

FAQs

1. **What is the browser compatibility for logging into the TCS Connected Intelligence Data Lake for Business Portal?**

Please check whether you are using Mozilla Firefox 18 or above and Google Chrome 26 or above.

2. **What is TCS Connected Intelligence Data Lake for Business Software for AWS?**

TCS Connected Intelligence Data Lake for Business Software for AWS combines powerful business data lake capabilities with the agility and flexibility of public cloud infrastructure. Benefits include self-service, under-an-hour provisioning, and pay-as-you-go — plus the ability to extend an on-premises ecosystem into the AWS environment. Additional advantages:

Rapid and easy procurement — quick proofs of concept for Data Lake exploration

On-Demand or Term subscriptions — pay only for what you use, optimize your spend

By incorporating AWS for deploying a business data lake analytics environment, we make it easier for companies of all sizes to become data-driven with best-in-class offerings.

3. **For whom is TCS Connected Intelligence Data Lake for Business Software for AWS intended?**

Any organization of any size can use and benefit from TCS Connected Intelligence Data Lake for Business for AWS, which combines the power of the TCS Connected Intelligence Data Lake for Business with the convenience of AWS for a win-win value proposition. The ability to self-provision, launch quickly, pay-as-you-go, and integrate with a wide variety of complementary software tools are all key benefits.

4. **Is TCS Connected Intelligence Data Lake for Business for AWS single-node or multi-node (i.e. SMP or MPP)?**

TCS Connected Intelligence Data Lake for Business Software for AWS is single-node. Higher levels of scalability will be available in the future.

5. **How secure is TCS Connected Intelligence Data Lake for Business Software for AWS?**

TCS Connected Intelligence Data Lake for Business Software for the AWS Marketplace offers all of the same security options that exist for any Amazon EC2-hosted offering. For example, EC2 compute instances are located in a Virtual Private Cloud (VPC) with an IP range that you specify. You decide which instances are exposed to the Internet and which remain private. Security Groups and networks access control lists (ACLs) allow you to control inbound and outbound network access to and from the instances. You can connect your existing IT

infrastructure to resources in the VPC using industry-standard encrypted IPsec VPN connections.

Furthermore, you can provision your EC2 resources as Dedicated Instances, which are Amazon EC2 Instances that run on hardware dedicated to a single customer for additional isolation. We refer you to AWS for details. For more information on Amazon EC2 security, refer to the Amazon Web Services: Overview of Security Process document.

6. What compliance levels are in place?

The same compliance levels that exist for all AWS Marketplace-deployed software are in place for TCS Connected Intelligence Data Lake for Business Software for AWS Marketplace. We refer you to AWS Marketplace for those specifics. Ultimately, you are responsible for ensuring that you meet whatever requirements are applicable for your industry, operating environment, and end-user expectations.

7. What AWS resource monitoring capabilities are available?

Amazon Cloud Watch is a web service that provides monitoring for AWS cloud resources and applications, starting with Amazon EC2. It provides you with visibility into resource utilization, operational performance, and overall demand patterns — including metrics such as CPU utilization, disk reads and writes, and network traffic. You can get statistics, view graphs, and set alarms for metric data. Amazon CloudWatch aggregates and stores monitoring data that can be accessed using web service APIs or Command Line Tools.

8. What is AWS Marketplace?

AWS Marketplace is an online store that helps customers find, buy, and immediately start using the software and services they need to build products and run their businesses. Visitors to the marketplace can use AWS Marketplace's 1-Click deployment to quickly launch pre-configured software and pay only for what they use. AWS handles billing and payments, and software charges appear on your AWS bill.

9. What do I need in order to subscribe to software on AWS Marketplace?

Any customer eligible to use AWS products is also able to shop on AWS Marketplace. You must have an Amazon EC2 account. Software charges appear on the same monthly bill as AWS infrastructure charges.

10. How do I pay for AWS Marketplace subscriptions?

If you are using a TCS Connected Intelligence Data Lake for Business AMI product that has an hourly fee, you will be billed by AWS for your use of this software. At the beginning of the month, you will receive a bill from AWS for your AWS Marketplace charges, which will include one or both of these types of charges:

The hourly fee for the software multiplied by the number of hours an instance with this software was running.

Your infrastructure charges such as EC2 usage, EBS usage, and Bandwidth usage.

For EC2 resources, each instance will run until one of the following occurs: you terminate the instance, the instance shuts itself down, or the host terminates due to software or hardware failure. Partial instance hours consumed are billed as full hours.

11. What happens to AWS Marketplace when TCS updates the TCS Connected Intelligence Data Lake for Business?

When TCS publishes an update to an AMI to which you are subscribed, you will be sent a notification email from AWS Marketplace that contains information about the latest version.

12. What service level can I expect?

TCS Connected Intelligence Data Lake for Business Software for the AWS Marketplace offers all of the same service levels that exist for any EC2-hosted offering. We refer you to AWS for details.

13. Which software I can use to develop an UI on top of TCS Connected Intelligence Data Lake for Business?

You can use any open source or commercial software to develop UI. Amazon Quick sight or Tableau are our suggestions.

14. Where can I find the user guide for developing use cases using TCS Connected Intelligence Data Lake for Business?

Please follow the below link to access the user guide

<https://dss.tcs.com/wp-content/uploads/2019/01/tcs-connected-intelligence-data-lake-user-guide.pdf>

15. Where can I find the assets – data/data models used in the airline delay demo use case in the how to videos?

Please follow the below link to access the demo data/data models

<https://dss.tcs.com/wp-content/uploads/2019/02/demoassets.zip>

List of Abbreviations

Abbreviation	Expanded Form
CRUD	Create, Read, Update and Delete
CSV	Comma Separated Value
DIP	Data Induction Platform
ID	Identification
IUX	Intelligent Urban Exchange
JDK	Java Development Kit
RDBMS	Relational Data Base Management System
XML	Extensible Markup Language
BI	Business Intelligence
CEP	Complex Event Processing
CIP	Connected Intelligence Platform
ETL	Extract Transform Load
FI	Final Inspection
PMML	Predictive Model Mark-up Language
REL	Release
SCD	Slow Changing Dimension
TCS	Tata Consultancy Services Ltd.
UI	User Interface
API	Application Program Interface
ARTS	Association of Retail Technology Standards

CAS	Central Authentication Service
DDL	Data Definition Language
DS&S	Digital Software and Solutions
GUI	Graphical User interface
HLD	High Level Design
KPI	Key Performance Indicator
LLD	Low Level Design
PSS	Shared Services Group
URL	Uniform Resource Locator
URM	Universal Record Management
EDW	Enterprise Data Warehouse
ETL	Extract, Transform and Load
FTP	File Transfer Protocol
HDFS	Hadoop Distributed File System
HTTP	Hypertext Transfer Protocol
IEG	Information Exchange Gateway
IoT	Internet of Things
JMS	Java Message Service
JSON	JavaScript Object Notation
LDAP	Lightweight Directory Access Protocol
MQTT	Message Queue Telemetry Transport
PMML	Predictive Model Markup Language
REST	Representational state transfer

SSO	Single Sign-on
SQL	Structured Query Language

Troubleshooting

1. **I am accessing the URL in a browser but the view is not clear and the panels are overlapping/hidden. What could be the issue?**

Please check whether you are using Mozilla Firefox 18 or above and Google Chrome 26 or above.

Due to slow internet connectivity this kind of issue may happen. Login through Amazon workspace would give better and fast results.

2. **I am trying to upload a data model but it is showing error "Invalid File Type". What could be the problem?**

Please check file type of the data model template. File type should be in MS Excel (.xls or .csv format) and open office spreadsheet

3. **I am trying to upload a data model but it is showing an error "*internal content is corrupted or is invalid as per the specified format. Check logs for more details*".**

The template you're using may not be correct. Please use correct template for the data model creation. Refer user guide for correct template.

4. **I am trying to upload a file template for creating data source in the CIP Portal. It is showing error "*Please input a file with proper template*". What could be the problem?**

The file should contain either the names of columns only or data with first row (header) containing the name of the columns. Please use correct template for the data source creation. Refer user guide for correct template.

5. **While running the pipeline, I am getting an error in the data transformation step - "*local directory or remote directory might be null! Data source must be a sink*". What could be the cause?**

There is no source input file in the desired path. Before pipeline creation it is mandatory to create sftp and hdfs path. Sftp path is the remote directory where we store the source files and local is the hdfs path.

6. **While running pipeline I am facing error - "*invalid source input data in the data transformation step*". What could be the cause?**

Please check the following:

- i. Correct data source is selected.
- ii. Is Header Present: Check this box while creating data source, if the data file contains names of the columns in the first row.
- iii. Provide correct properties of the fields (Field Type – INT, String and so on., Field Size – No. of characters if type is string, Field Format – DD/MM/YY for date and so on.) Refer user guide for list of field properties.

- iv. Check if proper delimiter character is chosen. Delimiter Character: Character that defines the separation between columns, for example, Comma (,) or Semicolon (;) or Tab (\t) and so on. Refer user guide.
- v. Finally check the datatype of the data being uploaded matched with the datatype in the metadata while creating data source.
- vi. Check the given sftp and hdfs path physically exist.

7. I am getting the error "*Please verify the query*" while running a dataset. What could be the cause?

Check that the table name is preceded by the schema name and the data store name. Check the SQL query syntax.

Don't put any semi colon (;) or dot (.) at the end of query statement

8. I am unable to test a dataset. What could be the cause?

Please check whether Apache Drill is up and running. You can check drill status in ambari console.

9. Apache drill is up and running. I am still unable to test or run a dataset. What could be the cause?

Please check whether you are using a valid data store.

10. While trying to save data store I am getting invalid data source error.

Please make sure path is setup during project creation. Also use the private IP obtained from AWS console instead of URL.