IRI (Voracity) and Data Fabric

Andy Hayler ©Bloor January 2024

Innovative Routines International (IRI) was founded in 1978 and is headquartered in Melbourne, Florida. Incidentally, IRI should not be confused with another venerable data-driven company called IRI (Information Resources, Inc) based in Chicago, which just rebranded as Circana after a merger.

Innovative Routines International is the inventor of CoSort, a highly efficient sort package adapted from mainframe computing to migrate sort jobs off the mainframe. CoSort is also used to accelerate other large-scale data transformations, and is sometimes embedded within other (ETL, wrangling, vertical application) products to do just this. IRI is also the developer of Voracity, a data management platform powered by CoSort that offers a wide range of functionality, from data discovery and data integration through to data governance, data masking and analytics.

Voracity could be regarded as the Swiss Army Knife of data management platforms, as it covers such a wide range of useful functionality. IRI has customers in banking, government, healthcare, and other industries including telco, retail and energy. Customers include LexisNexis, Fidelity Investments, Comcast and Rolex. IRI tends to excel in situations where there is a need to deal with complex data transformations at scale, or mask PII in multiple sources.

IRI uses data definition files (DDF) to catalog and map data sources to targets, and to apply and track business rules such as masking fields containing PII or other sensitive data. There is support for various aspects of data governance, and some ability to visualise the data, though the product does not currently have a full knowledge graph. IRI uses the DDFs as part of its robust 4GL program called SortCL for data manipulations that include data conversion, transformation and reporting, as well as data quality functions like profiling, cleansing and enrichment. The product can be deployed either on-premise, in the cloud, or in a hybrid environment. The various functions of the product are driven via a Workbench product with a graphical user interface built on Eclipse.

These capabilities mean that IRI has many of the core capabilities to support a data fabric architecture. There is a data catalog, and extensive capabilities to allow data to be accessed from a wide variety of sources. Indeed, the history of the company means that it actually carries out data transformations extremely efficiently, to the extent that it is often used to accelerate ETL processes in other vendor products. They are currently missing a true knowledge graph, but that is a feature that would appear to be relatively easy to add given the extensive core underlying capabilities of the product that are already in place.

IRI can use machine learning for unstructured data discovery, such as searching through legal depositions for particular elements of interest or creating structure for virtual or static mashups with other sources in a dta fabric. They plan to use AI to enhance their data profiling capabilities and generate SortCL scripts for a variety of new applications.

IRI is a long-established vendor with a proven track record and mature capabilities in its data management platform. It does not spend heavily on marketing and to match the name recognition of larger players in data management, but its impressive roster of customers and well-established use by technology partners demonstrate the maturity of its IP stack. IRI products can play a part within a modern data fabric architecture, and its strengths in dealing with very efficient data transformation and assorted data wrangling make Voracity a platform well worth looking at, especially if you have particularly challenging data volumes or performance demands concerning around production data transformation and migration jobs, or a need to find, mask or synthesise data in development environments.

Further Reading – Daniel Howard articles:
The IRI Platform, _Data Masking_,
Test Data Management, _Test Data Generation_