LTIMindtree contribution to OSDU Forum

Introduction

The Open Subsurface Data Universe (OSDU) is a vision of multiple Oil and Gas (O&G) companies like Shell, Schlumberger, BP, Equinor, Dell Technologies, Devon, Epam, Exxon, Chevron, Microsoft, and Teradata to build a standard inter-operational platform for operating different geotechnical applications with well-defined rules and policies. It is an organization called OSDU Forum, a collaboration of different industries that combines multiple exploration and production datasets into one platform.

According to the March 2021 survey, there are over 194 members of the OSDU Forum, including oil and gas operators, service providers, and cloud service providers. The OSDU data platform is a reference architecture designed to support cloud-native data platforms. It's not a stage that can be readily used for production data. The OSDU Forum members aim to develop tools and services to curate subsurface data and develop codes to be used by clients and third parties on their cloud environments as per their business requirements. The codes are publicly released for platforms like core services, entitlement and obligation services, data loading and ingestion, and EDS. These codes can be used to automate the workflows created in different projects. These codes are released after thorough testing performed by developers and domain team members from the forum. This helps users to take help from OSDU-defined standard codes and implement them directly to their working project rather than doing the basic research and preparing the codes. The client developers and team members are highly benefitted from these pre-structured data frameworks and codes that ensure their legacy dataset is converted and visualized in any application that is OSDU native.

The core concept involves isolating data from applications, so users can access data by any OSDU-compatible data platform.

LTIMindtree boasts a distinguished legacy of executing domain-centric engagements in OSDU data platforms. As proud silver members of OSDU, our techno-domain team is adept at comprehending diverse O&G data sources and honing their expertise in varied OSDU platforms. Our core services, showcased below as case studies, encompass OSDU Entitlement & Obligation, Audit & Metrics, Schema Validation, and other OSDU platforms. Our services facilitate management's assessment of dynamic policies within OSDU environments while empowering users to validate multiple schemas against results that O&G majors have embraced. Mercury Release R3 is the first release that includes a full open-sourced data platform contributed by Schlumberger. From Mercury Release R3-M10 through R3-M17, our LTIMindtree team has been instrumental in developing a Key Performance Indicator (KPI) dashboard, delivering domain consulting, executing project management, and testing activities.

Business Challenges

Different exploration and production companies face issues in managing, organizing, storing, and accessing their subsurface data that is in pace with the increasing volume of data and growing advanced technology to adapt multiple data formats. Since both operating and service-providing companies are facing the same issues, they collaborate to understand the business challenges and adopt standard rules to foster innovation and increase efficiency in business outcomes.

Companies have continuously faced multiple challenges in migrating subsurface data to the OSDU environment. Different initiatives taken into account for data transformation are dealing with multiple challenges that can be categorized in the following terms:

- Data is an important asset that remains under-utilized if not properly managed and curated.
- Data structuring, management, and proper understanding to handle different formats of data.
- Silosed data result in a lack of better collaboration between cross-domain users.
- Certification policy for users and documentation.
- Data ingestion processes and dashboard development for business and management purposes.
- Data authenticity can be a bigger challenge in the absence of lineage information.
- Testing the core services and applications in different environments.
- Data entitlement and obligations.
- Workflow and process performance checks.
- Automation for schema validation and checks.
- Application Programming Interface (API) testing in various environments (IBM, Azure, AWS, GCP) of the platforms of OSDU is a challenge.
- Companies follow an old system of data management (hard copies) that is not compatible with the future technology of the industry.

Multiple companies are facing the above-listed challenges, and the solution can be better identified when multiple companies work together rather than trying to find the solution in isolation. Users with different data set experiences can work together on a common platform to create a framework to resolve such issues instead of working in silos, with little knowledge of other complexities data may have.

Various business challenges to which LTIMindtree has contributed are described below in detail:

| OSDU entitlement and obligation

Data security and access have been a significant concern for users to access their data in different cloud environments. Since data is the owner's property, she can manage access by selecting which users can view or edit the data. This entitlement process is an important platform in OSDU where entitlement and obligation policies and APIs are tested and run to authorize data accessibility by multiple users.

One of the primary challenges to enabling users with data access is determining the nature of the data to which they are entitled to access. Confirming whether the user belongs to the data group, service group, and the service they are endeavoring to access entails separate processes that must be verified to ensure data security. Service authorization is necessary to corroborate whether the client or service invoking another service has adequate approval.

| OSDU audit and metrics

Multiple users seek a solution to measure performance KPI checks during various OSDU activities like search, delivery, ingestion, governance, and platform usage for the data hosted in OSDU in a multi-cloud environment. Since measurement and analysis of various platform and data level KPIs will help management to understand business workflows, the client wanted a solution to be built for these KPIs. Alert generation and notification on usage and performance issues will help clients handle business problems and challenges.

| OSDU schema validation

In the OSDU data platform, the schema is defined as a structure that provides information for the data record files in JSON format. It also defines whether the data record is a floating point, string, integer, list, or other data type. Understanding whether search API indexes the information related to the schema is an important element that should be validated before data search or ingestion. So, users should create the schema and provide the necessary information in a manifest file before migrating and ingesting to the cloud environment.

Since schemas are defined based on data types, the query of the respective API will generate results for the defined schema. This API testing is a time-consuming manual check process of responses against the standard OSDU schemas.

Software testing suite

Users from multiple domains are migrating their data from on-prem to the cloud, and face challenges in accessing the data and understanding data formats and types. Much time is invested in data search, visualization, loading, ingestion, and testing different collections in different environment variables. Since API testing helps in the creation, full search, and filtered search of the schema resources in the open data for industrial service.

Solutions

LTIMindtree's techno-domain team has been continuously helping the community find solutions to the challenges users face in cloud environments.

Different OSDU services and LTIMindtree contributions have been listed below in Figure 1,2,3:

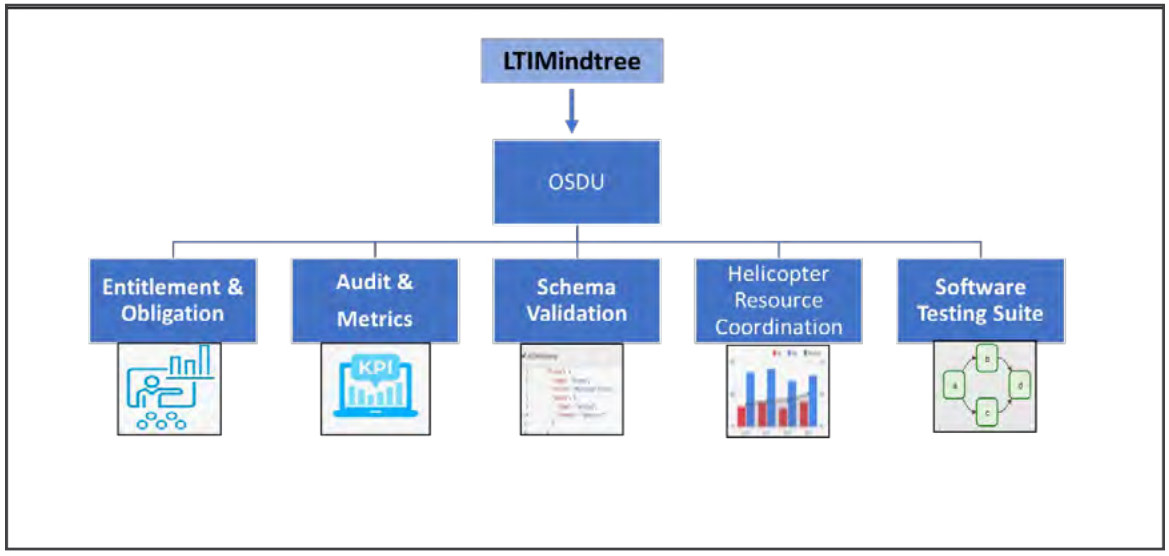


Figure 1. LTIMindtree contribution towards OSDU data platforms

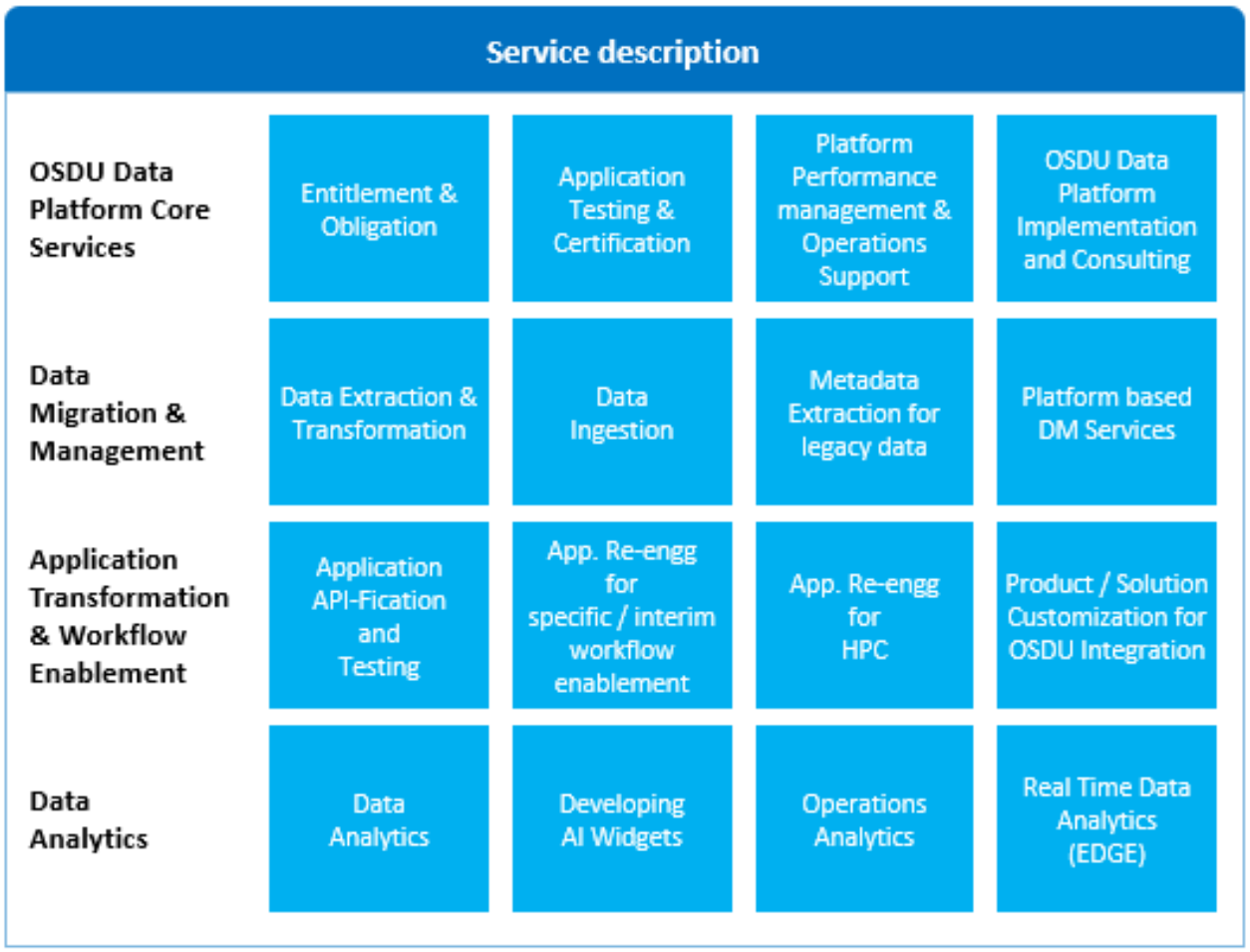


Figure 2. Different OSDU services and their description

LTI's Contribution to OSDU (Open-source Contribution)	
Certification	• Automation for OSDU Schema Validation • OSDU Application Certification Policy guideline documentation
Software Testing Suite	• Writing automation scripts for OSDU API Testing using Postman / Newman since Mercury Release -Milestone 10
Audit & Metrics AND Reporting & Dashboarding	• Developing a solution for OSDU Data Platform Operations area using LTI's CloudEnsure data governance platform. Solution work for 11 KPIs donated to the OSDU forum for GCP ,Mercury Release-Milestone#12 . Azure / AWS integration is in progress.
Entitlement and Obligations (E&O)	• Proposed architecture design for policy implementation • Primary contributor E&O Incubator project • Policy development using Open Policy Agent (OPA) • Enhancement work for OSDU Milestone # 8
Helicopter Resource Coordination	• Dashboard development using PowerBI • Helicopter schedule data ingestion into OSDU data platform

Figure 3. LTIMindtree contributions to different OSDU platforms

OSDU entitlement and obligation

We have been involved in designing the solution and writing code to deploy the OSDU policy manager in the Rego language. We have implemented the code for multiple use cases like economic sanctions, joint ventures, Production Sharing Contract (PSC), data-room visits, etc.

OSDU data policy is a user-friendly solution to enable authorization in the data ecosystem. The solution defines policies based on various legal entities and business rules via the user interface. These policies are created and customized based on various conditions like legislation, commercial, and governance and will be validated based on the entitlement tags stored for various data types. The policy engine returns an error message if the user request doesn't meet the policy rules/conditions.

Some highlights are:

- Open-source policy engine that unifies policy enforcement
- Implementation of policy as code
- Rego queries for assertions on data stored in Open Policy Agent (OPA).
- Access to data based on legal tags

LTIMindtree techno-domain team has been a primary contributor to the entitlement and obligation (E&O) ncubator project, where the team has been developing OPA.

OSDU audit and metrics

LTIMindtree generated various platform and data level KPIs for audit, metrics, and better governance.

The KPI metrics were generated in terms of the following:

- Data ingestion/search/delivery
- Data governance, platform performance, and traction
- Data volumes and usage
- Data quality metrics
- Utilize cloud provider agents for measuring cloud platform metrics
- Scheduler jobs for collecting various statistics

The techno-domain team has contributed towards finding the solution for the OSDU data platform operations area using LTIMindtree's CloudEnsure data governance platform. Solution work for 11 KPIs has been donated to the OSDU forum for GCP, Mercury Release-Milestone#12.

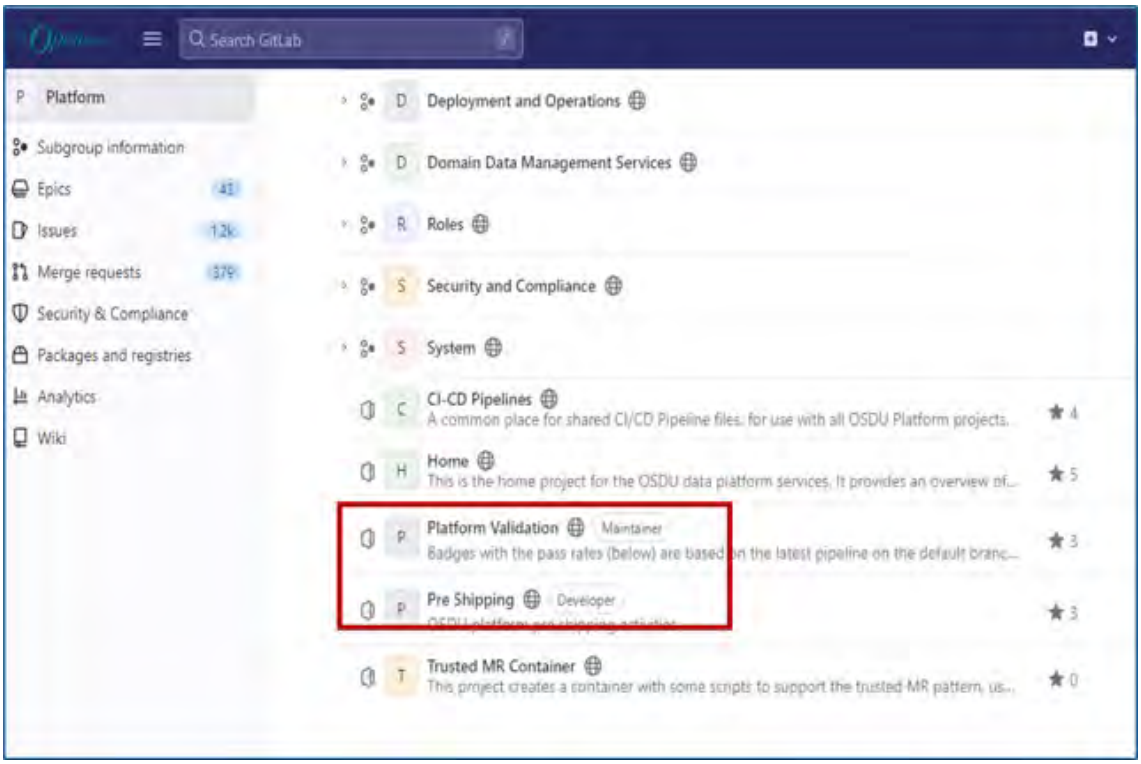
OSDU schema validation

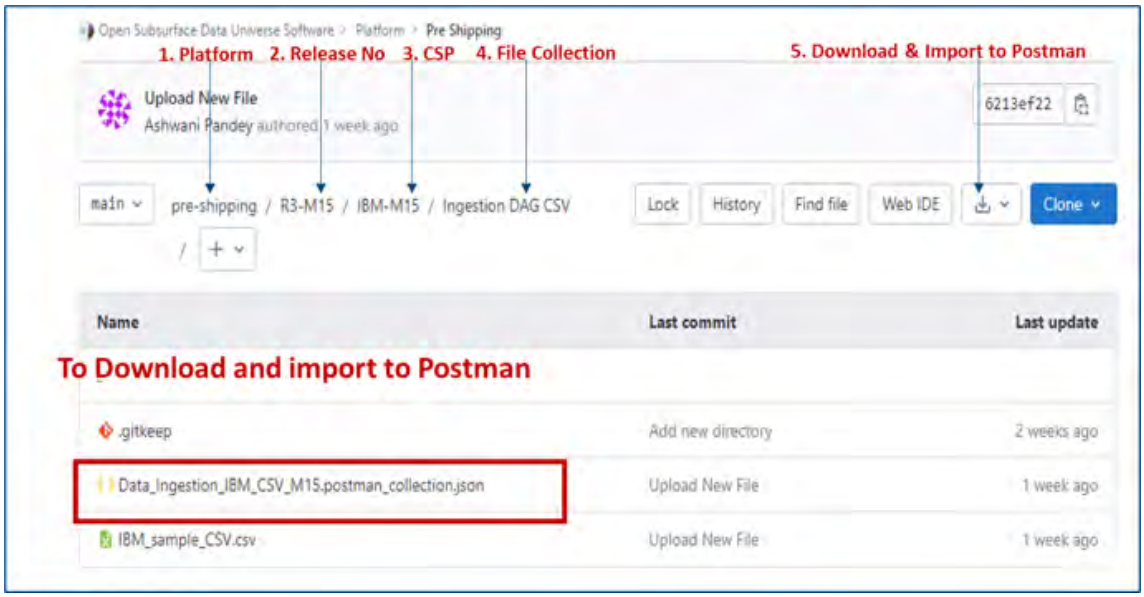
LTIMindtree team members have participated in different OSDU forums like platform validation, pre-shipping, data loading and ingestion, Edge, and Kafka platforms. The team has actively participated in testing different collections like Manifest and CSV ingestion, geospatial consumption zone testing, and ingestion by external reference. The team has further contributed towards testing API collections from R3-M10 to R3-M17 in Platform and Pre-shipping environments.

The team is also continuously working with the OSDU teams to understand the APIs developed by the OSDU forum and test the response before release. The testing of these APIs needs domain experience on multiple subsurface data types and formats that OSDU standard formats have recognized. The OSDU Project Management Committee (PMC) has recognized team efforts and leads.

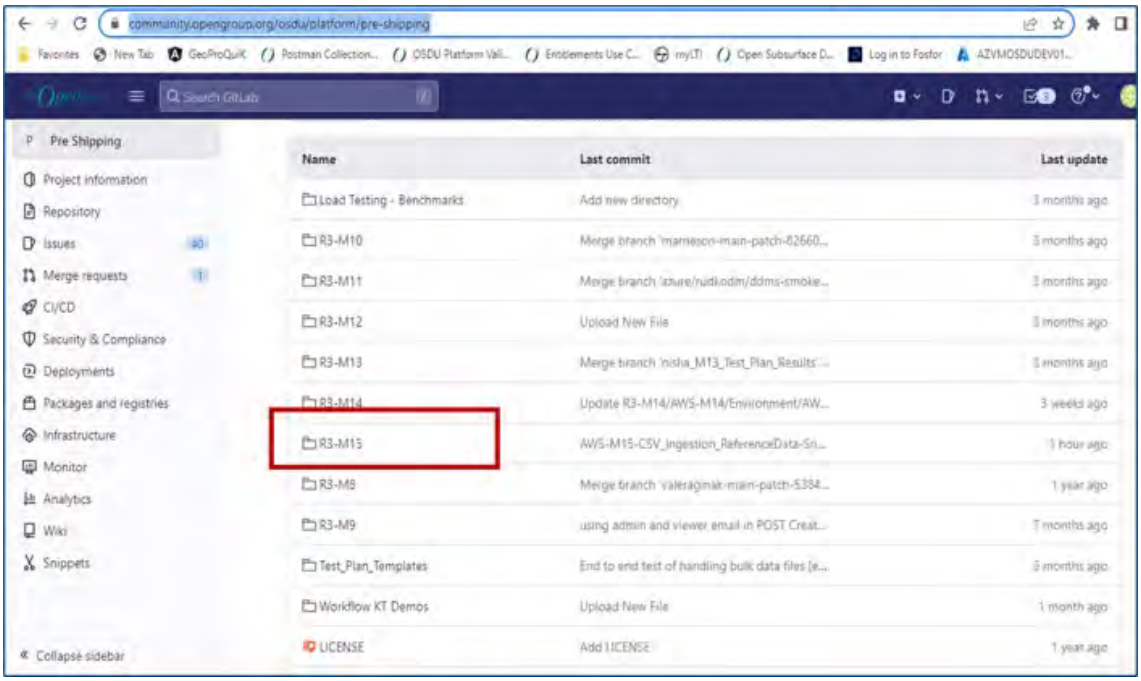
The testing workflow includes the following steps.

- Define a scope of testing activity like specific data type/workflow in OSDU.
- Define the test plan with use cases along with success/failure criteria.
- Putting the test cases in place for Postman Collection Structures
    - API and data type object-centric (ex., Legal API, Storage API, etc.)
    - Data workflow-centric (ex., Ingestion workflow)
    - Data Management Services (DMS) (ex., Seismic DMS, Wellbore DMS)
- Add/update testing cases in postman/pre-shipping collection
- Continuous Integration (CI)/Continuous Deployment (CD) pipeline
    - Generate CI/CD pipeline
    - Automatize from the collections (Newman)
    - Validate and execute testing cases
    - Report and publish testing results

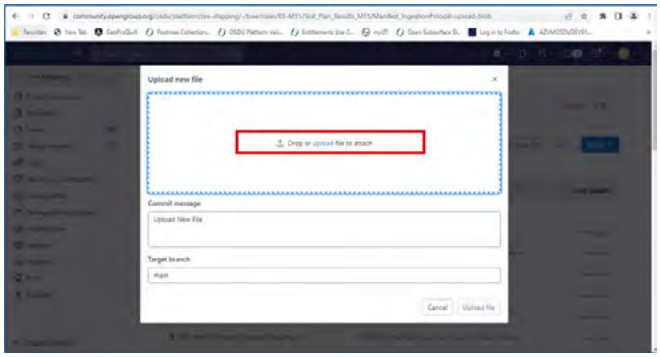


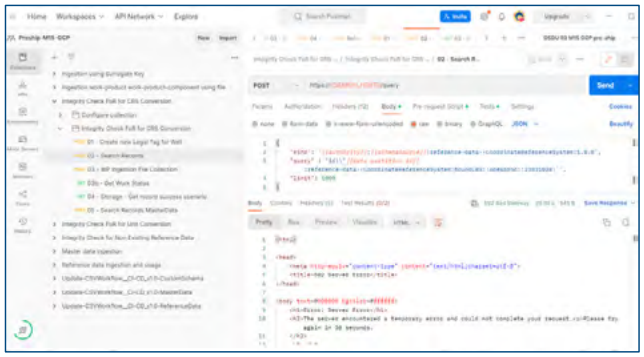


Select the folder with the release number you are testing (R3-M15)



Once the testing is done, the file is uploaded in the pre-shipping/platform environment at a defined path. The issue can be created in the defined path if an error appears while testing.





Business benefits

As mentioned above, the OSDU platforms will greatly benefit major O&G industry users including:

1) Technical and domain users and geoscientists
2) Business leaders and communities
3) Digitization communities
4) O&G companies, including third parties

The benefits of the OSDU data platform are profound, and some of them can be briefed as below:

> Helps in handling "data silos" between cross-domain teams that increases chances of creating a single platform to access exploration, production, and development data.
> Time-consuming data transfers can be handled in lesser time once a user has access to multiple geotechnical applications that can increase the speed of data migration.
> With the help of global data availability, faster and improved business strategies can be built up with more confidence and a greater success rate.
> Understanding of enhanced technological applications and software to meet requirements.
> Seamless access to the trusted data set.
> Helps save time searching data and focus on domain development and quick results.
> Increases chances of cross-domain collaboration.

So they need the OSDU so that seamless accessing, sharing, managing, and handling of data can be achieved with industry standards. The forum aims at helping operator companies with data residency requirements. Different data laws can be imposed on data residing in a country on the cloud or on-premise.

LTIMindtree team has been involved in addressing the challenges described above through their contribution to different platforms of the OSDU forum. Additional business benefits that can be highlighted in multiple domains have been listed below:

OSDU entitlement and obligation

The OSDU policy implementation has been a great help to OSDU data platform users, who have gained the following benefits:

- User-friendly interface to manage policies
- Policies to enforce entitlement to data
- Dedicated service for policy management
- Policies that define who can access what
- Open Policy Agent REGO for policy implementation
- Scalability and performance are determined as critical factors for evaluating the policy engine

OSDU audit and metrics

The technology stack used are Python, ReactJS, Flask, and MongoDB.

The solution helped clients in multiple aspects, as follows:

- Multi-cloud one governance, continuous monitoring, and metrics monitoring.
- Easy compliance adherence, seamless application integration, structured prioritization, and remediation.
- Use OSDU API for metadata access to generate various statistics.
- 11+ KPI metrics donated to OSDU for the platform and data governance.
- 100+ additional business rules available for KPI monitoring.
- Scalability and performance are determined as critical factors for evaluating the policy engine

OSDU schema validation

- Automated testing for search and delivery APIs
- Verified and validated R1 APIs as per the defined OSDU standard schemas
- Time saved

The solution built by the LTIMindtree team has helped users develop a dashboard that can save their time in automated validation of the API-generated schema.

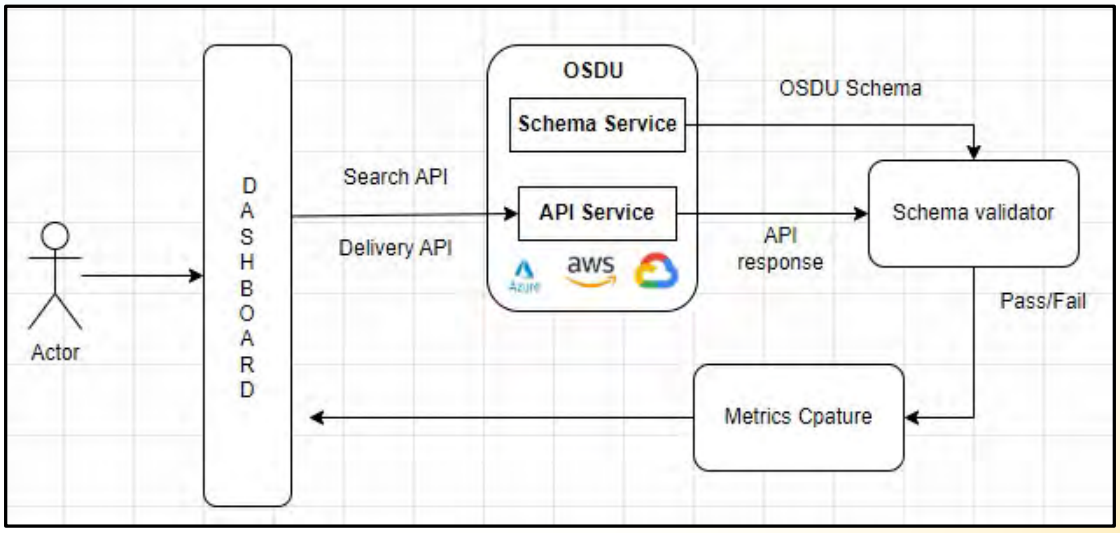
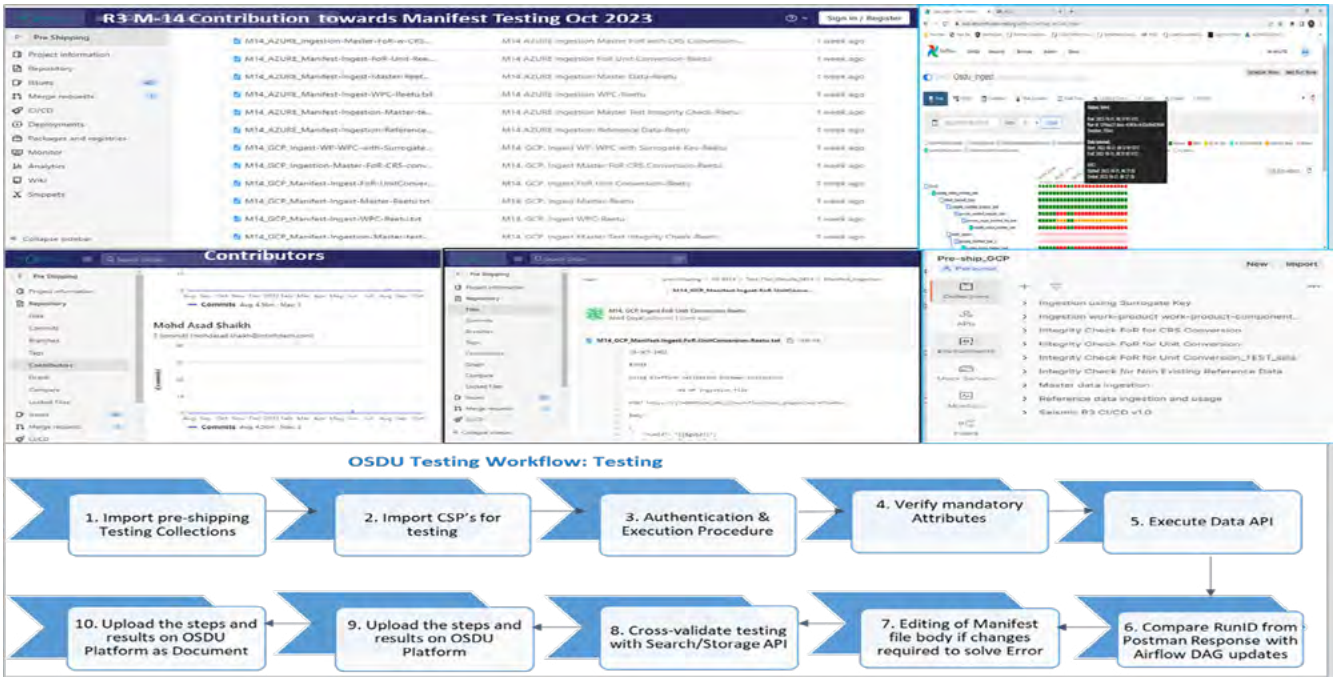


Figure 4. Above diagram shows the steps followed when a user makes a request that goes through search and delivery API to get response as pass or fail.

Software testing suite

Geologists, geophysicists, developers, and testers from different domains are working on the platform and pre-shipping environment for all Cloud Service Providers (CSPs). Testing seismic and well data API in GCP, IBM, AWS, and Azure has helped the forum understand the data type, different technologies, core services platform, and ingestion procedures. The technology landscape includes Python language, Rego language, Postman collections, and different CSPs. This has further helped the users in the following ways:

- OSDU Manifest Ingestion Testing for Platform and Pre-ship Testing teams
- Seismic API's Collection Testing using Sdutil for segy to zgy conversion workflow through sdutil libraries.
- Users can search, filter and add selected data to the cart.
- Handle SEG Y data migrating to different cloud environments
- Establish a platform-based approach to data storage and management
- Reduce human intervention through API testing
- Testing results help users in data search
- Proper documentation of folder structure helps users to find API endpoints and save time
- Data ingestion and migration to cloud environments through OSDU standard workflows help users to store their valuable datasets in a secure environment



Conclusion

This article highlights how LTIMindtree has used techno-domain skills to contribute to various OSDU forums. With different Mercury Release, the OSDU platform collaborates with different team members where well, seismic, real-time, and production data are considered during testing phases.

LTIMindtree also brings its highly differentiated capabilities and IP solutions to help customers deliver accelerated modernization of their entire subsurface area and innovate for value. We will continue leveraging our deep knowledge of the subsurface domain to deliver accelerated solutions and value to the OSDU Data Platform. For more information on this, please contact **Manesh.Parmar@lntinfotech.com**, **Shankar.Velappan@lntinfotech.com**, and **Reetu.Ragini@ltimindtree.com**

References

https://community.opengroup.org/osdu/platform/security-and-compliance/home/-/wikis/Design

https://www.openpolicyagent.org/

https://osdu.pages.opengroup.org/platform/domain-data-mgmt-services/seismic/open-vds/index.html

https://community.opengroup.org/osdu/platform/data-flow/ingestion/ingestion-workflow/-/issues/56

https://community.opengroup.org/osdu/documentation/-/wikis/OSDU-(C)/Architecture-Exploration-Topics/OSDU-R3-Architecture/Entitlements-Use-Cases

https://community.opengroup.org/osdu/platform/domain-data-mgmt-services/seismic/open-vds/-/blob/master/docs/connection.rst

Authors

REETU RAGINI

Reetu is a geoscientist with 15+years of experience in Upstream Oil & Gas industry. In her career across multiple O&G companies she has worked extensively in exploration and development projects. Her core competencies include seismic and well data interpretation, real-time data analysis, and subsurface data management. She has also worked as a part of Managed Services Team and successfully delivered multiple projects in application modernization, data migration and data management. She is highly skilled in G&G Petro-technical software. She is also a Certified Scrum Master.

SHANKAR VELAPPAN

Shankar has 20+ years of experience in the IT Industry in Data Analysis and Application development. He has experience working with major Oil & Gas companies in implementing solutions for Subsurface Data Management and Drilling applications. He is an industry expert in implementing solutions in areas of Data audit, migration, quality analysis, business rules creation & dashboarding. He has rich working experience in various G&G applications, Data Models and cloud-native platforms like OSDU.

LTIMindtree is a global technology consulting and digital solutions company that enables enterprises across industries to reimagine business models, accelerate innovation, and maximize growth by harnessing digital technologies. As a digital transformation partner to more than 700 clients, LTIMindtree brings extensive domain and technology expertise to help drive superior competitive differentiation, customer experiences, and business outcomes in a converging world. Powered by more than 84,000+ talented and entrepreneurial professionals across 33 countries, LTIMindtree — a Larsen & Toubro Group company — combines the industry-acclaimed strengths of erstwhile Larsen & Toubro Infotech and Mindtree in solving the most complex business challenges and delivering transformation at scale. For more information, please visit www.ltimindtree.com.