Cross Data Integrator
XDI

3Insys
Insightful•Innovative•Intelligent Systems

Highlights

- *Delivers a single "Source of Truth" for company data*
- *Enables data driven business decisions*
- *Removes data silos*
- *Enterprise Data Warehouse*

The **Cross Data Integrator (XDI)** from 3Insys integrates data from files, databases, and other storage systems and delivers it in a single view for the company. The unified data can be consumed by Business Intelligence (BI) applications or stored a centralized location for making confident decisions driven by company data.

The data a company generates is a critical asset. As utilities modernize and their ecosystem grows, the amount of data produced can become overly massive. With each system storing data in different formats and locations, having access to all data when critical business decisions need to be made is imperative.

As a complete data integration solution, the **XDI** connects to remote data sources and transforms the data to a common format using defined schemas. Depending on each organization's needs, **XDI** can include an enterprise **Data Warehouse (DW)** as a centralized repository to store and manage the data.
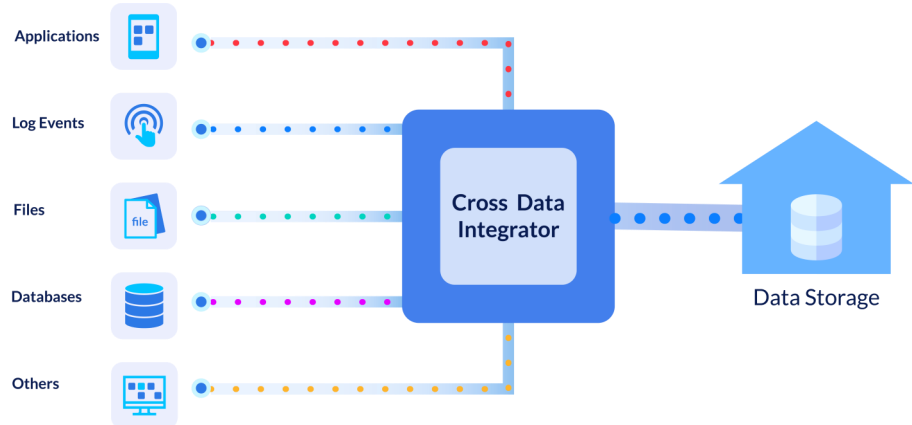
Data integration uses schemas and data cleansing techniques to unify the data and exchanged between data services and data sources.

XDI provides high performance integrated data processing from multiple sources where the data will be cleaned, denormalized, and aggregated by applying subject-oriented business rules. **XDI** supports integrated, subject-oriented, volatile, and non-historical atomic data used specifically in support of transactional processing and timely operational reporting. It provides a foundation for collective, up-to-date information integrated from operational staging tables, and facilitates operational reporting in real-time or near real-time.

3Insys
Insightful•Innovative•Intelligent Systems

The **Cross Data Integration (XDI)** solution employs an elastically scalable modern architecture that can handle large volumes of data, perform complex operations on multiple types of data, and provide detail business and process insights at near real-time. It's based on self-contained REST API based microservices that impart scalability, reusability, resiliency, better fault isolation, deployment flexibility, maintainability, and ease of adopting changes.

The microservices impart the notion of 'Separation of Concerns' with each service performing as self-sufficient and participating as loosely-coupled component of the data warehouse architecture. The architecture has three Important operational sectors with responsibilities helmed by microservices for the data collection to data presentation pipelines.

```
Data            →    Data         →    Data
Collection           Analysis          Presentation
```

Data Collection:

The data collection pattern may be deployed as an Extract, Transform, and Load (ETL) or as an Extract, Load, and Transform (ELT) process. The data collection microservices provide the following functionality:

- **Data Extraction**: Data is gathered from multiple source systems that may be homogeneous or heterogeneous systems, such as RDBMS, No SQL, Spreadsheets, XML, JSON, CSV, TSV, Flat files, and other sources. The microservices can be deployed to collect data on-demand, on-schedule, event-based or as a combination. The microservices will also individually cater to the performance characteristics such as volume, latency, burst-mode, or spike-mode.

- **Data Cleaning**: Data that is gathered is parsed to find errors and rectify in compliance with pre-defined rule-templates. Individual microservices will perform the independent error monitoring and cleaning operations. Multiple microservices will work in a pipeline to mitigate all the relevant error types as dictated by the templates.

- **Data Transformation**: Data is converted from the source/legacy format to a uniform DW format irrespective of homogeneous or heterogeneous nature of the source data. Each microservice is tasked with a conversion responsibility such as format conversion, enrichment, collation, aggregation, joining, splitting, derivation, sorting, filtering, etc. Multi-step conversion will also be performed in a microservices' pipeline, where the order of conversion is established by rules.

- **Data Loading**: Transformed and error-free data is loaded into the data warehouse. The loading can be done as full load or incremental load. Incremental loading takes place at regular intervals that can be streaming increments or batch increments. During loading the data schema components such as indices, partitioning and views are created as dictated by the rules. Independent microservices will handle the individual responsibilities and complete combination functions as a pipeline.

Insys
Insightful•Innovative•Intelligent Systems

Data Analyses:

Data analysis can be done in relational format or multi-dimensional format. Multi-dimensional structure will be modeled based on facts and dimensions tables.

- *Tables Management*: *Microservices based modeling tools will help in creating the multi-dimensional data structures as facts and dimensions tables.*

- *Data Modeling*: *Microservices will help in pre-calculating, creating complex aggregations, and operational sets as models that capture for patterns, trends and relationships.*
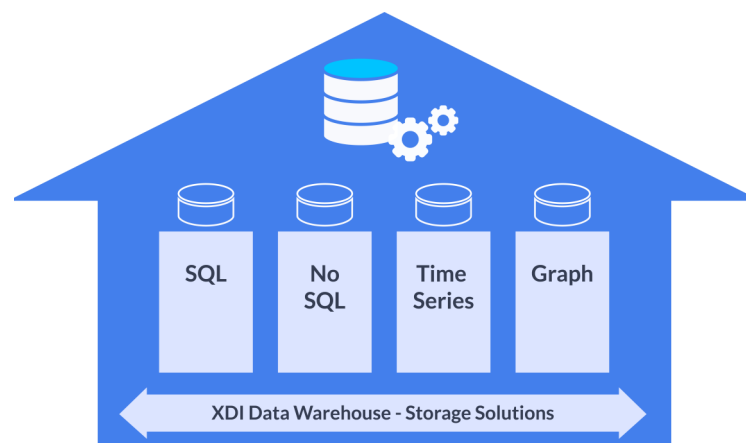
Data Presentation:

Data presentation facilitates the viewing of information from the data models – relational or multi-dimensional by relevant business stakeholders in their desired formats. Presentation framework will have single responsibility microservices to present different types of information / data to a variety of clients such as browsers, mobile devices or stand-alone applications. The microservices will transform and format the information/data to suit the clients.

Data Warehouse

An important aspect of storing data in a data warehouse is that the data from multiple sources is integrated and transformed in the context of a subject or functional area before storage, instead of directly loading the transactional data from the sources.

XDI Data Warehouse features include:

- **Flexible Storage**—SQL, No SQL, Time Series, or Graph.

- **Multi-Tiered Architecture**—Direct, through domain-specific data marts, or through data marts ad OLAP cubes.

- **Data Storage Consolidation**—Consolidate data from data lakes, data hubs, data virtualization, data federation, operational data stores, and data warehouses.

- **Security**—enterprise grade security to protect the data from extraction to when it's consumed or archived.

- **Scalable**—Easily scalable to grow as business needs change.

- **Fully Manageable**—Straight forward, easy-to-use interface for controlling the accessibility to data and the processes.

The **XDI Data Warehouse** supports the engineering of data from various sources into an accurate representation to support queries from users and applications in an unambiguous way. In the architecture, this is achieved by establishing each of the integrated applications as 'Systems of Record' (SoR) and the data warehouse as the 'Source of Truth' (SoT). As the SoT, the data warehouse will maintain only one source for every data element or entity by collating the data for the target entities from the applications.

The data warehouse component of the architecture will hold data extracted from transaction systems, operational data stores, and external sources and optimize it for retrieval. The optimization exercise involves data denormalization during which data may be duplicated to increase query performance using unique dimensional data models. It provides a single, central, and consistent store to support fast searches, complex queries, business and process analytics, and data mining.

Benefits of the **XDI Data Warehouse** include:

- Transform data into business intelligence

- Provide a single point of access by consolidating data from multiple sources

- Maintain a "Source of Truth" from entities in the system

- Reduce data duplication from applications and transactions

- Reduce stress on transactional data sources

- Optimize the repository for read access, flexibility, and high-speed query performance

- Utilize Master Data Management models with related hierarchies

- Improve data integrity, quality, and consistency

- Virtualize data in real-time

- Fully integrated with the **Cross System Integrator (XSI)** and **Cross Security Manager (XSM)**.

XDI employs a multi-layer defense strategy to protect critical data, individual privacy, and assure compliance, utilizing access control to the microservices, limited connections to XDI, and encryption.

T: (+1) 323-488-2846   **E:** contactus@3insys.com   **W:** www.3insys.com

A: 601 S Figueroa St #4650, Los Angeles, CA 90017

3Insys

Insightful•Innovative•Intelligent Systems