

# AGILE DATA INTEGRATION AND VISUALIZATION

---

## Delivering Rapid Business Values with xAQUA<sup>®</sup> Insight



### Abstract

Large amount of siloed data with rapidly increasing volume, variety and velocity (3V) is continuously compounding the data integration and visualization challenges. Traditional Extract Transform and Load (ETL) based data integration initiatives uses data driven waterfall approach that can't deliver rapid business values to the customers. Moreover, the 3V challenges associated with modern data are making the ETL based approach less and less efficient. The agile data integration and visualization solutions need to efficiently address the 3V challenges using a top-down business driven approach and necessary technical capabilities to establish the foundation for efficient Data Integration, Interoperability and Insight (3i).

## Introduction

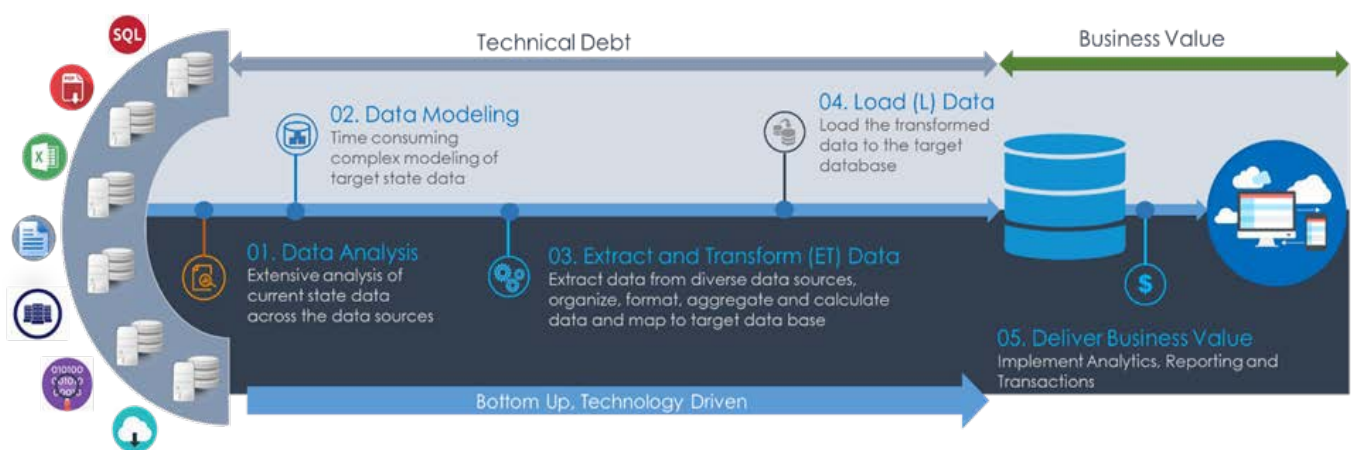
Modern enterprises are striving to explore the hidden knowledge locked within the data. Achieving this, is however, is a daunting task as challenges around data integration and interoperability are only increasing. Siloed data located in application silos, legacy systems, paper documents, electronic documents, and the web impede an organization's ability to aggregate, standardize, contextualize, and visualize data. Concerns around data integrity, quality, and source trustworthiness further compound the difficulties. In year 2001 Gartner analyst Doug Laney introduced the term "3V" to present the three dimensions of the challenges associated to enterprise data – the *volume, variety and velocity*.

## Challenges of Traditional ETL Based Approach

In the ETL based approach, data is extracted from multiple data sources, transformed to fit into the target data model and loaded into the target database. In this approach, the source datasets are analyzed, and a target data model is designed so that data from multiple data sources can be loaded into a single target data model. In most cases, data is cleaned, and data quality is improved before the data is loaded into the target database.

The ETL based approach uses bottom-up data driven waterfall model where the entire source datasets need to be analyzed upfront to design the target data model using schema-first approach and define the transformation rules for data load. Data from the source systems can then be extracted, transformed and loaded to the target database.

Traditional ETL based approach uses waterfall method for data integration and visualization and fails to produce quick business values for the customers.



---

Below are the key reasons why Traditional ETL Based Data Integration Approach Fails to Meet 21<sup>st</sup> Century Data Integration and Visualization Needs.

- ❑ Schema-first waterfall approach delays the delivery of the business values to the customer
- ❑ Business requirements need to be well understood at the beginning
- ❑ Rigid relational data model prevents semantic agility of data
- ❑ Challenges with integration of unstructured data using traditional ETL Tools
- ❑ Big bang ETL considers entire source datasets when business use cases need only subset of the integrated data
- ❑ Big bang data quality improvement uses data centric approach instead of business use case centric approach

## Agile Approach to Data Integration and Visualization

### Agile Data Integration Requires Top-Down Business Driven Approach

To deliver rapid incremental values to the business users, the data integration initiatives need to take a top-down business driven approach where only the subset of the source data that are required to implement the high priority business use cases are extracted, loaded and transformed in an iterative manner. This process continues iterative for the next set of business use cases delivering incremental business values to the business users.

Below are the key activities for agile data integration:

### 01 Ingest Data in Diversified Data Format into a Multi-Model Database

Establish capability to ingest structured, unstructured, text, and binary data in a Multi-Model database as-is without any transformation. This substantially cuts down the time and resource needed to extract, transform and load data using traditional ETL based approach.

### 02 Develop Incremental Semantic Data Models

No more big bang bottom-up data driven modeling of databases that results in rigid data models. Semantic data models are optimized to support the requirements for specific business use case(s). Because the target data models are defined for specific use cases, they can be changes very quickly and efficiently. As opposed to ETL approach, the semantic data models are developed iteratively that substantially cuts down upfront time for business services delivery.

### 03 Harmonize Data Incrementally to the Semantic Data Models

As opposed to traditional ETL based approach, data is harmonized to the semantic data models incrementally to support the specific use cases. This substantially cuts down upfront time for business services delivery.

### 04 Improve Data Quality Incrementally

The agile approach for data integration allows improving data quality incrementally. Traditional big bang data cleansing and data quality improvement initiatives often uses data centric approach for data quality improvement, whereas data quality requirements are specific to how the data is used – the data use case. The same data may be considered to have good quality for one use case and may not be for another usage scenario. The agile data quality improvement

takes top-down use case centric approach to define data quality requirements that are implemented, measured and improved iteratively.

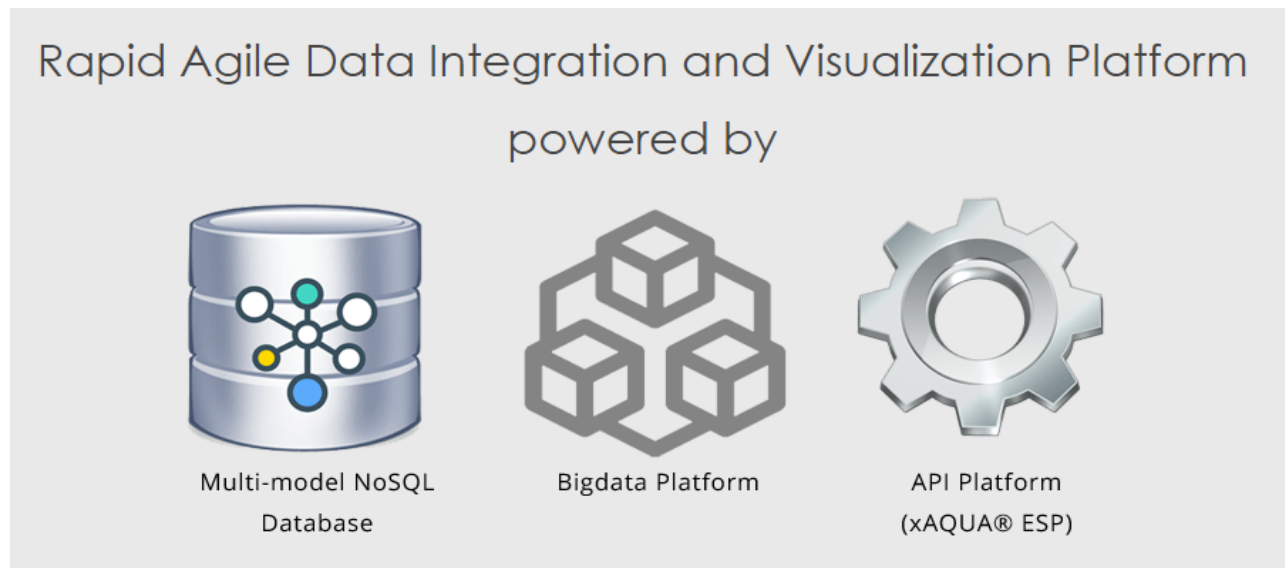
## 05 Establish Data Interoperability using Data as a Service (DaaS) API

Deliver the harmonized data through Data as a Service (DaaS) API for analytics data visualization.

## xAQUA<sup>®</sup> Insight

xAQUA<sup>®</sup> Insight is xFusion's solution for rapid agile integration, visualization and mastering of data.

xAQUA<sup>®</sup> Insight data integration solution is built on top of xAQUA<sup>®</sup> Enterprise Solution Platform (ESP) and combines the API First, Cloud First and DevOps First principles.



xAQUA<sup>®</sup> Insight leverages Multi-model NoSQL database and Bigdata platform for rapid agile data integration and visualization.

The integrated data is harmonized to semantic data models using top-down business driven approach. The harmonized data is then delivered through Data as a Service (DaaS) API for analytics data visualization.

The business-driven harmonization of data using xAQUA<sup>®</sup> Insight enables rapid agile data Integration, Visualization, and Mastering of Enterprise Data.

## xAQUA Insight address 3i<sup>®</sup> challenges with data

xAQUA Insight solution holistically addresses the three key elements of effective data integration and visualization: *integration, interoperability, and insight*—a quandary we address as '3i'<sup>®</sup>.

## xAQUA Insight Enables Agile Mastering of Data

xAQUA Insight solution comes bundled with Agile Data Mastering Framework (ADMF) that combines process, tools, technologies required to incrementally improve data quality using a business centric approach.



## Key Benefits of xAQUA<sup>®</sup> Agile DevOps

### FASTER

- Data Integration and Visualization

### BETTER

- Data Quality Management
- Flexibility of Data Model
- Risk Mitigation
- Performance and Scalability

### CHEAPER

- Development Cost
- Maintenance Cost
- Platform Cost

---

## About xFusion Technologies, Inc.

xFusion Technologies, Inc. is a State of California Incorporation, head quartered in Rancho Cordova, California, with offshore enters in Kolkata and Pune, India. xFusion is an expertise-based organization that specializes in serving various Public and Private sector organizations. xFusion offers a full spectrum of IT services in the areas of Enterprise Architecture, Legacy Systems Modernization, Custom Product Development, Outsourced Product Development, and Data Engineering. xFusion's xAQUA.io provides Platform Automation Functions as a Service (FaaS).

**xAQUA.io**

**General Platform Automation Functions as a Service (GPAFaaS)**

2893 Sunrise Blvd, Suite 202  
Rancho Cordova, CA 95742  
Phone: 916.668.6021  
Fax: 916 608 9697

info@xfusiontech.com  
www.xfusiontech.com  
www.xaqua.io