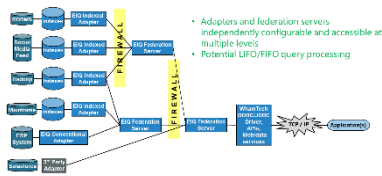


# SmartData Fabric® (aka EIQ Product Suite and EIQ Server)

## Basic and Advanced Training

Revision 1.9



## Contents

Introduction.....	3
Recommended Skills.....	3
Getting Started .....	5
Installation.....	6
Tutorial Preparation .....	6
Tutorial Scenario 1: EIQ Hybrid Adapter Configuration.....	6
Tutorial Scenario 2: Create an EIQ Federation Server .....	7
Tutorial Scenario 3: Master Data Management.....	7
Tutorial Scenario 4: Connecting to Tableau and Running Reports .....	9
Other Special Features.....	9
Create Indexed Views.....	9
Setup Realtime Updates through Query Based Polling.....	12
Configure Data Security and Access Control .....	16
Configure Row Level Security.....	18
Application Query Execution .....	19
Enabling SSL/TSL for Client Server Communication.....	20
Diagnostic Tracing/Logging .....	21
Audit Logging.....	23
Administration: Release Management, Backup, Maintenance, Recovery and Upgrades .....	24
Installation .....	24
Development .....	25
Testing .....	25
Backup and Maintenance .....	26
Recovery .....	27
Staging .....	28
Upgrades .....	28
Production .....	28
About WhamTech, Inc.....	29

## Introduction

This document is intended to list the high-level steps required for a basic implementation of Smart Data Fabric (SDF) (aka EIQ Product Suite). Many of the steps listed are available in a more detailed form in the freely accessible online user manual: [www.whamtech.com/sdf\\_help/v8/](http://www.whamtech.com/sdf_help/v8/). SDF is a comprehensive toolset that addresses a wide range of data management fundamentals. The following attempts to simplify the steps in the basic implementation. There are advanced aspects of these steps and referenced advanced features such as entity extraction, event processing and MDM, available in the user manual.

There are two types of training in the document, self-led training and instructor-led training. Some items are not available for hands-on learning, but are important for the process due to continuity and intended use later down the road. Instructor-led training is noted, all other topics are self-led.

New users should focus on steps 1-16. The rest of the steps in this document are more advanced steps and can be ignored until the basics are mastered.

## Recommended Skills

Before training begins, the team members need to be selected that are capable of performing the tasks that SmartData Fabric® tools require. The following are a list of roles and skills that team members should have some experience with before working with SDF®:

### Data Engineering Experience:

- Good relationships with data source system owners and DBAs
- Various data source types: Mainframes, RDBMS, Big Data SQL, NoSQL and non-SQL data
- Good working knowledge of file handling: FTP, encryption and compression
- General understanding of the function of scripts and system commands
- Running data extract, Changed Data Capture (CDC) and streaming data tools
- Experience with database connectivity (both between databases and between database systems and clients to databases) in a diverse computing environment
- Knowledge of stored procedures and data transformation
- Familiar with access control security, e.g., AD/LDAP, IAM, RBAC, ABAC/RLS and CLS
- Able to isolate problems between hardware and software, and provide information to appropriate development team(s)
- Highly developed, process-oriented skills for troubleshooting, problem solving and problem resolution
- Hands-on experience with RDBMSs, file systems, local area networks, distributed computing environments and information technologies

### ETL/Data Warehousing Experience (some overlap with the Data Engineering):

- ETL, data modeling and data architecture
- Data management, data warehousing and data marts
- Various data source types: Mainframes, RDBMS, Big Data SQL, NoSQL and non-SQL data
- Familiarity with unstructured data and text analytics
- Data discovery, profiling, transformation, quality, standards, security, lineage and governance
- Experience with Master Data management (MDM), metadata repositories and other systems of reference

- APIs – planning, design and implementation – and familiarity with HTTP and invoking web-APIs
- Work closely with data engineers, scientists and data/business analysts
- Work with business users to understand the business objectives
- Work with cloud and data architects to define robust architectures in cloud, on-premise, in data centers, SaaS and multi-cloud
- Workflow management

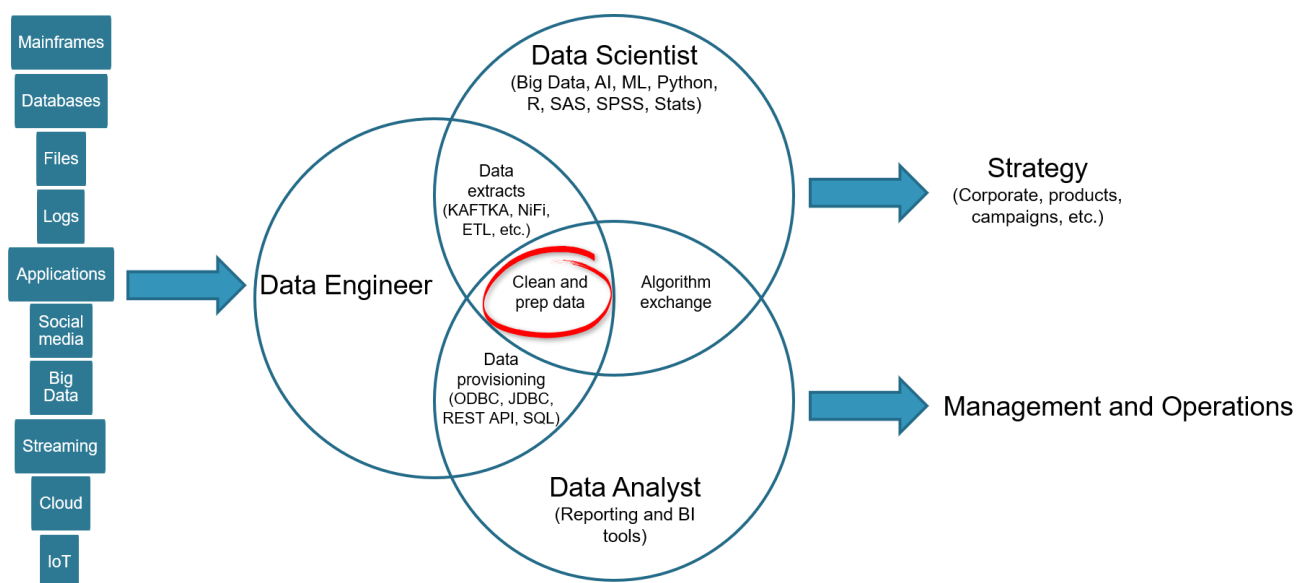
#### Data/Business Analysis:

- Learn, publish, analyze and improve management information dashboards, operational business metrics decks and key performance indicators
- Design, setup, execute and administer web-serving solutions.
- Solid experience in SQL
- Understanding of database theory and experience in at least one relational DBMS
- Knowledge on distributed applications/enterprise applications is plus.
- Strong knowledge of data structures, algorithms and designing for performance and scale.
- Familiarity with ETL
- Business Analysis
- Reporting/BI apps such as Tableau, Power Bi and QlikView
- Attention to detail
- Ability to troubleshoot
- Able to define and document issues, e.g., work with Jira

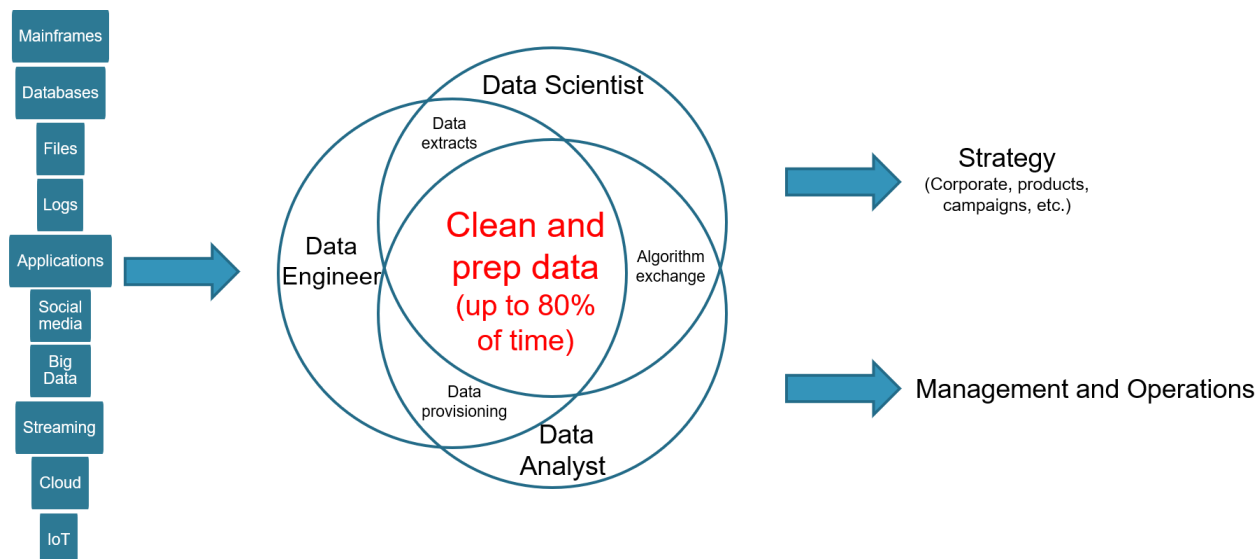
The number of team members with particular skills should be balanced. For example, a team of 5 members would be organized as such;

- Three (3) with ETL/data warehouse experience, as that is the majority of the work with our tools.
- One (1) data engineer, as they would be needed on the data discovery/sourcing/profiling side of things, assessing needs by data scientists and data analysts.
- One (1) data/business analyst who can help set up reports, BI and analytics.

WhamTech's goal is to enable the ideal data world and responsibilities depicted in the image below.



The reality is, Data Engineers, Scientists and Analysts spend a vast majority of their time cleaning and prepping data rather than their other daily tasks and responsibilities.



A well balanced team with the right skills and training in SDF® will help enable that ideal data world and avoid the reality.

## Getting Started

Step 1: Attend WhamTech-led session “What is SDF?” for an introduction to the SmartData Fabric® and EIQ Product Suite Tools, as well as a live demonstration of WhamTech’s capabilities.

Step 2: Explore the supplementary [Brief Technical Overview](#) presentation for more details about the SmartData Fabric® and EIQ Product Suite.

Step 3: View the following demo presentations on WhamTech’s [documents](#) page.

- ☐ [WhamTech EU General Data Protection Regulation \(GDPR\) Demo](#)
- ☐ [WhamTech Financial Services Demo](#)
- ☐ [WhamTech Mutli-level Security Demo](#)

Step 4: Watch the tutorial videos on [whamtech.com](#) or [YouTube](#) for video demonstrations on how the product works and for more depth on advanced capabilities. New users should focus on:

- ☐ [EIQ Tools Demo](#) - an introduction to the EIQ Product Suite tools and features.
- ☐ [Financial Services Demo](#) - more of the advanced capabilities of the SmartData Fabric® and how it can leverage an organization’s financial data.

Step 5: Become familiar with the [FAQ](#) section and [Knowledge Base](#) for answers to specific questions about the platform as well as supplementary blog posts that help convey the WhamTech ideology.



After completing the steps below, a follow up Q&A meeting should be scheduled so WhamTech can address any remaining questions. This can include any questions on features that do not appear in the tutorial steps and any clarifications on integration capabilities.

## Installation

Step 6: Read the [SmartData Fabric® Quick Start Guide](#) for installation information.

Step 7: Install the EIQ Product Suite on a Virtual Machine or other work environment.

Step 8: Configure the Admin user during the EIQ Product Suite install.

Step 9: Install ODBC drivers for all popular DBMSs on the same Virtual Machine.

*Note: In some cases, a pre-configured training VM may be provided. If so, ignore steps 6-9.*

## Tutorial Preparation

Step 10: Obtain sample databases from the Dropbox link provided.

Step 11: Read the [Tutorial Overview](#) for insight on what the tutorial scenarios will cover.

Step 12: Begin following the tutorial scenarios, using the sample databases as testing.

The tutorial scenarios guide users through the basic functions of the EIQ Product Suite such as creating EIQ Indexes, configuring an EIQ Hybrid Adapter, federating adapters, building MDM, and running reports on data using Tableau.

## Tutorial Scenario 1: EIQ Hybrid Adapter Configuration

Step 13: Work through Tutorial Scenario 1 to learn how to create an EIQ Hybrid Adapter on an operational data source. An EIQ Hybrid Adapter is one of the various configurations available to users to connect to data and access it. This particular adapter is special, as it allows the user to index, cleanse, standardize and transform the data that needs it (the basic functionality of an EIQ Indexed Adapter) and then access data that does not through a basic conventional adapter (EIQ Conventional Adapter functionality) at the same time.

- [Tutorial Scenario 1](#)
  - For more information see:
    - [EIQ Server RTI Tool](#)
    - [EIQ Server Configuration Tool](#)

At the end of Tutorial Scenario 1, users should know how to:

- connect to a data source with the EIQ RTI Tool,
- build an EIQ Index,
- register data sources,
- create a virtual data source,

- become familiar with a Business Data Dictionary
- perform a dictionary mapping to a Business Data View,
- create materialized views,
- create business views,
- configure an EIQ Hybrid Adapter, and
- query the data through the adapter using PGAdmin.

## Tutorial Scenario 2: Create an EIQ Federation Server

Step 14: Work through Tutorial Scenario 2 to learn how to configure an EIQ Federation Server. This tutorial introduces a new data source designed to imitate a data warehouse and will federate the operational data source from tutorial scenario 1 and the new data warehouse to make them queryable as if a single source.

Note: EIQ Server code can be configured for all three main EIQ Adapters (EIQ Hybrid Adapter, EIQ Indexed Adapter, and EIQ Conventional Adapter) and EIQ Federation Server, therefore, very similar terms and methodology used.

- [Tutorial Scenario 2](#)
  - For more information see:
    - [EIQ Server RTI Tool](#)
    - [EIQ Server Configuration Tool](#)

At the end of Tutorial Scenario 2, users should know how to:

- be more proficient with the EIQ Index build process,
- have a better understanding of creating virtual data sources,
- perform consistent and accurate dictionary mappings,
- have a deeper understanding of the importance of Business Data Views,
- create consistent materialized views and business views for the purpose of federation,
- register a VDS as a Remote EIQ Data Source for the purpose of federation,
- create an EIQ Federation Server, and
- have a familiarity with querying federated data.

## Tutorial Scenario 3: Master Data Management

Step 15: Work through Tutorial Scenario 3 to learn how to define entities and build master data.

A major benefit of the EIQ Product Suite is the ability to create and maintain master data from data across multiple data sources, in a distributed manner, without copying master data to a single and centralized location. The option remains to centralize or regionalize as required. Distributed MDM (a) improves/facilitates integration with distributed data sources, (b) improves performance and (c) reduces data loss risk. MDM with the EIQ Product Suite optionally provides a seamless and automatic use of master data as part of normal query processing that uses separate master data indexes to isolate any and all related operational/transactional data in multiple data sources. Once operational/transactional data is isolated and read, master data can be optionally substituted for similar attribute operational data, resulting in very clean

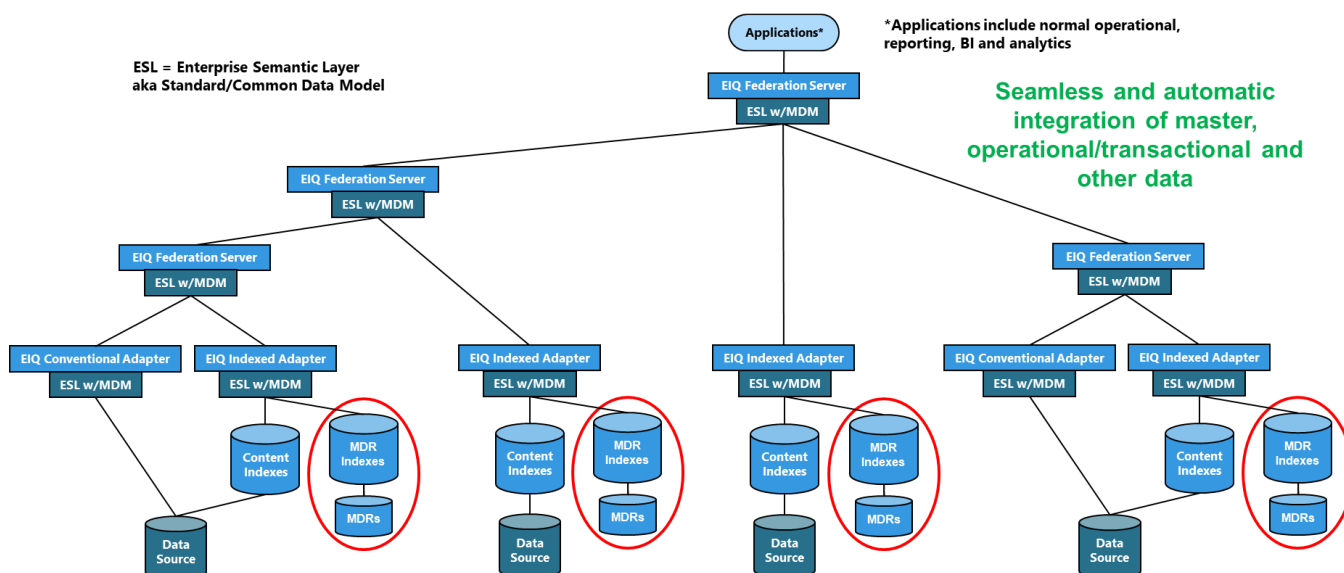
results data. This is all the while not putting query load on the underlying data sources. This is accomplished by taking advantage of built-in data federation and Link Indexes™ of EIQ Product Suite.

The master data generation process is based on finding matching attribute values for master data entity attributes. Master data entities can be simple or complex. Simple entities may have one or two attributes whereas complex entities may have other entities as attributes. An example of simple entities can be a contact EMAIL or PHONE\_NO. An example of complex entities is a PERSON having first, middle and last name(s), fuzzy first, middle or last name(s), ADDRESS, EMAIL, PHONE\_NO, DOB and SSN as attributes.

Internally, the master data build process takes the following steps for each entity type configured for master data generation:

1. Binning entities based on a selected subset of entity attributes using exact match of raw or fuzzy version of the attribute values.
2. Linking records within each bin.
3. A detailed scoring of entities in a bin based on closeness of attribute values using edit-distance algorithms.
4. Grouping and filtering matches based on given cutoff threshold (records in same group are considered as belonging to the same entity).
5. Merging matching entities based on given merge rules to generate a master data record for that entity.

The following diagram illustrates distributed Master Data Records and Indexes:



There are also options to mask, tokenize and/or encrypt master data, e.g., using third-party Format-Preserving Encryption (FPE).

All entries in master data records are indexed and can be joined to normal content indexes and Link Indexes for seamless and automatic integration and real-time updates.



- [Tutorial Scenario 3](#)

- For more information see:

- [EIQ Server Configuration Tool](#)
    - [Master Data Management Help](#)

At the end of Tutorial Scenario 3, users should know how to:

- add, delete and edit entities,
- access the MDM build area of the EIQ Server Configuration Tool
- configure matching rules for building MDM,
- understand matching attributes and how to customize matching parameters,
- build MDM,
- map MDM to its own unique Business Data View,
- create a Business View for MDM, and
- query MDM.

## Tutorial Scenario 4: Connecting to Tableau and Running Reports

Step 16: Work through Tutorial Scenario 4 to learn how to connect the EIQ Federation Server as a data source to Tableau, running reports on MDM, federated indexed data, and federated conventional source data as if it all originated from a single source.

SmartData Fabric® can seamlessly integrate with Tableau to enable the multitude of benefits of the EIQ Product Suite™ for Tableau's advanced BI and analytic capabilities. The following document will detail how to configure the EIQ Adapters and Federation Server so that they are ready for a Tableau connection and the various ways that connection can be made.

- [Tutorial Scenario 4](#)

- For more information see:

- [Tableau Integration Help](#)

At the end of Tutorial Scenario 4, users should know how to:

- connect Tableau to an EIQ Adapter or EIQ Federation Server,
- run a report on top of two federated EIQ Hybrid Adapters with MDM, and
- have a significant understanding of the benefits EIQ Adapters and Federation Server can provide to an organizations daily operations.

## Other Special Features

Some of these features are documented on [https://www.whamtech.com/sdf\\_help/v8/](https://www.whamtech.com/sdf_help/v8/) -> EIQ Special Features; others, such as Event Processing, are in process of being documented.

## Create Indexed Views

Indexed Views in the EIQ Product Suite are materialized views with multiple uses:

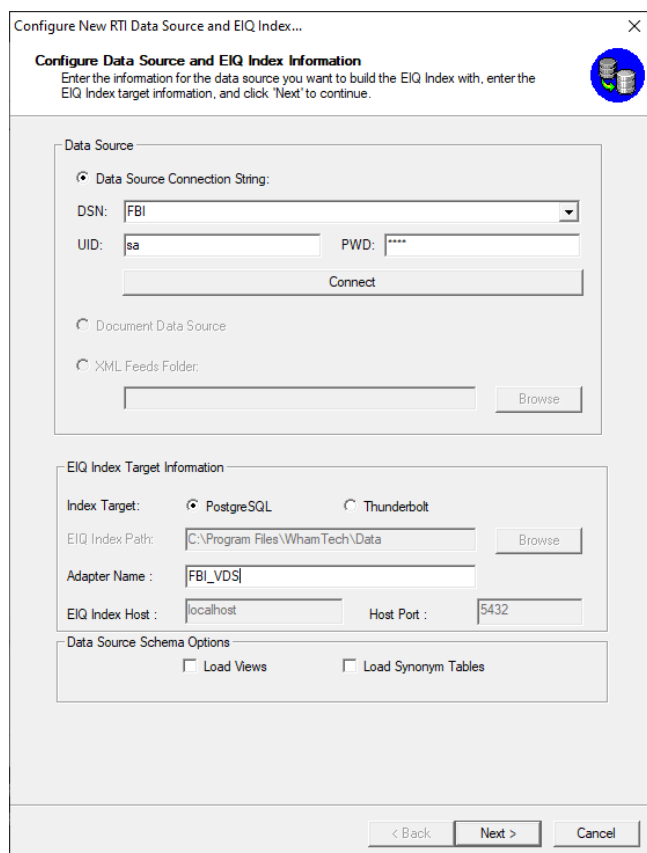
- They improve query performance for most frequently submitted and resource intensive queries.
- They provide a flattened view for applications to use.

An indexed view is defined using either the EIQ Server RTI Tool or any EIQ Server client by specifying the SQL statement to create the view. The views are maintained in real-time by EIQ RTIS. It updates EIQ Indexes when the original data changes.

Indexed Views are covered in Tutorial Scenario 6.

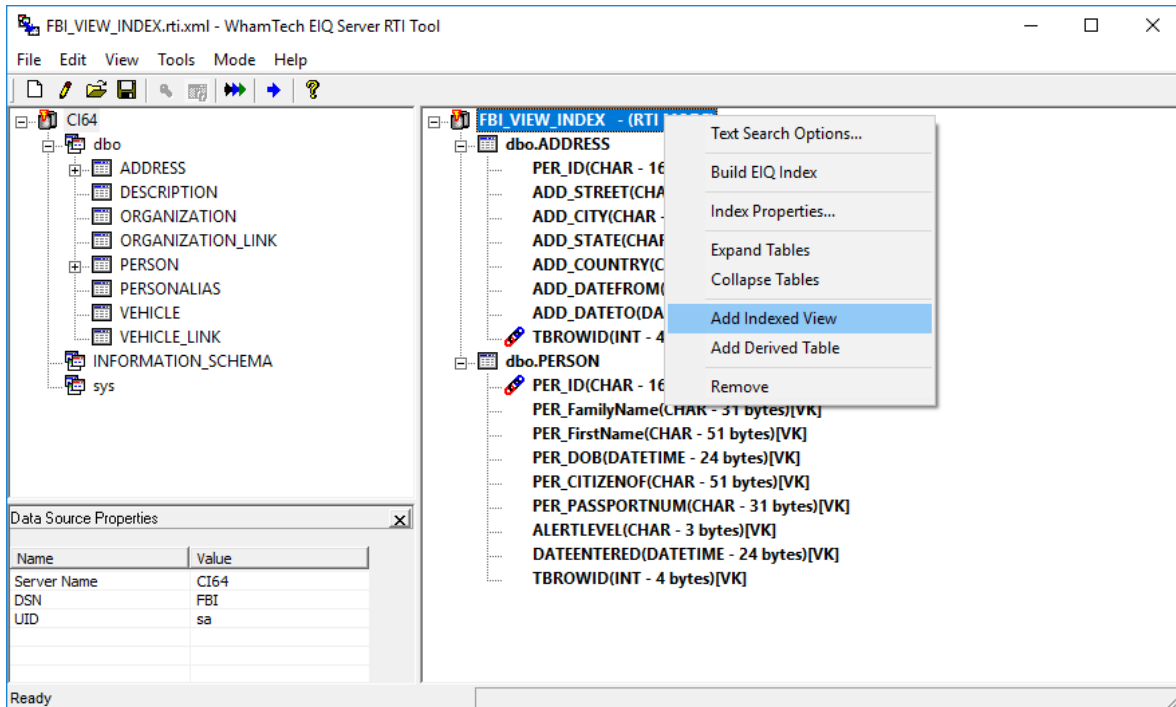
Step 17: Open the EIQ Server RTI Tool.

Step 18: Create a new index or open an existing one.

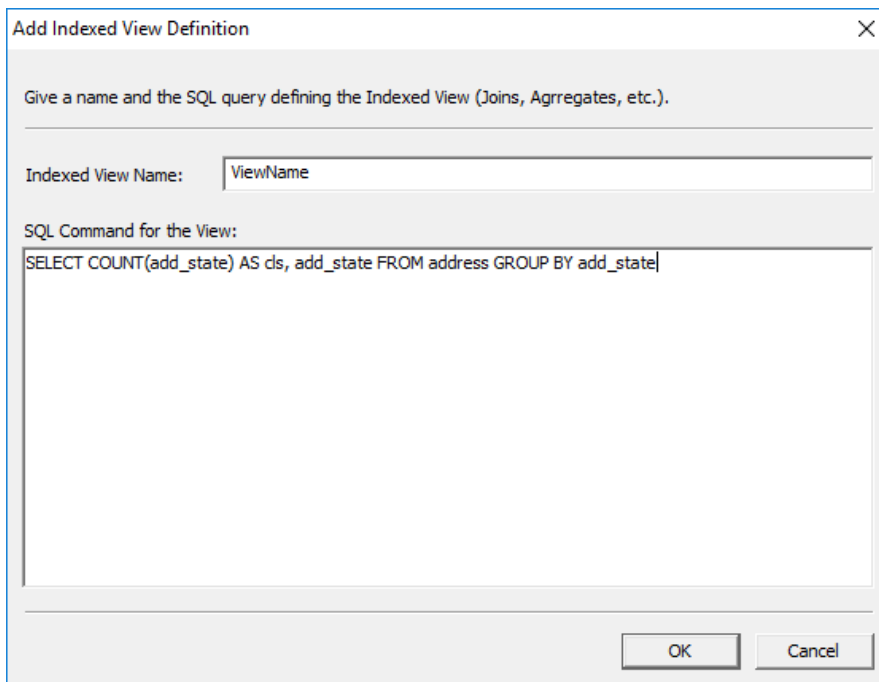


Step 19: Switch to RTI mode.

Step 20: Select Add Indexed View.



Step 21: Define the indexed view using SQL.



Step 22: Specify initial size of the view.



Indexed View Initial Size

Enter the initial number of rows for the indexed view:  
1000

Enter the maximum number of rows for the indexed view:  
10000

Referenced table information  
"dbo"."ADDRESS" has 15158 records

Group fields:  
"dbo"."ADDRESS"."ADD\_STATE"

OK Cancel

Step 23: Build or re-build the EIQ Index.

FBI\_VIEW\_INDEX.rti.xml - WhamTech EIQ Server RTI Tool

File Edit View Tools Mode Help

CI64

- dbo
  - ADDRESS
  - DESCRIPTION
  - ORGANIZATION
  - ORGANIZATION\_LINK
  - PERSON
  - PERSONALIAS
  - VEHICLE
  - VEHICLE\_LINK
  - INFORMATION\_SCHEMA
- sys

Data Source Properties

Name	Value
Server Name	CI64
DSN	FBI
UID	sa

Ready

FBI\_VIEW\_INDEX - (RTI MODE)

- dbo.ADDRESS
  - PER\_ID(CHAR - 16 bytes)[FK]
  - ADD\_STREET(CHAR - 31 bytes)[VK]
  - ADD\_CITY(CHAR - 31 bytes)[VK]
  - ADD\_STATE(CHAR - 51 bytes)[VK]
  - ADD\_COUNTRY(CHAR - 31 bytes)[VK]
  - ADD\_DATEFROM(DATETIME - 24 bytes)[VK]
  - ADD\_DATETO(DATETIME - 24 bytes)[VK]
  - TBROWID(INT - 4 bytes)[PK]
- dbo.PERSON
  - PER\_ID(CHAR - 16 bytes)[PK]
  - PER\_FamilyName(CHAR - 31 bytes)[VK]
  - PER\_FirstName(CHAR - 51 bytes)[VK]
  - PER\_DOB(DATETIME - 24 bytes)[VK]
  - PER\_CITIZENOF(CHAR - 51 bytes)[VK]
  - PER\_PASSPORTNUM(CHAR - 31 bytes)[VK]
  - ALERTLEVEL(CHAR - 3 bytes)[VK]
  - DATEENTERED(DATETIME - 24 bytes)[VK]
  - TBROWID(INT - 4 bytes)[VK]

ViewName

## Setup Realtime Updates through Query Based Polling

EIQ RTIS provides a query-based polling option to monitor for changes in the data sources and apply them to the corresponding EIQ Indexes. As opposed to polling a message queue for change messages, in query-based polling EIQ RTIS directly polls the data source tables or associated CDC table for changes at a frequency specified by the user.

Step 24: Activate SQL Server Agent if it is not already running.

Step 25: Activate CDC on the SQL Server databases if CDC is the update method of choice.



Step 26: Open the EIQ Update Configuration Tool.

EIQ Update Configuration Tool

SmartData Fabric™  
EIQ Update  
Configuration Tool

Protected by U.S. Patents  
Copyright © 2017 WhamTech, Inc.  
www.whamtech.com

EIQ Server Address:  
127.0.0.1

EIQ RTIS Address:  
127.0.0.1

User Name:  
admin

Password:  
\*\*\*\*\*

☐ Always use these settings

Connect Cancel

Step 27: Create a new Task.

EIQ Update Configuration Tool...

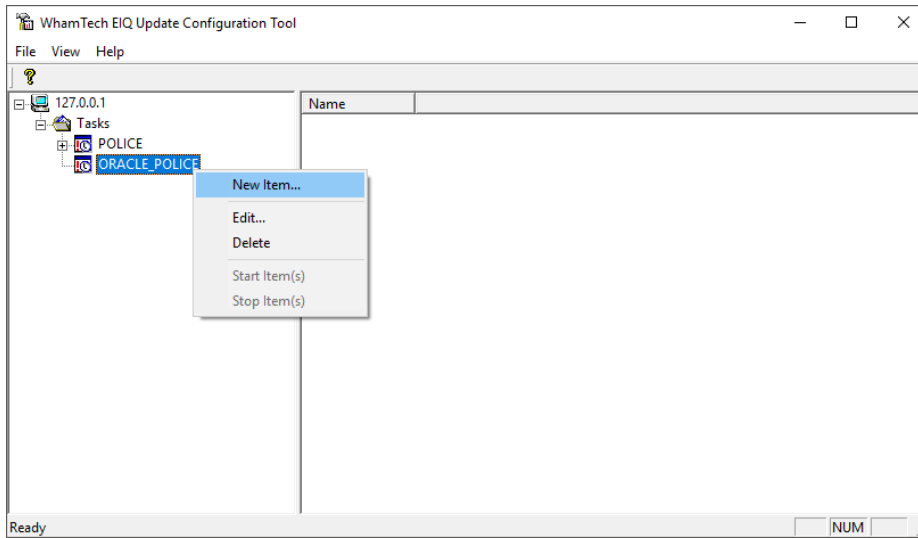
**Configure Task**  
Enter a task name. Press 'Next' to continue.

Task Name:  
ORACLE\_POLICE

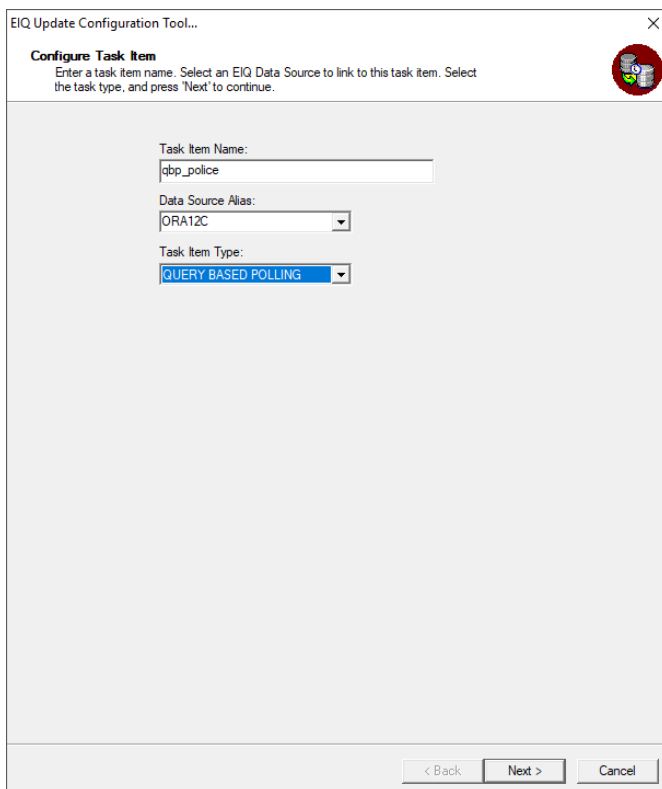
☒ Use Virtual Data Source

< Back Next > Cancel

Step 28: Create a new Task Item.



Step 29: Select Query Based Polling as the Task item type.



Step 30: Enter the name and credentials for the data source.

Step 31: Select the data source platform. SQL Server for CDC; Other for Oracle.

Step 32: Set the Polling Frequency.

Step 33: Schedule Polling if desired.

Step 34: Select specific column names for Oracle data sources.



EQ Update Configuration Tool

**Change Query-based/Rest API Configuration Setup**  
Configure the Query-based Polling or REST API settings, specify the other task parameters, and click 'Finish'.

Task type: Query based Polling

**QueryBased Polling Data Source**

Data Source Name (DSN): ORA12C

User ID (UID): \*\*\*\*\*

Password: \*\*\*\*\*

Test Connection

Data Source Platform: OTHER

**Rest API Credentials**

User Name:

Password:

File Path:

Browse

**Polling Frequency for updates**

Poll Interval (seconds): 5

Start/Last Timestamp: 03-DEC-18 8.50.00 AM

Initial Timestamp:

Batch Size: 5000

☐ Allow SQL Timestamp

**Schedule Polling**

Start date: 12/ 3/2018

Occurs at: 9:00:00 AM

Occurs every: day(s)

☐ Enable Deletes

**Table specific data:**

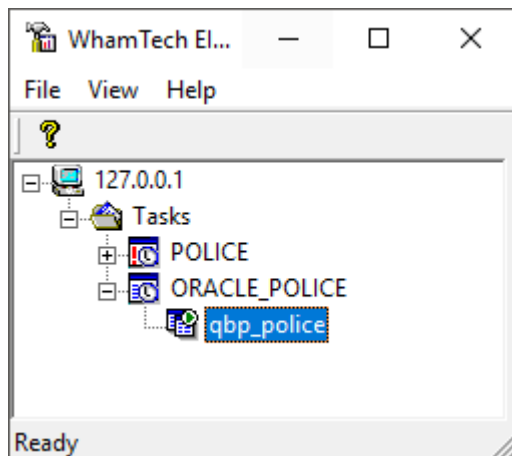
Please edit the Column Name and the Value by double clicking the field. ☒ Show only Timestamp columns

Schema Name	Table Name	Column Name1	Column Name2	Value
dbo	INCIDENT			
dbo	LOPERSON			
dbo	PERSINCIDENT			

< >

< Back Finish Cancel

Step 35: Start the Task.

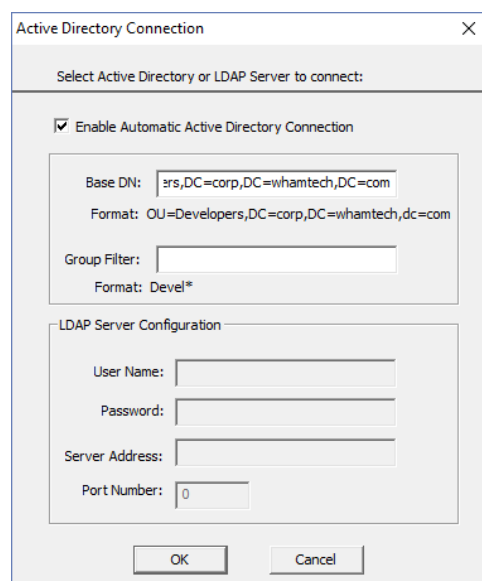


## Configure Data Security and Access Control

The EIQ Product Suite allows Admins to control how data is accessed, how data is viewed, and who is allowed to view specific data. This function is provided through the EIQ Server Configuration Tool under the 'Management' Tab. Keeping data secure is vital to organizational functionality and day-to-day operations.

Step 36: Open the EIQ Server Configuration Tool and navigate to the 'Management' tab.

Step 37: Connect to a Domain, LDAP Security Group, or OVD.



Active Directory Connection

Select Active Directory or LDAP Server to connect:

☒ Enable Automatic Active Directory Connection

Base DN:

Format: OU=Developers,DC=corp,DC=whamtech,dc=com

Group Filter:

Format: Devel\*

LDAP Server Configuration

User Name:

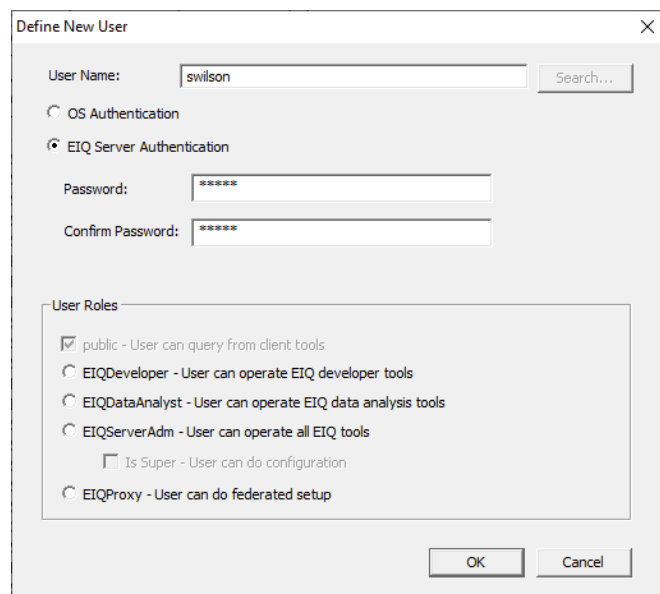
Password:

Server Address:

Port Number:

OK Cancel

Step 38: Create a new user; selecting OS Authentication to import a Domain user.



Define New User

User Name:  Search...

☐ OS Authentication

☒ EIQ Server Authentication

Password:

Confirm Password:

User Roles

☒ public - User can query from client tools

☐ EIQDeveloper - User can operate EIQ developer tools

☐ EIQDataAnalyst - User can operate EIQ data analysis tools

☐ EIQServerAdm - User can operate all EIQ tools

☐ Is Super - User can do configuration

☐ EIQProxy - User can do federated setup

OK Cancel

Step 39: Create a new Role.



New Role

Role Name:

Employee

OK

Cancel

Step 40: Add User to Role.

WhamTech EIQ Server Configuration Tool on <Josh-Laptop.corp.whamtech.com> (adminsUPER)

File Help

Virtual Data Sources | Dictionary Mapping | Data Sources | Advanced Settings | Management

Refresh

Virtual Data Sources

Client Sessions

Users

Roles

SDFBuild

SDFConfig

SDFQuery

SDFConfig

SDFProxy

EIQDeveloper

EIQDataAn

EIQServer/

SDFdomain

SDFQuery\_

SDFRevers

Manager

Employee

Role m

Access

Supply Cha

SupplyChai

SDFQuery\_

SDFQuery\_

QueryUser

SDFQuery\_

FedProxyGi

SDFQuery\_

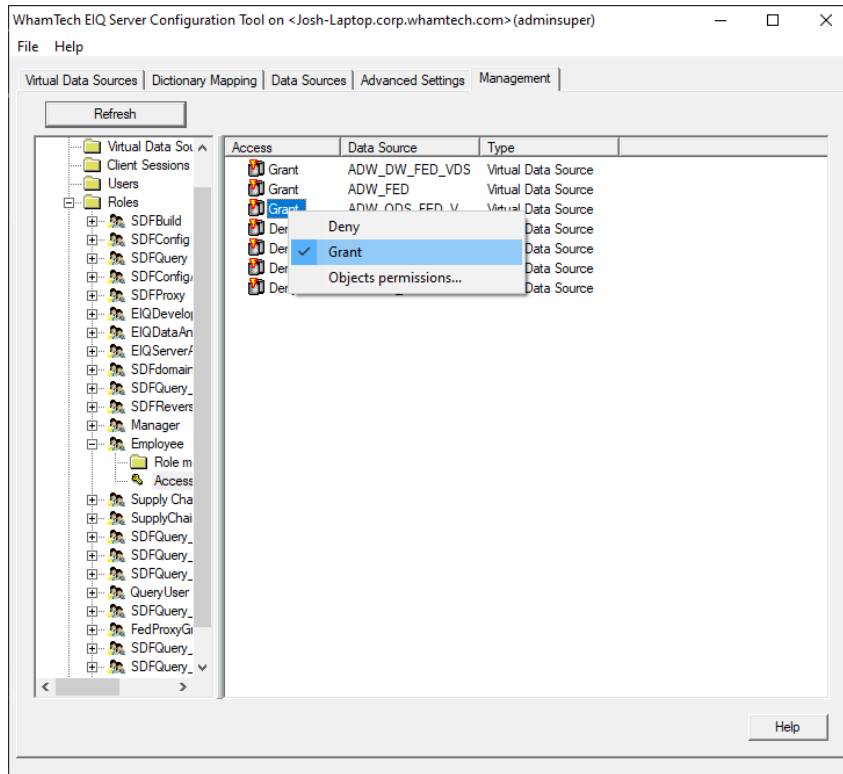
SDFQuery\_

Name	Type	Member
dataanalyst	USER	NO
corp\briany	USER	NO
adminsUPER	USER	NO
postgres	USER	NO
PM_FED_PM_...	USER	NO
queryuser1	USER	NO
proxyuser	USER	YES
serveradmin	USER	NO
swilson	USER	YES
configadmin	USER	NO
tstark	USER	NO
corp\sridham	USER	NO
srogers	USER	NO
corp\vn	USER	NO
joshp@corp	USER	NO
corp\vinayk	USER	NO
WORKS_FED...	USER	NO
ADW_DW_FE...	USER	NO
developer	USER	NO
ADW_ODS_F...	USER	NO

Add...

Help

Step 41: Grant role privileges to desired data sources.

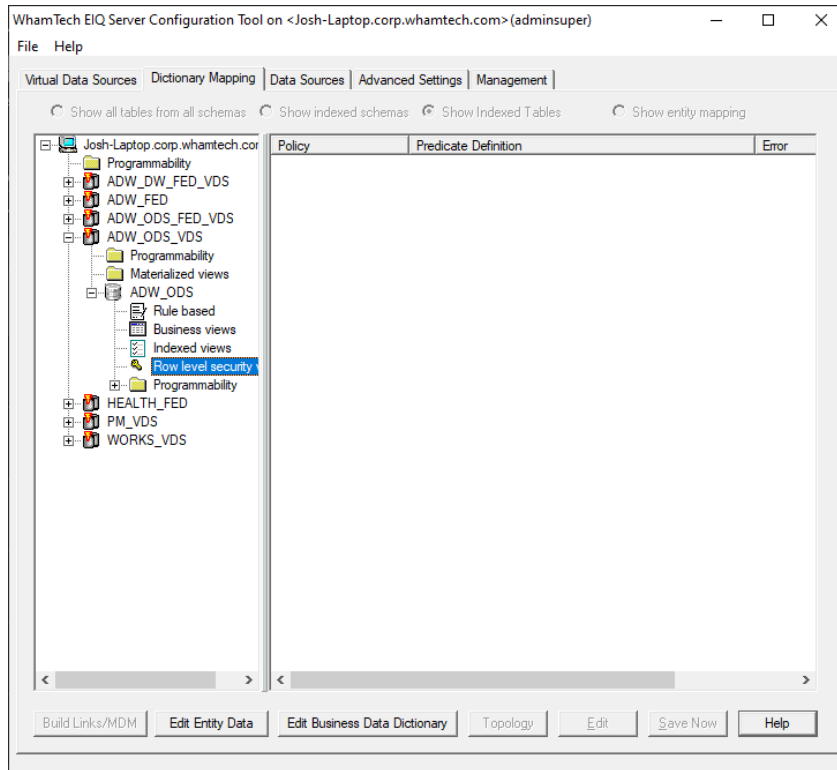


## Configure Row Level Security

Row-level security enables administrators to control access to specific rows in a database table based on the user executing the query. What differentiates this method of security is that users can still access all of the tables, but specific rows are hidden to them defined by a view that preforms a security check based on user attributes.

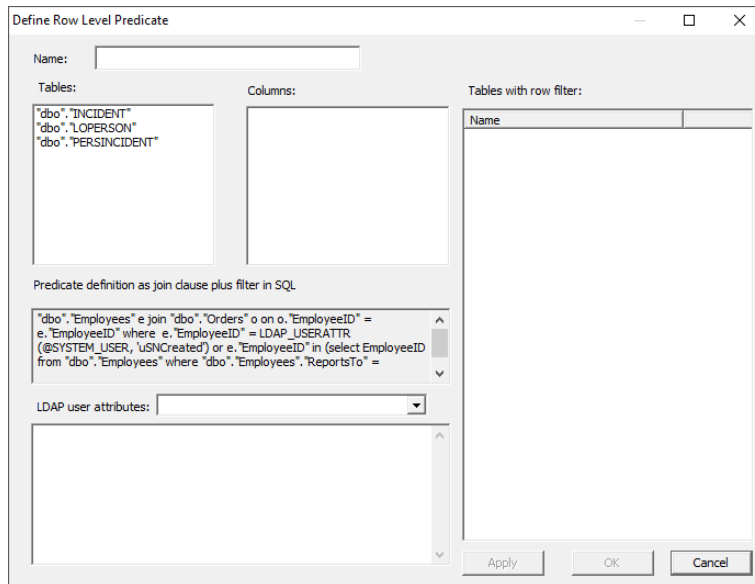
Step 42: Open the EIQ Server Configuration Tool and go to the 'Dictionary Mapping' tab.

Step 43: Select a Virtual Data Source and expand the nodes until 'Row Level Security view' is available.



Step 44: Create a new row level security policy.

Step 45: Use the dialog to define the row level security policy.



## Application Query Execution

Applications, such as BI, reporting, analytics, CRM and others, connect to any of the EIQ Products, i.e., EIQ Adapters and EIQ Federation Servers through standard ODBC and JDBC drivers, REST APIs and other means, similar to databases, and submit standard ANSI SQL for query execution.

In addition to ANSI SQL, EIQ Products can support query languages with conversion to ANSI SQL. Some PL/SQL is already supported by EIQ Products, but development is underway for more complete PL/SQL to ANSI SQL conversion. SPARQL to SQL third-party conversion tool reviews are also underway. And, testing is underway for Spark SQL and other Spark query options, e.g., ML. Other efforts are focused on enabling applications to work against any data source or sources through query transforms and inverse data transforms to and from standard data views, i.e., completely decoupling application queries from any schema dependency.

Part of the EIQ Products value proposition is being able to write back to data sources as well as read from them. This has been developed in conjunction with data services (REST APIs) and new workflows developed in BPM software for interactive interoperability solutions in healthcare.

Depending on the standard data view used, e.g., HL7, FHIR, NIEM or XBRL, data can be represented simultaneously as (1) data object/semantic/ontological models through APIs and virtual graph database/triple store, (2) relational views where a query constraint determines the fact table and (3) NoSQL/Big Table.

## Enabling SSL/TSL for Client Server Communication

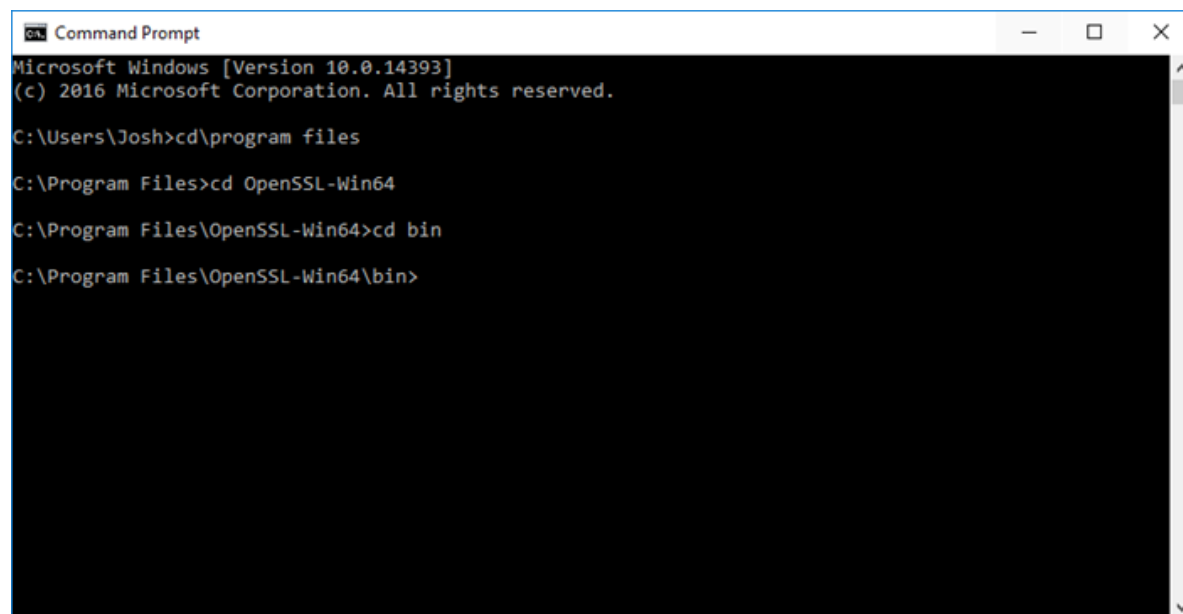
TLS/SSL is used in applications to create secure communications over computer networks and inhibit eavesdropping. It is also used to ascertain the identity of the client at the other end of a connection.

The EIQ Product Suite uses OpenSSL to enable secure communications across server and client modules. For more details on enabling SSL/TSL, click [here](#). The following steps are optional.

Step 46: Download and install OpenSSL for Windows.

Step 47: Make sure the path to the .dll files are added to the %path% environment variable for the Windows user account

Step 48: Open the command prompt and change the directory to the location of the OpenSSL installation.



```
Microsoft Windows [Version 10.0.14393]
(c) 2016 Microsoft Corporation. All rights reserved.

C:\Users\Josh>cd\program files
C:\Program Files>cd OpenSSL-Win64
C:\Program Files\OpenSSL-Win64>cd bin
C:\Program Files\OpenSSL-Win64\bin>
```



Step 49: Use OpenSSL to generate a key and certificate by creating a batch file.

Step 50: Go to the working directory and run the batch file.

The output of the previous steps generates 5 files:

- key.pem
- cert.pem
- clientkey.pem
- clientcert.pem
- trusted.pem

The first 2 files (key.pem, cert.pem) strictly relate to the server. The next 2 files (clientkey.pem, clientcert.pem) strictly relate to the client. The trusted.pem file is needed both at the server and client side.

Step 51: Create a subdirectory .eiq under user account's home directory if it does not already exist. For example, ("C:\Users\USER\_ACCOUNT\eiq").

Step 52: Copy clientkey.pem, clientcert.pem and trusted.pem files to .eiq directory on the client machine and rename the clientkey.pem to key and clientcert.pem to cert.pem.

Note: If the client and server programs are both running on the same machine, clientkey.pem and clientcert.pem are not needed, since the client is local to the server. In this case, just copy key.pem, cert.pem and trusted.pem to the .eiq directory on the server machine.

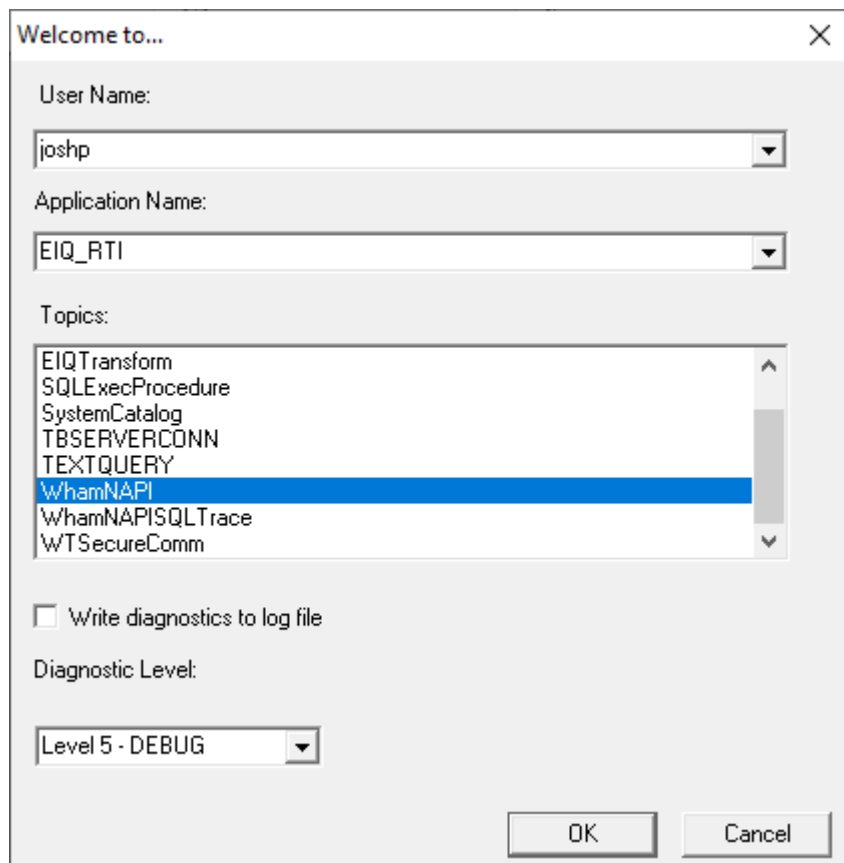
Step 53: Restart/Start EIQ Server after making these changes.

## Diagnostic Tracing/Logging

The EIQ Product Suite provides an in-house diagnostic tool that is used to view diagnostic messages from EIQ Product Suite components. This diagnostic trace aids users and developers in tracking down and diagnosing any issues with EIQ Product Suite components.

Each EIQ Product Suite component (executables and libraries) writes diagnostic messages to Windows message sinks. The EIQ Diagnostics Tool taps into these message sinks and displays messages as they arrive. Users must open a diagnostic window for each component they are interested in monitoring.

Upon launching the tool and signing in, users are presented with this window:



The first drop down is the User Name. Generally, this is the user profile running the EIQ Server.

Next is the Application Name. This dropdown menu shows the list of EIQ applications and services the user has launched.

Then, is the Topic box. This box shows all of the possible topics to monitor under the selected application.

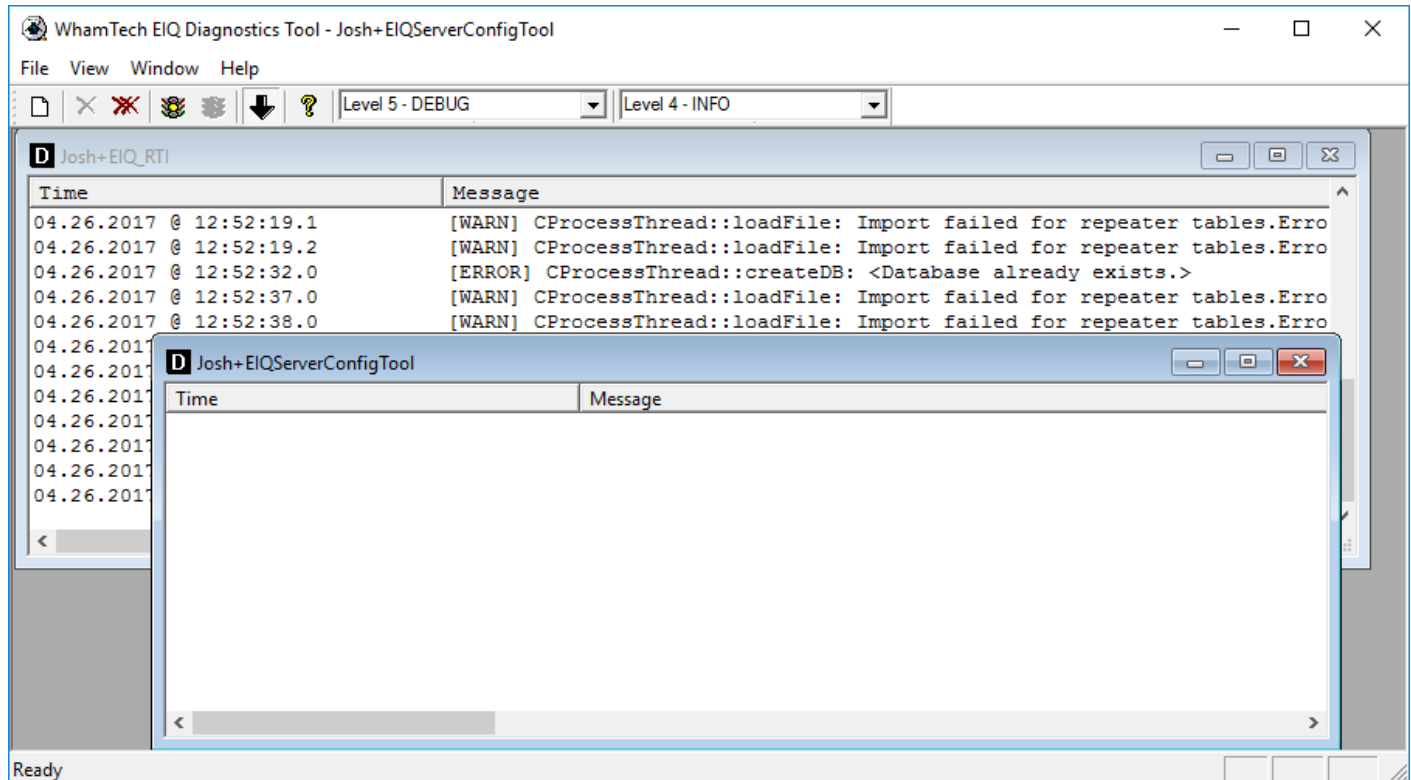
When users select a component, by default the 'Write diagnostics to log file' option is selected. This writes diagnostic messages to a log file named after the component under the 'WhamTech\log' folder in the install path.

And finally, the Diagnostic level. This dropdown will show the possible monitoring levels. Selecting a level will show all of the levels above it.

The possible diagnostic levels are:

- Level 1 - Fatal: A very serious error has occurred that can lead to the application aborting.
- Level 2 - Error: An error has occurred but the application may still continue to run.
- Level 3 - Warning: There is a potentially harmful situation occurring.
- Level 4 - Info: Provides informational messages highlighting the procedures of the application.
- Level 5 - Debug: Provides details on informational events useful in debugging the application.

Once the appropriate selections have been made, the diagnostic window will open in the larger pane behind the selection window. Multiple logging windows can be opened and monitored at the same time. Organize them by clicking 'Window' from the toolbar and selecting 'Cascade' or 'Tile'.



## Audit Logging

WhamTech Audit Logs are security logs captured by the EIQ Server Suite. The logs track significant temporal events that enable administrators or investigators to reconstruct suspicious activities in the event of a security breach.

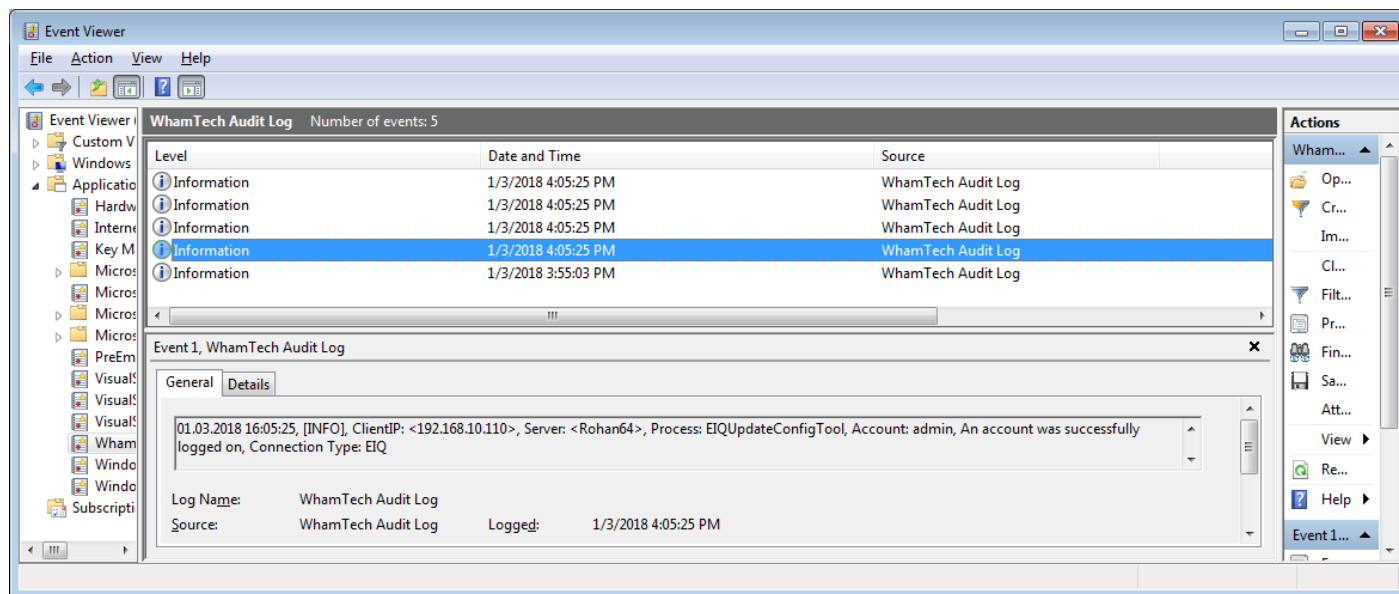
At a high level it captures:

- User authentication attempts, connections and disconnections
- Additions, deletions and modifications of User definitions, grants and privileges
- User database accesses and successfully and unsuccessfully executed queries

On Windows, WhamTech Audit Logs are captured using the Windows NT Event Logging system.

- Open the **Event Viewer** application

Audit Logs are viewed in the WhamTech Audit Log under Applications and Services Logs:



The audit logs will capture a wide range of messages from data source management, user account control, queries, index builds and task item activities.

## Administration: Release Management, Backup, Maintenance, Recovery and Upgrades

The software release management guidelines for the SmartData Fabric® solution should include the following stages:

- Development
- Testing
- Staging
- Production

Ideally, each stage in the release management process should be completed in a discrete environment, separate from the other environments. Realistically, it may be necessary to combine one or more of the environments due to hardware, time, or other resource constraints. At a bare minimum, the production environment should be separate from the other environments.

Modern datacenter and cloud platforms provide several built-in features that help with configuring and instantiating the environments for each of the phases as needed. The features such as software configurable automatic backup/restore, high availability, load balancing, resource scaling, tiered storage etc. ease this process considerably.

## Installation

SmartData Fabric® and EIQ Product Suite installation requirements vary based on several factors: the purpose of installation (developer or production deployment), types of data sources, number and size of the EIQ Indexes, query processing needs, redundancy, update frequency, etc.





## Systems Configuration Considerations:

- A virtual machine (VM) with SmartData Fabric® pre-installed can be used as a basis for creating multiple development machines for multiple developers to work in parallel on different data sources.
- For development, sever machines may not be needed (Windows 7/Windows 10/Linux desktop OS machines may suffice).

For more specific system and hardware requirements, please see the [EIQ Product Suite Quick Start Guide](#). This online document also helps users install the EIQ Product Suite, start the servers, and setup client access to data sources.

## Development

In this phase, the main focus is to understand the data sources, data source connectors, schema/data model, data profiling, data quality, transforms, common data model, standard data view mapping, MDM, and CDC/Polling options, and build data source adapter and federation configuration(s) per solution requirements.

Development users require flexibility to explore various options during the development phase and should be provided with sufficient access privileges on the development system(s) and data sources.

It may be sufficient to use a single machine for configuring multiple adapters and federation server or use additional machines as needed. At this time the following limits apply:

- A single instance of WhamTech's EIQ Server per machine.
- A single configuration client per EIQ Server at a time.

The development phase may use development/test data sources with simulated/generated data that mimics actual production data source data as much as possible. However, access to actual data will help with the data cleansing/standardization process.

The following items generated during development can be reused in future stages:

- RTI Maps/transforms for data standardization and cleansing
- Adapter configuration/standard data view mapping
- Federation, MDM configuration, developed client application queries
- CDC/Polling configuration to update indexes
- Sizing parameters for hardware/systems configuration for test/staging/production environment

## Testing

In this phase, the environment is setup with specific tests in mind. These typically include tests for functionality, security, access control, performance and scale.

VMs and artifacts generated during development phase can be re-used/re-purposed/applied in the new environment with some modifications suitable for the new environment. For example, data sources adapters may need to point to a different data source with larger dataset.

The following are the most common tests in this phase:

- Performance and Scale tests: These tests typically require setting-up an environment different from development environment. For example, each of the adapters and federation servers need dedicated

machines with configurations that are appropriate for production-level performance and scale under target loads.

- Back-up/recovery scenario tests: These tests include scenarios for backing-up configurations/indexes etc. for recovering/restoring after a planned or unexpected outage. These tests should cover various boundary situations and would help improve the reliability and robustness of the solution.
- Process automation: Automating various tasks in testing, staging and production phases using scripting should be considered to improve efficiency, reliability etc.

As with the development phase, the following additional generated objects can be saved and reused in future phases:

- Load parameters, performance metrics
- Security and other compliance parameters
- Test scripts
- Automation scripts
- Input for staging and production systems specification

## Backup and Maintenance

During or after the testing phase is a good time to establish a backup and maintenance strategy. Certain objects from the development phase and future phases can be reused, and objects and artifacts generated by the EIQ Product Suite during production are vital to keeping operations moving.

These output artifacts along with WhamTech software should be saved for reuse and maintained in a repository/folder for efficiency, traceability and recoverability. Saving images of VMs should also be considered for easy replication/recoverability.

Especially once the environment has moved to production, it is recommended that admins backup configuration related files regularly to enable quick recovery to the most recent working state in case of failures. Before copying configuration files, make sure that all EIQ Product servers are stopped. Stop the servers through 'Windows Services'.

The following are EIQ Product Suite files that need to be backed up regularly:

- Index Files: The index files built through the EIQ Server RTI Tool in "Profiling" and "RTI" mode are stored in "C:\Program Files\WhamTech\Data" by default, under folders named after the respective index names. If you chose a folder other than default folder when building indexes, copy the index files from that folder.
- EIQ Server Configuration Files: All the EIQ Product Server configuration information, such as virtual data sources, registered data sources, and Standard Data View mappings, are stored in the "Settings\EIQConfig" folder under the EIQ Server Installation folder. By default, this is under "C:\Program Files\WhamTech\". The "EIQConfig" folder needs to be backed up.
- EIQ RTIS Configuration Files : All the configuration information related to EIQ RTIS is stored in files under "Settings\EIQUpdateConfig". These files can be backed up from the default "C:\Program Files\WhamTech\Settings\EIQUpdateConfig" folder. Make a copy of the entire folder "EIQUpdateConfig".
- JMS Bindings: The information related to JMS bindings is stored in "C:\Program Files\WhamTech\JMSConnectionFactory" by default.

Apart from these files, admins should make note of relevant system environment variables. This information is essential for restoring or re-creating the system environment for an EIQ Product deployment. The information includes:

- Data Sources: Any information related to data sources including client configuration details
- ODBC DSNs: Names of the ODBC DSNs for the related data sources
- Logical Drives and Paths: The full path information for each of the files backed-up above
- MSMQ/JMS Queues: Names and properties of MSMQ and JMS Queues

## Recovery

Should a disaster or other event where files are lost occur, there are a few recommended ways to recover those files from the backups.

### Scenario 1: Restoring to a Last Known Configuration

If configuration files get corrupted for any reason or there is a need to revert back to a known working configuration, all that an EIQ Server Admin needs to do is to replace the present copy of files with the old copy from backup.

Before you proceed, first make sure to stop the EIQ Product related servers. Replace the files mentioned in the previous section with the files in the backup.

### Scenario 2: Moving/Re-deploying EIQ Product Suite Configuration

In the event a machine stops working or there is another critical system failure, the configuration can be moved to a new environment. When moving or redeploying an existing EIQ Product Suite deployment configuration to a new system, there is one extra major step involved: an EIQ Admin needs to re-create the original system environment on the new system. This includes installing data source client software, re-creating DSNs with original names, re-mapping logical drives, creating file/folder paths, creating queues, etc.

Steps for recreating the system environment:

- Set up the required data sources by attaching the databases with the appropriate names and settings.
- Create ODBC DSN for the databases with the appropriate names and related databases.
- Create logical drive maps and paths.
- Load the backup files and store them at the right places.
- Create the related message queues and make sure that the names are consistent.

Once the original environment is faithfully re-created on a new system, copy the backup files to their corresponding locations.

## Staging

Staging environments should be setup to mimic production environment, and, in some cases, prepared to be switched into production. Sizing, load, performance and other metrics collected in the previous phases should be used to configure the systems for staging and production environments.

In this phase, data source adapters are built from scratch against full data sets. Often, the data sources used for building adapters are replicated versions of original production systems. In some cases, even production data sources are used. Federation and MDM are built against these adapters. Testing by end users can start in this phase.

Apart from functionality, load and performance verification; plans for backup/recovery/restore, software/systems upgrade and other scenarios, such as adding new data source adapters to an existing solution, should be reviewed and addressed in preparation for deploying into production. Any enterprise requirements for compliance etc. should also be addressed at this phase before deploying in production.

## Upgrades

During Staging and Production, there may be updates to the SmartData Fabric®. These could be anything from bugfixes, new and/or requested features or updated query support. These updates are supplied as a patch.

When upgrading the EIQ Product Suite software, follow any special installation and un-installation instructions for the upgrade. If there are no special instructions given, then follow the standard un-installation and installation procedure. Standard installation procedures leave any EIQ Product Suite deployment configurations untouched.

## Production

The production environment is usually setup similar to the staging environment. In this phase, the solution is deployed against production data sources and is called by production applications used by end users.

Plans should be developed and tested for outages and automatic switching back to a backup environment.

There can be an initial gradual roll-out plan for the first time that involves limiting access to a certain type of users, certain applications, etc.



## About WhamTech, Inc.

WhamTech, Inc. (WhamTech) is a privately-held US-owned Delaware Corporation established in October 2000 and based in Dallas, Texas. WhamTech's mission is to develop indexed adapter-based data virtualization, federation integration and interoperability technology software products. WhamTech develops these products to anticipate, meet and exceed the demands of customers seeking an alternative to the conventional approaches of data warehouses, federated data access with conventional adapters and enterprise search. Our goal is to provide a better and more seamless way to work with data, by changing the way data is accessed, analyzed, integrated, shared and interoperated. WhamTech adapter products are based on independent, cleansed indexes that execute both structured and unstructured queries for data warehouse capabilities, including results when data sources are unavailable and archive.

Information on WhamTech solutions, sales and services, and partnership and investment opportunities can be obtained through [whamtech.com](http://whamtech.com).

Copyright © 2023, WhamTech, Inc. All rights reserved. This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. Composite and Denodo are registered trademarks of their respective owners. Other names may be trademarks of their respective owners.