

Infrastructure Management

Transform the way you manage data for unmatched performance.





Table of Contents

1. Introduction

- 1.1 Empowering businesses for data-driven excellence 02

2. Infrastructure Mangement

- 2.1 Compute Farm 02
- 2.2 Data Lake 06
- 2.3 Data Cache 09
- 2.4 Notebook Configuration 11

3. Benefits

13

4. Conclusion

14



I. Introduction



1.1 Empowering businesses for data-driven excellence

In today's rapidly advancing technological landscape, organizations heavily rely on their technical infrastructure to support their operations and achieve their goals. In ConverSight, Infrastructure Management plays a critical role in overseeing and optimizing the resources and components that form the backbone of an organization's technical infrastructure.

Infrastructure Management encompasses essential components critical to organizational success. The Compute Farm facilitates the efficient creation of clusters and enables distributed computing. The Data Lake serves as a centralized repository for both structured and unstructured data. The Data Cache optimizes system performance through high-speed data storage. Furthermore, effective configuration and management of Notebooks enable collaborative, reproducible data analysis and model development. By proficiently managing these infrastructure aspects, organizations can enhance performance, achieve greater scalability and improve operational efficiency.

Through exploration of the features and benefits of each component, users gain a clear understanding of their significance in establishing a resilient and efficient infrastructure environment. This knowledge empowers users to make informed decisions and maximize the potential of their infrastructure. They also gain insights into how these components synergize to create a robust and well-optimized infrastructure environment.

2. Infrastructure Management

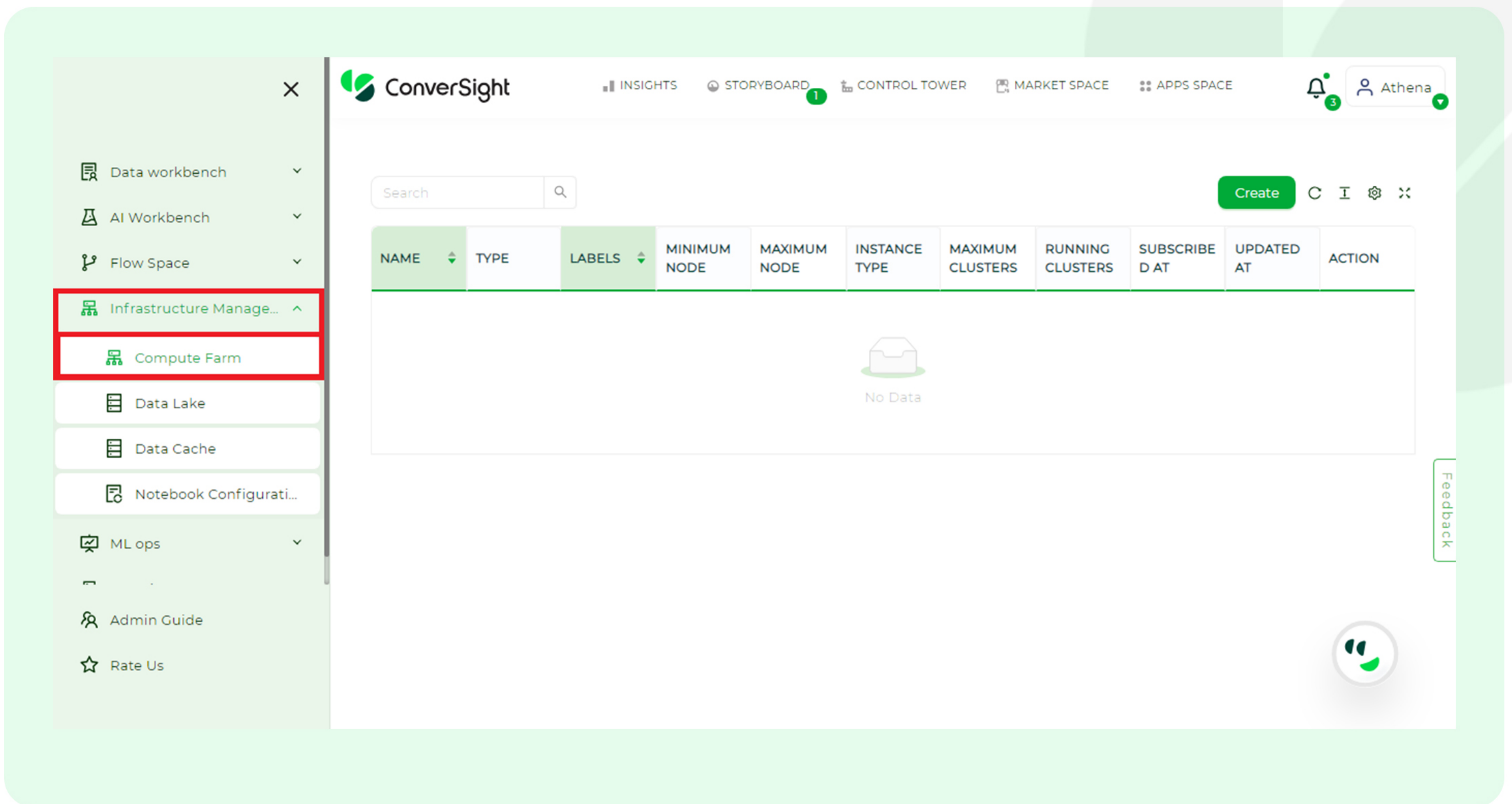


2.1 Compute Farm

The Compute Farm within ConverSight facilitates cluster creation and management. A cluster represents a collection of nodes that collaborate to execute flows, here each node can be a physical machine or a virtual instance in a cloud environment. These clusters can have a fixed size or dynamically scale based on resource demands. Leveraging clusters offers advantages such as high performance, scalability, flexibility and simplified management.

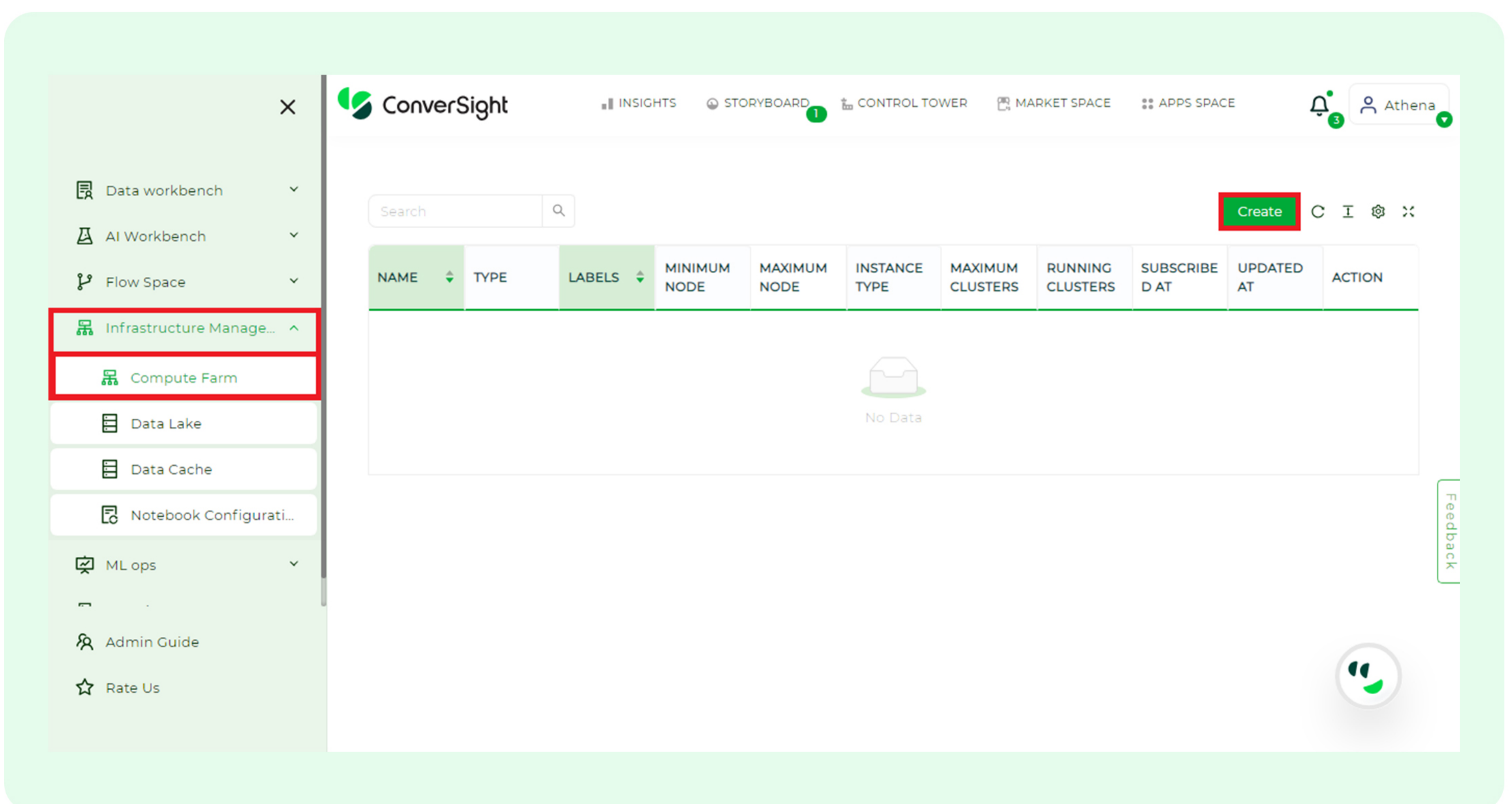
Accessing Compute Farm

To access the Compute Farm in the ConverSight Platform, navigate to the configuration section and select **'Compute Farm'** under **'Infrastructure Management'**.



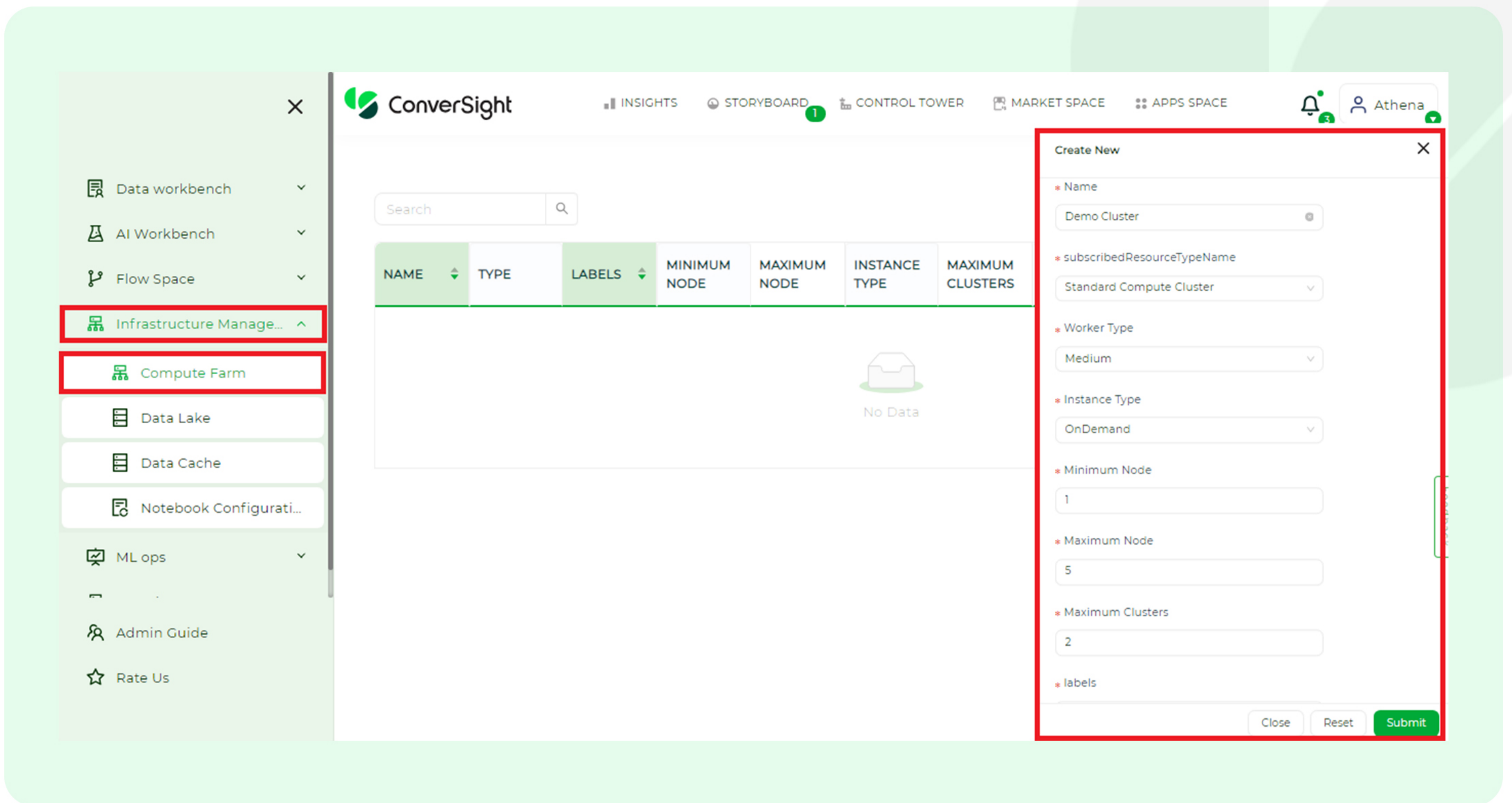
Creating a Cluster

Under the Compute Farm feature, users can create a cluster and manage clusters. This functionality allows users to group together multiple computing resources, such as servers or virtual machines to form a unified system. During the cluster creation process, users have the ability to define specific cluster specifications. These specifications include parameters such as the number of nodes or machines in the cluster. By specifying these cluster specifications, users can customize the cluster to meet their unique requirements and workload demands. This level of customization enables users to optimize the performance, scalability and resource utilization of the cluster.



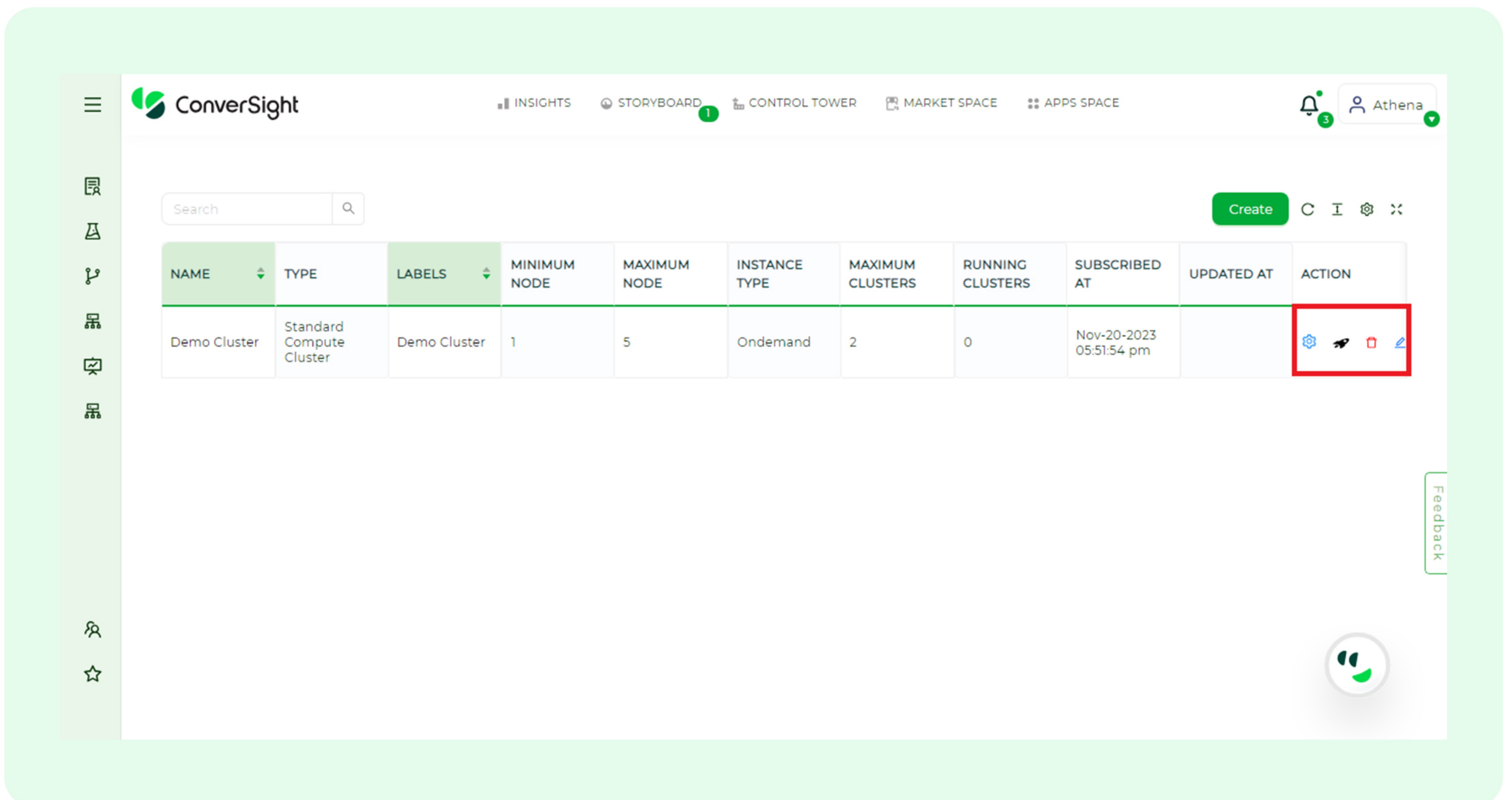
Upon clicking the **'Create'** option, the following screen will be displayed. Users can input the necessary details as shown in the table below. Finally click on **'Submit'** to create the cluster.

SPECIFICATION	DESCRIPTION
Name	The Cluster name can contain alphabets and numbers with length ranging from 3 to 32 characters.
Subscribed Resource Type Name	Users can choose either the Standard Compute Cluster or GPU Compute Cluster from the dropdown.
Instance Type	Users have the flexibility to choose between On-Demand instances or Spot instances based on specific needs.
Minimum and Maximum Nodes	Minimum and maximum number of nodes required.
Maximum Clusters	Choose the number of Clusters based on size and complexity.
Labels	Users can add labels as per their requirement.
Subscription Level	Users can choose among the three subscription levels. Dataset Cluster – Only used for Dataset. Organization Cluster – Used by the entire organization. User Cluster – Can be used only by a specific user.
Enable Cluster Culling	When the Cluster Culling toggle is activated, users can specify a particular time for the automatic termination of a cluster.
Allow Athena to Start Cluster on Need	If this toggle is enabled, Athena will start a Cluster with the chosen subscription level.



Cluster Management

In the ConverSight platform, users have the flexibility to take control of Clusters by initiating them manually or allowing Athena to start a Cluster automatically based on needs. Furthermore, users can customize the specifications of existing Clusters, monitor Cluster management activities and remove Clusters when required. These robust functionalities empower users to efficiently manage Clusters and seamlessly perform desired operations within the platform. With ConverSight, users have the tools they need to optimize their Cluster management experience.



- ⚙️ **'Manage Clusters'** provides access to a detailed view of both terminated and currently running Clusters. The Clusters are displayed in a tabular format, showcasing information such as the Cluster name, head node, initiation time, start time, status and available actions.
- 🚀 You can run a Cluster in the ConverSight platform by clicking on the Rocket icon, when you hover over the rocket icon users can view the popup that says, **'Start a Cluster'**.
- ✏️ After creating a cluster in ConverSight, users have the flexibility to update or edit the cluster specifications at any time. To accomplish this, simply click on the **'Edit'** icon in the actions menu and make the necessary adjustments according to your specific needs.
- 🗑️ The **'Delete'** icon serves the purpose of terminating a running Cluster. To delete a Cluster permanently, users must terminate all running clusters and then use the Delete icon on the Compute Farm page which displays all Clusters.

2.2 Data Lake

Data Lake represents a centralized repository meticulously designed for the purpose of storing, processing and ensuring the security of substantial volumes of structured and unstructured data. It possesses the capability to retain data in its raw format and seamlessly process it, irrespective of its size, offering a robust and efficient solution for data management and analysis.

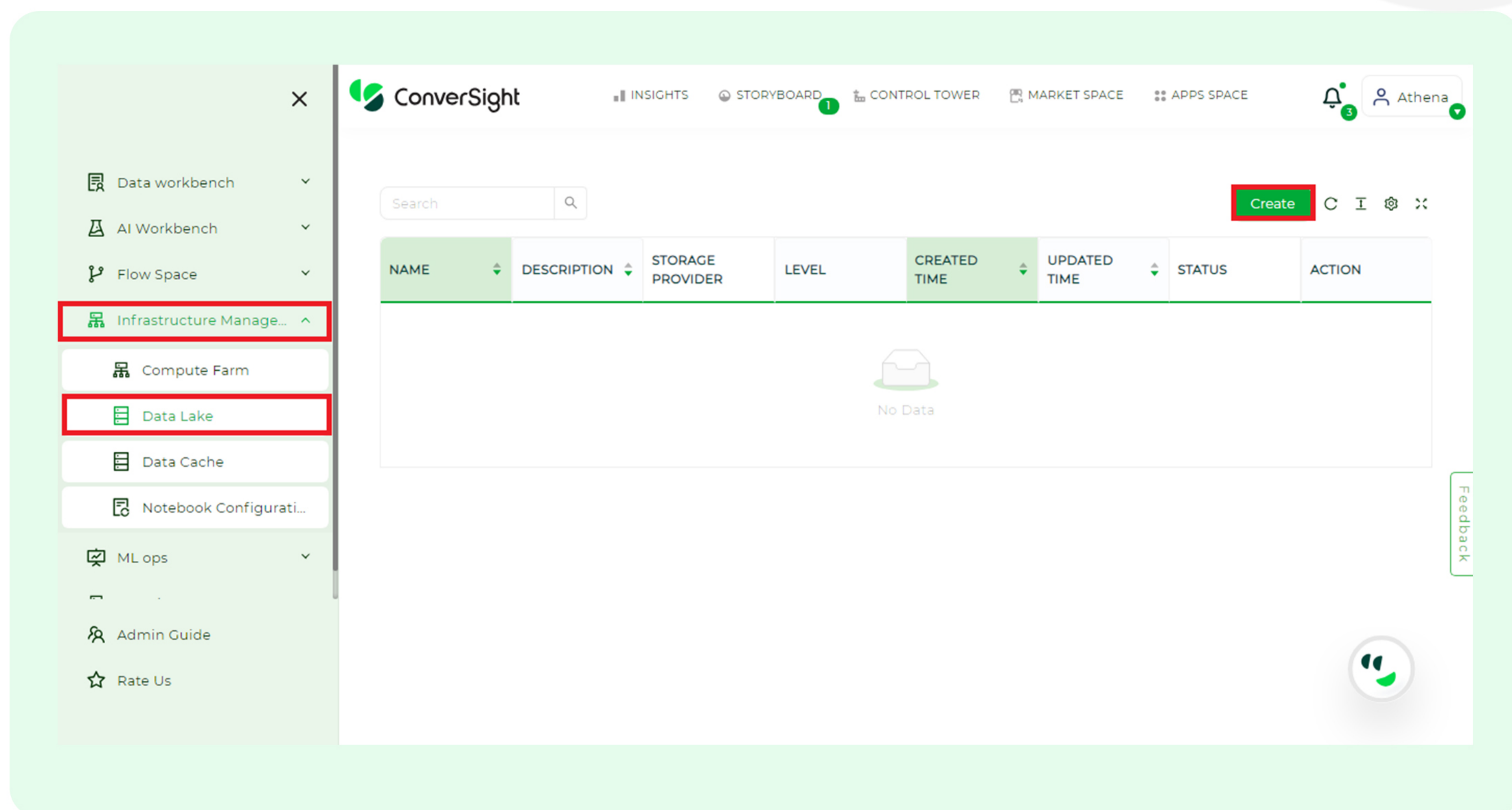
Accessing Data Lake

To access the Data Lake in the ConverSight Platform, navigate to the configuration section and select **'Data Lake'** under **'Infrastructure Management'**.

The screenshot displays the ConverSight platform interface. On the left, a navigation sidebar is visible with the following items: Data workbench, AI Workbench, Flow Space, Infrastructure Manage... (highlighted with a red box), Compute Farm, Data Lake (highlighted with a red box), Data Cache, Notebook Configurati..., ML ops, Admin Guide, and Rate Us. The main content area shows the 'Data Lake' configuration page. At the top, there is a search bar and a 'Create' button. Below this is a table with the following columns: NAME, DESCRIPTION, STORAGE PROVIDER, LEVEL, CREATED TIME, UPDATED TIME, STATUS, and ACTION. The table is currently empty, displaying 'No Data' in the center. The top navigation bar includes 'INSIGHTS', 'STORYBOARD', 'CONTROL TOWER', 'MARKET SPACE', and 'APPS SPACE', along with a user profile for 'Athena' and a notification bell.

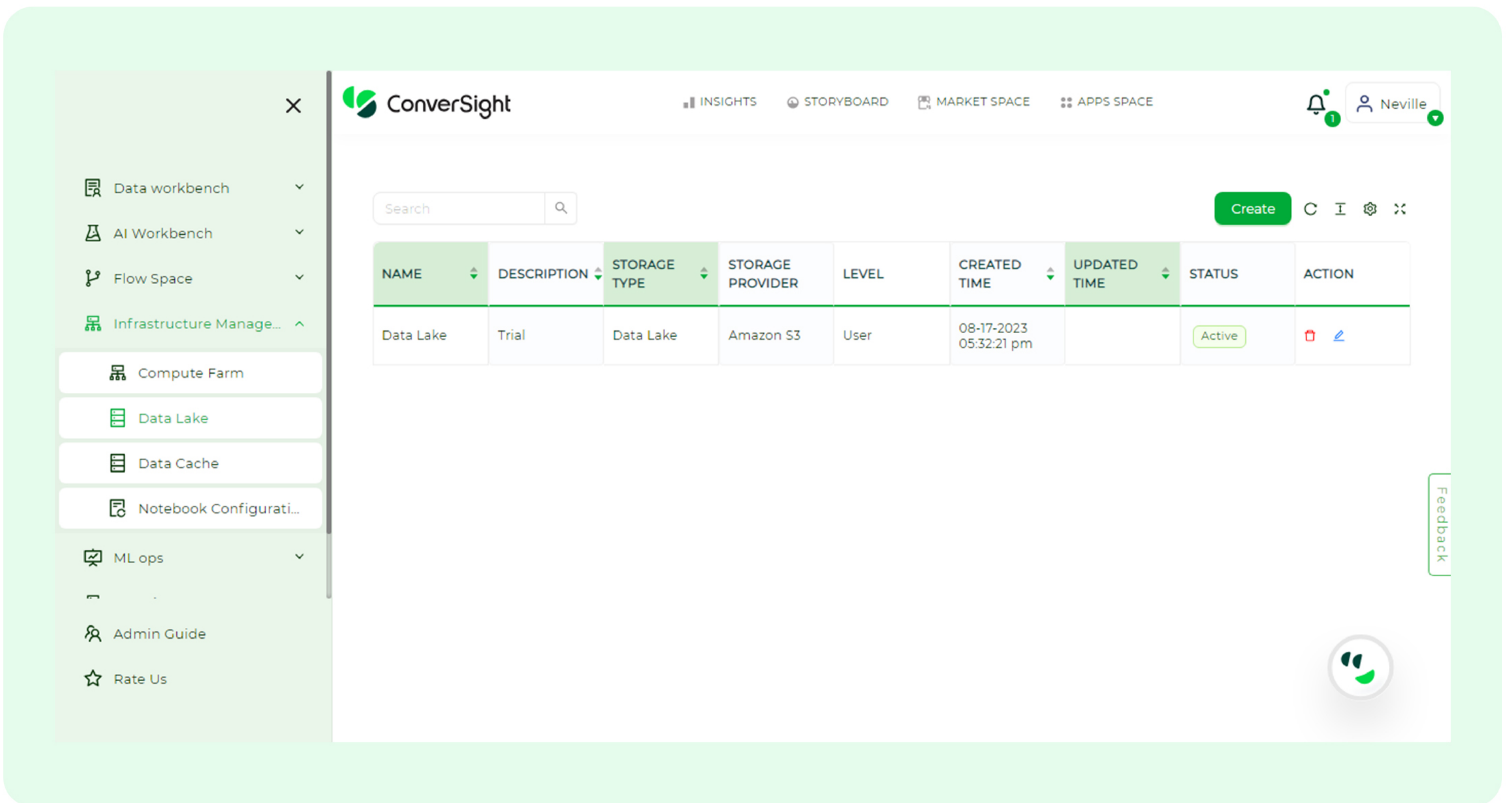
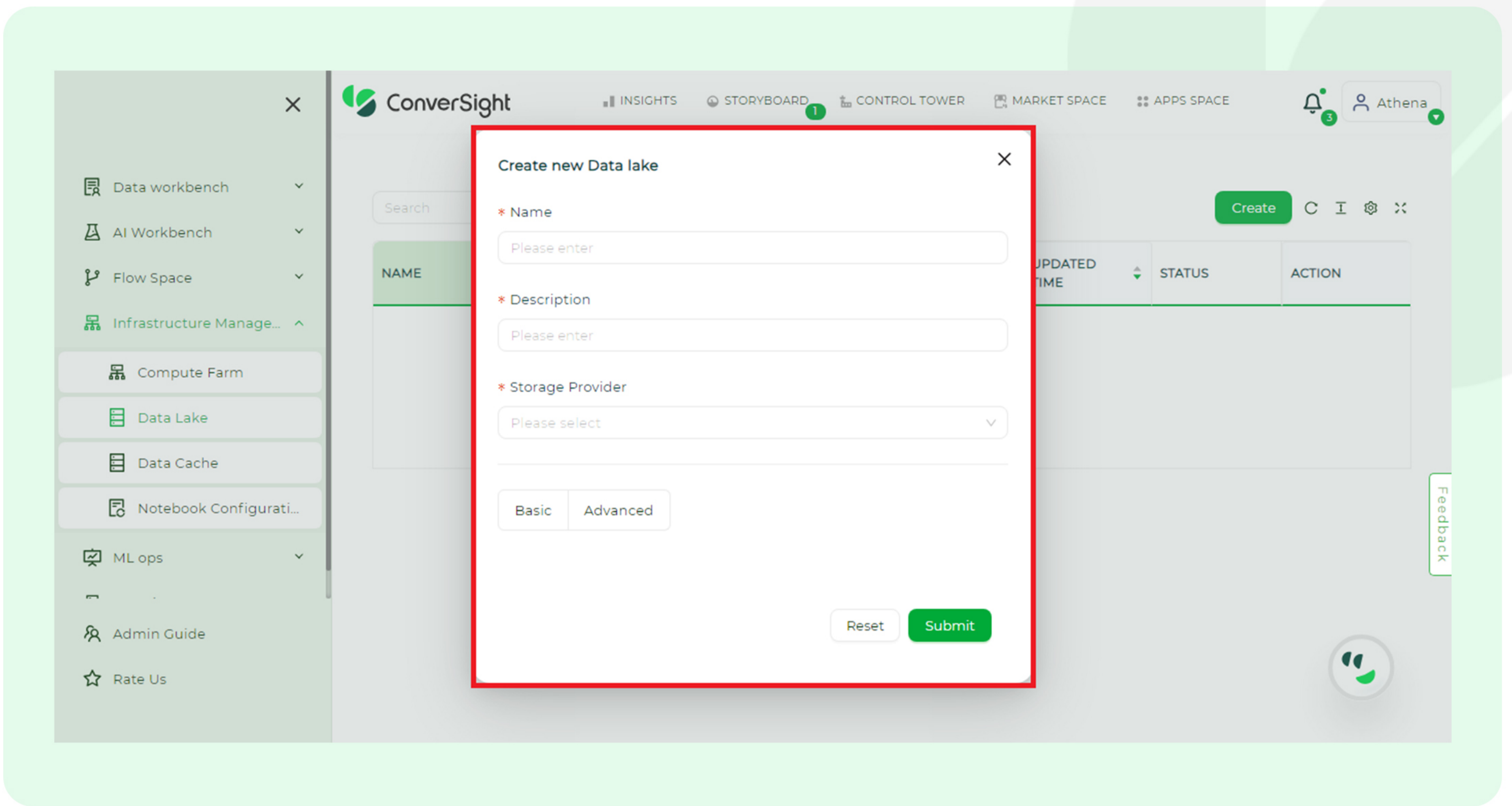
Creating a Data Lake

To create a Data Lake, simply click the **'Create'** button. This action enables the creation of a centralized repository for storing and managing large volumes of structured and unstructured data. Data Lake facilitates efficient data storage and retrieval, supporting advanced analytics and facilitating data-driven decision-making. This straightforward process allows users to establish a customized Data Lake tailored to their specific needs and requirements.



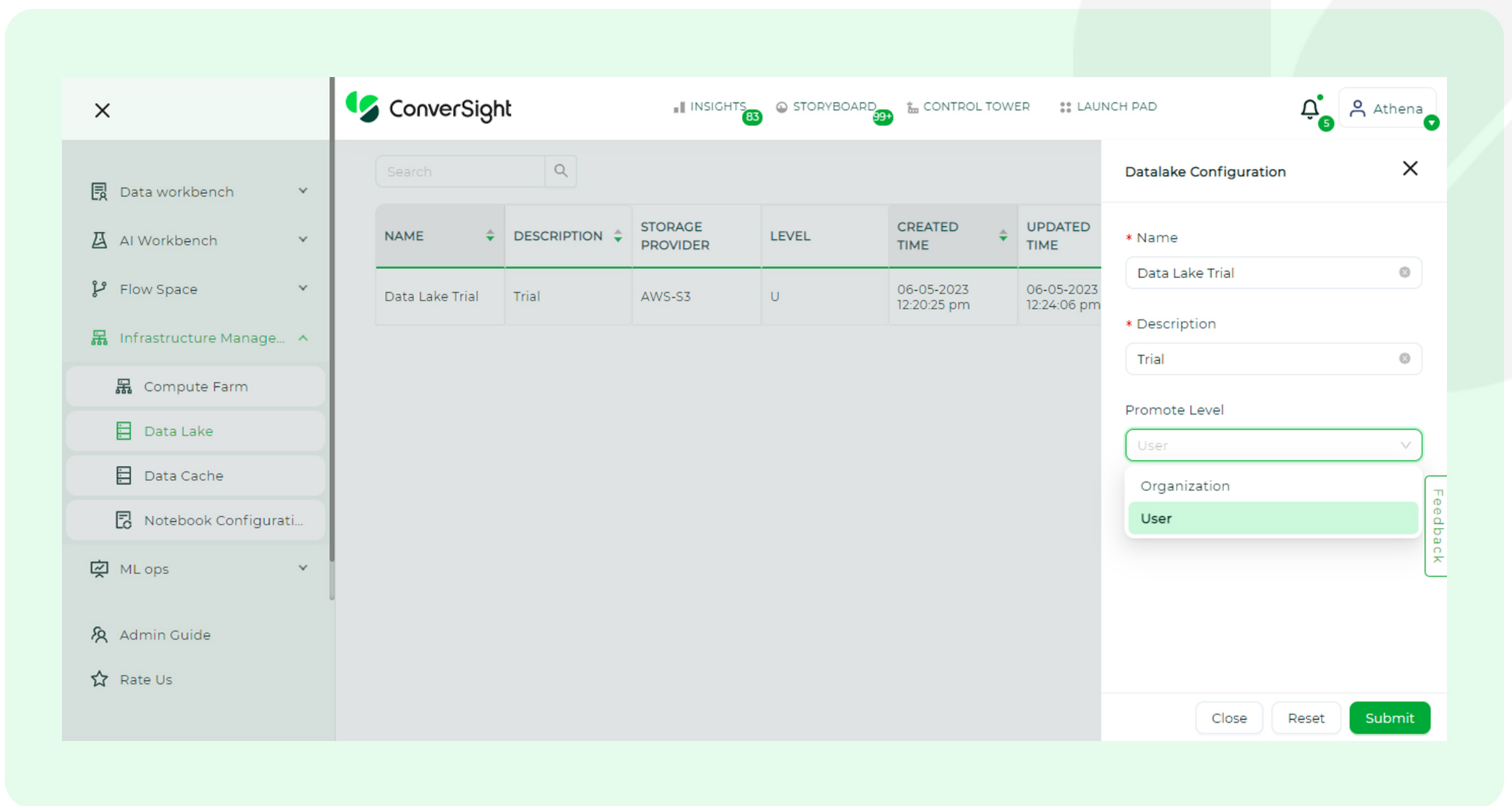
In the Create New page, you will need to input the following information to create a new Data Lake.

SPECIFICATION	DESCRIPTION
Name	Provide a name for the Data Lake.
Description	A Description provides additional information about the Data Lake for better understanding.
Storage Provider	Users must select the appropriate storage provider from the dropdown.
Basic / Advanced	Basic – Allows users to utilize database of ConverseSight. Advanced – Allows users to utilize custom databases.



Updating Data Lake

Upon the creation of a Data Lake, users are granted the opportunity to further customize and refine its configuration. This pivotal capability, as visually depicted in the accompanying image, enables users to precisely modify the Data Lake according to their specific requirements and promotional hierarchy levels like Organization level and User level.



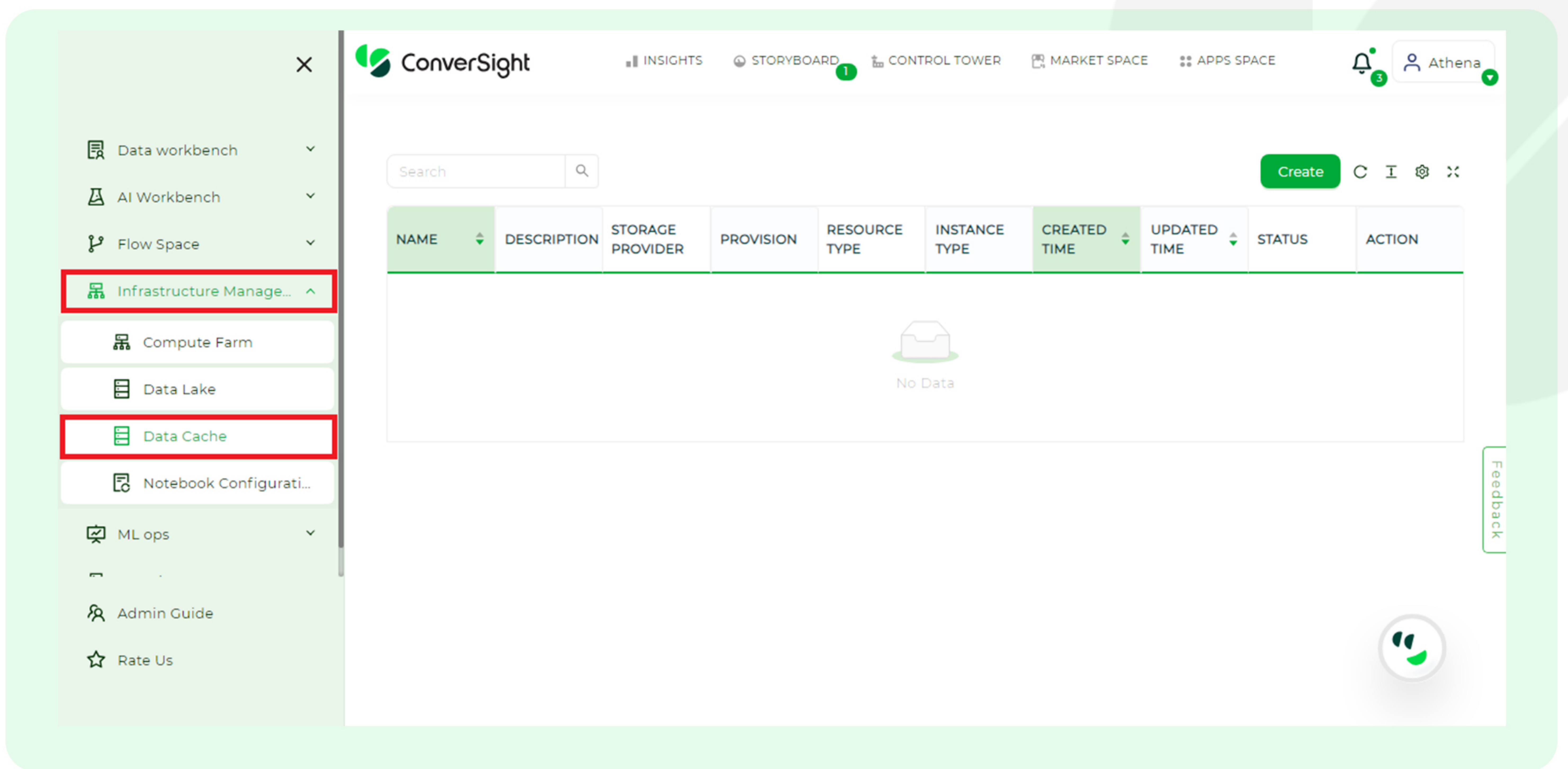
Data Lake can be deleted by clicking on the **'Delete'** icon associated with the respective Data Lake.

2.3 Data Cache

Data Cache is a memory space where data is stored in a specific database, allowing for easier and faster access when needed. This accelerates application performance and enhances overall efficiency.

Accessing Data Cache

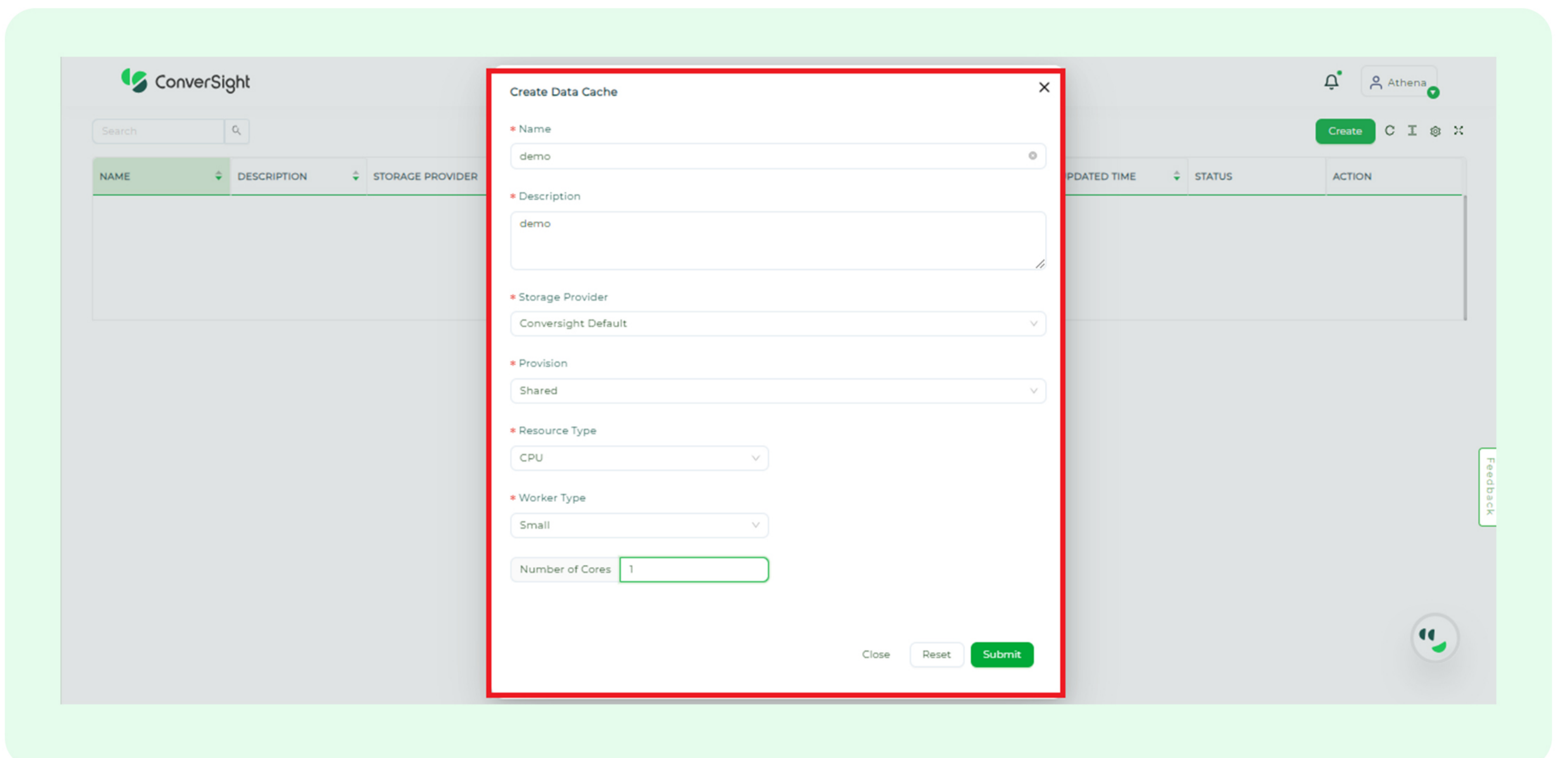
To access the Data Cache in the ConverseSight Platform, navigate to the configuration section and select **'Data Cache'** under **'Infrastructure Management'**.

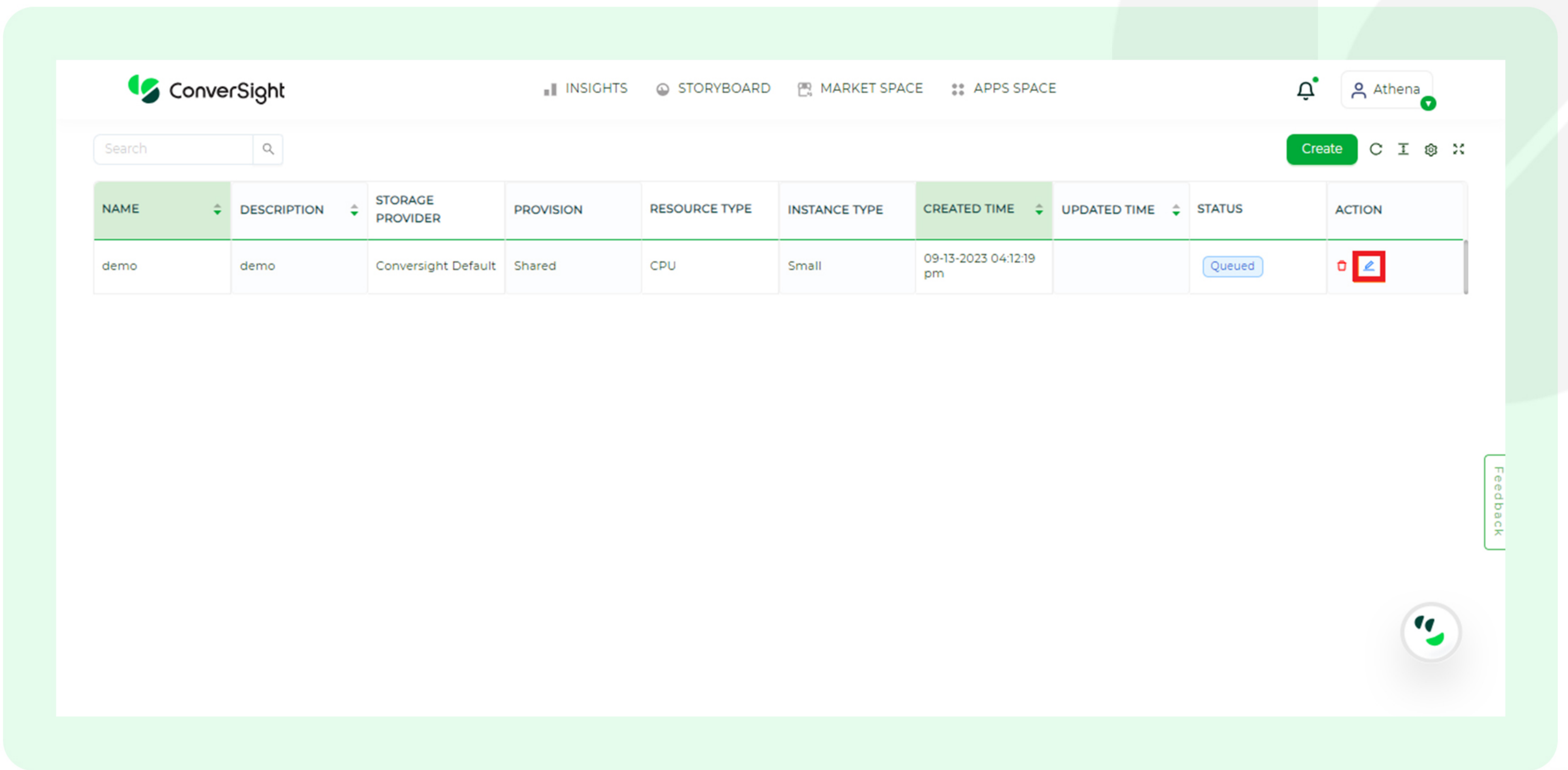


Creating Data Cache

To create a Data Cache, simply click the **'Create'** button. Users can specify the Name, Description, Storage provider, Provision, Resource type, Worker type and Quantity of cores. This systematic approach allows users to meticulously configure their Data Cache, ensuring optimal customization and precise control over caching operations.

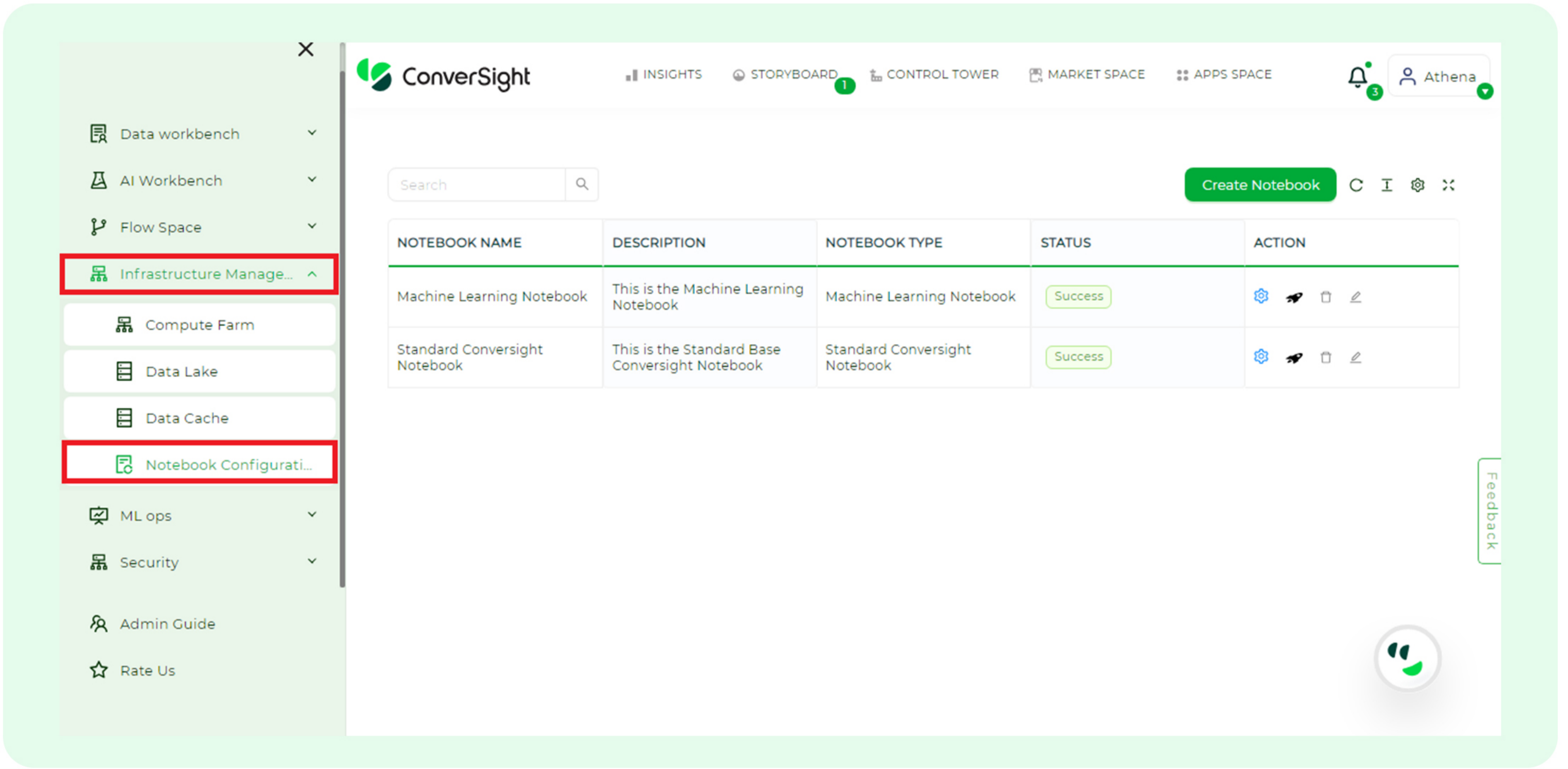
By default, the storage provider is **'ConverSight Default'**. Provision encompasses Dedicated User and Shared Users, depending on the users that can be altered to meet their specifications.



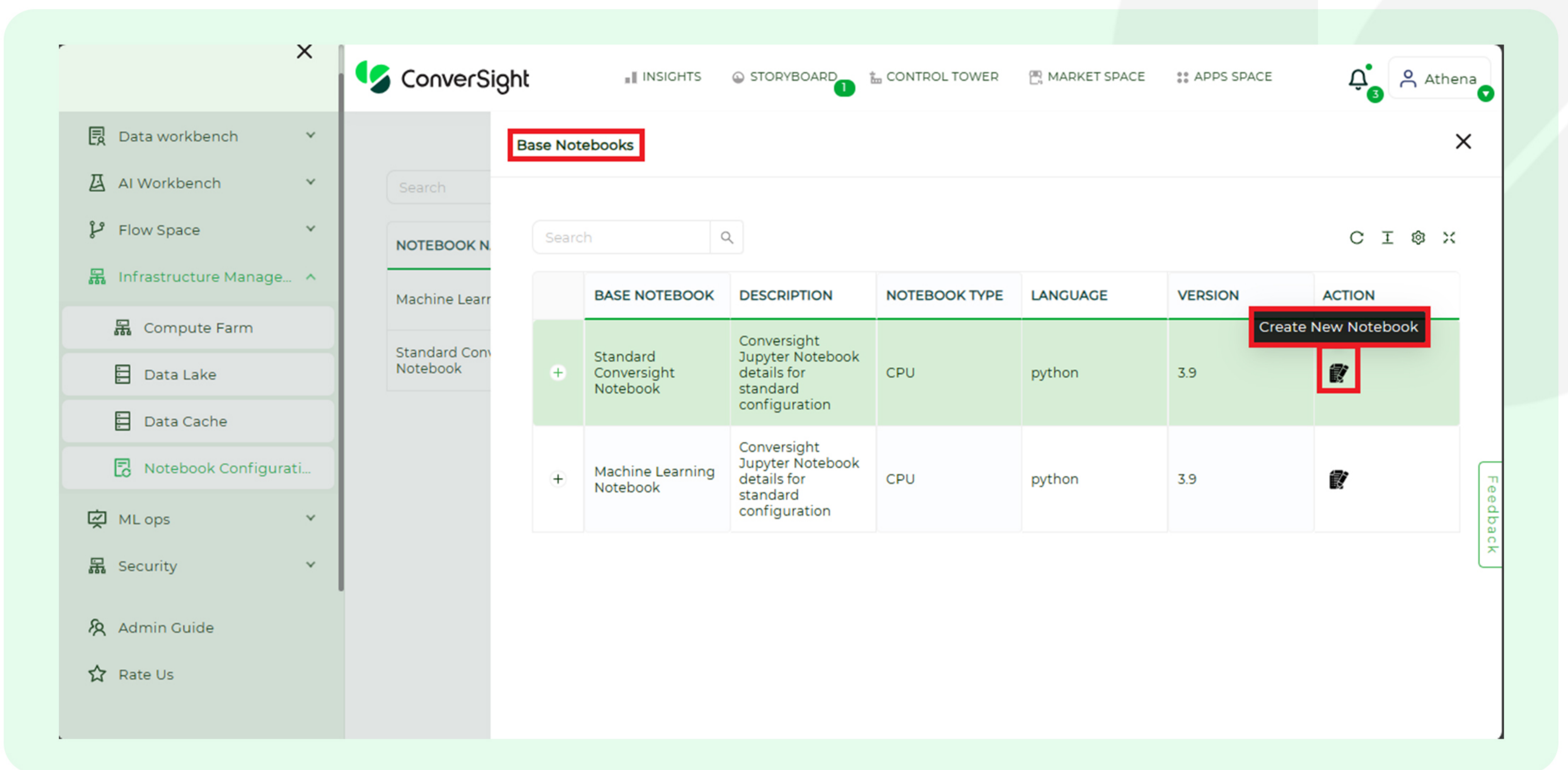


2.4 Notebook Configuration

To access the Notebook Configuration in the ConverSight Platform, navigate to the configuration section and select **'Notebook Configuration'** under **'Infrastructure Management'**.



To create a Notebook Configuration, simply click the **'Create Notebook'** button. Users can choose which Notebook Configuration they prefer and create new Notebook using the Edit icon in the Action column.



Upon clicking the **'Edit'** icon users can enter the necessary details as shown in the table.

ARGUMENT	DESCRIPTION
Name	Provide a name for the Notebook Configuration.
Description	A Description provides additional information about the Configuration for better understanding.
Python Version	Displays the version of Python used.
Notebook Type	Displays the type of Notebooks available.
Instance Type	Depending on the specific needs, users have the flexibility to choose between small, medium and large instances.
Verified Packages	Users can select all available packages from the dropdown.
Add Custom Packages	Users can add any custom packages in addition to the verified packages.

Finally click on **'Submit'** to create the new Notebook Configuration.

3. Benefits



4. Conclusion

ConverSight's Infrastructure Management empowers businesses to establish a robust and efficient technical infrastructure. By effectively managing components such as Compute Farm, Data Lake, Data Cache and Notebook Configuration, organizations can optimize performance, scalability and operational efficiency. Users gain a clear understanding of the significance of each component in creating a resilient infrastructure environment. With the ability to customize and refine configurations, users can make informed decisions and maximize the potential of their infrastructure. Whether it's cluster creation, centralized data storage, high-speed caching or collaborative data analysis, ConverSight's Infrastructure Management provides the tools and functionalities to support data-driven excellence.

Join our customers who have accelerated growth with ConverSight



About ConverSight

ConverSight's Adaptive Analytics platform uses conversational AI, Natural Language Processing and machine learning to converge the distance between humans and data through data stories, presenting the meaning of data in the most effective, personalized and efficient form possible. ConverSight's patented AI business assistant, Athena, connects distributed databases to answer questions and Augment the consumers through 4 key functions: Information on demand, Automated Story Telling, Proactive Insights, and Recommended Actions.

For more information, visit www.conversight.ai

