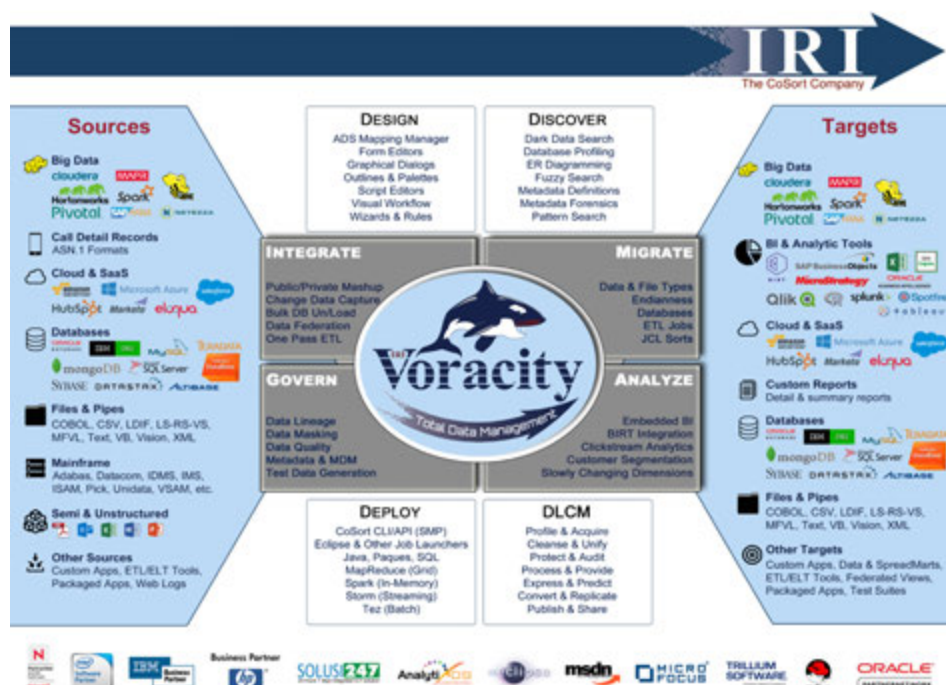
Produce any number of database, file, and report targets -- with real production or realistic test data -- in multiple legacy and modern formats at once. Consolidate more work into the same I/O pass to save time in job design and execution time.



# Data Life Cycle Management

Use these essential curation capabilities within Voracity:

- **Profile & Acquire** Discover, classify, and extract data and metadata as you connect to disparate data sources. Define custom data structures, mask 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 at the field level as you filter, transform, report, or hand-it off. Encode, encrypt, hash, pseudonymize, redact, tokenize, etc. and log in XML.
- **Process & Provide** Integrate, migrate, govern, and analyze data in the same job and I/O pass in CoSort or Hadoop. Visualize and feed test or real data targets in any format.
- **Express & Predict** Aggregate, cross-calculate, and format data in detail, summary and trend reports, or, hand-off results to your analytic tool or BIRT visualizations.
- **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.
- **Publish & Share** Federate, save, or populate multiple targets at once, and connect to them and their metadata in secure repositories like GIT for change tracking and version control.



## Interactive Job Design Features

Use any combination of these Voracity 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
- AnalytixDS Mapping Manager for spreadsheet-style, 'code-free ETL'

## Automatic ETL Tool Migration

Through IRI's alliance with Analytix DS, legacy ETL tool and SQL architects can auto-migrate their jobs to cheaper, faster, and simpler Voracity workflows. Hadoop users can leverage Voracity capabilities in whole or part. And through IRI's "Gulfstream" API for metadata integration, ISVs and solution developers can consume or build Voracity workflows seamlessly.

Copyright © 2016, Innovative Routines International (IRI), Inc. All Rights Reserved. CoSort, FieldShield, NextForm, RowGen, and Voracity are registered trademarks of IRI, Inc. FACT is a trademark of DataStreams Corp. (CoSort Korea). Other product, brand, or company names are, or may be, (registered) trademarks of their respective holders.