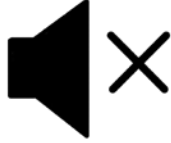


Sep 1, 2020

# EDC New Feature: Data Asset Analytics

Srinivas Pai

# Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available to view on our **INFASupport YouTube channel** and **Success Portal**. The link will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

# Feature Rich Success Portal

---



Bootstrap trial and  
POC Customers



Enriched Customer  
Onboarding  
experience



Product Learning  
Paths and Weekly  
Expert Sessions



Informatica  
Concierge with  
Chatbot integrations



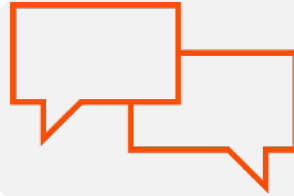
Tailored training and  
content  
recommendations

# More Information



## Success Portal

<https://success.informatica.com>



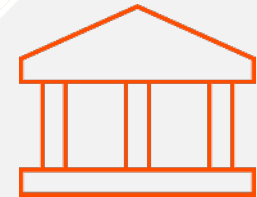
## Communities & Support

<https://network.informatica.com>



## Documentation

<https://docs.informatica.com>



## University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

# Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

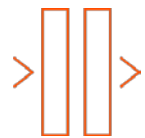
Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

# Agenda

- Overview
- Installation/Configuration/Administration
- Functionality (Chart Dashboard/Report)
- Demo
- FAQ/Troubleshooting
- Q&A

# Data Asset Analytics (DAA) Overview

Optimize data asset value with analytics on data asset inventory, usage, enrichment, and user collaboration in your data catalog



Measure



Analyze

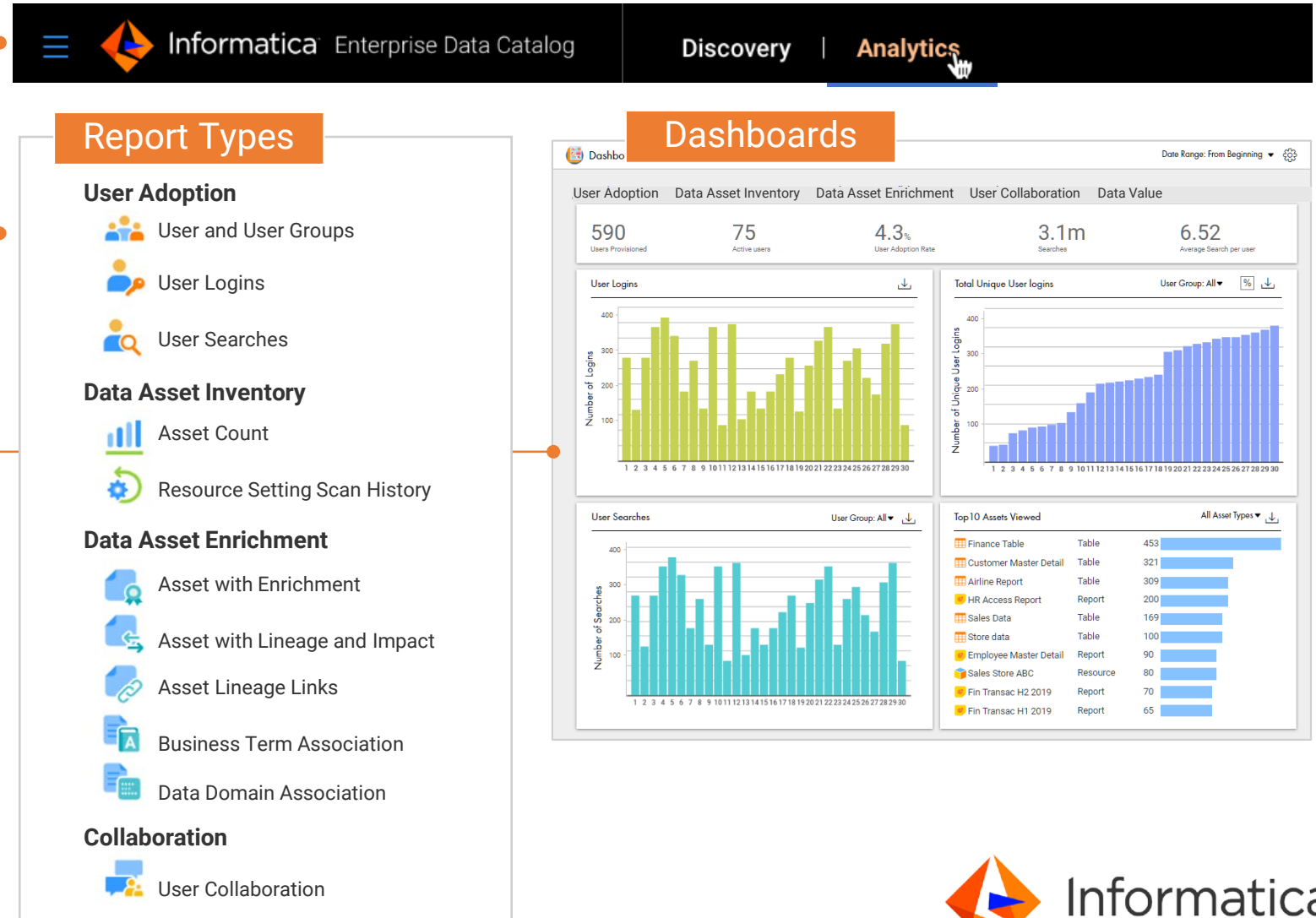


Optimize

# EDC: Data Asset Analytics

## Capability highlights

- Available with EDC 10.4.1 (Q2 2020)
  - Industry's First Data Asset Analytics
- Comprehensive Report Types
  - Predefined event (audit) history datasets
  - Preview and download event history data
  - Access event datasets directly from BI Tools
- Out-of-Box Data Asset Dashboards
  - Instant visibility to data asset usage, inventory, enrichment, collaboration, and data value
  - Near real-time key metrics and trends
  - Quick analysis with built-in filters
  - Storytelling for data assets





# Data Asset Analytics:

Visibility to data asset usage, inventory, enrichment, collaboration and data value



• Track Key Metrics



▶ Monitor Activities



▶ Undercover Patterns



▶ Spot Trends



▶ Improve Data Value



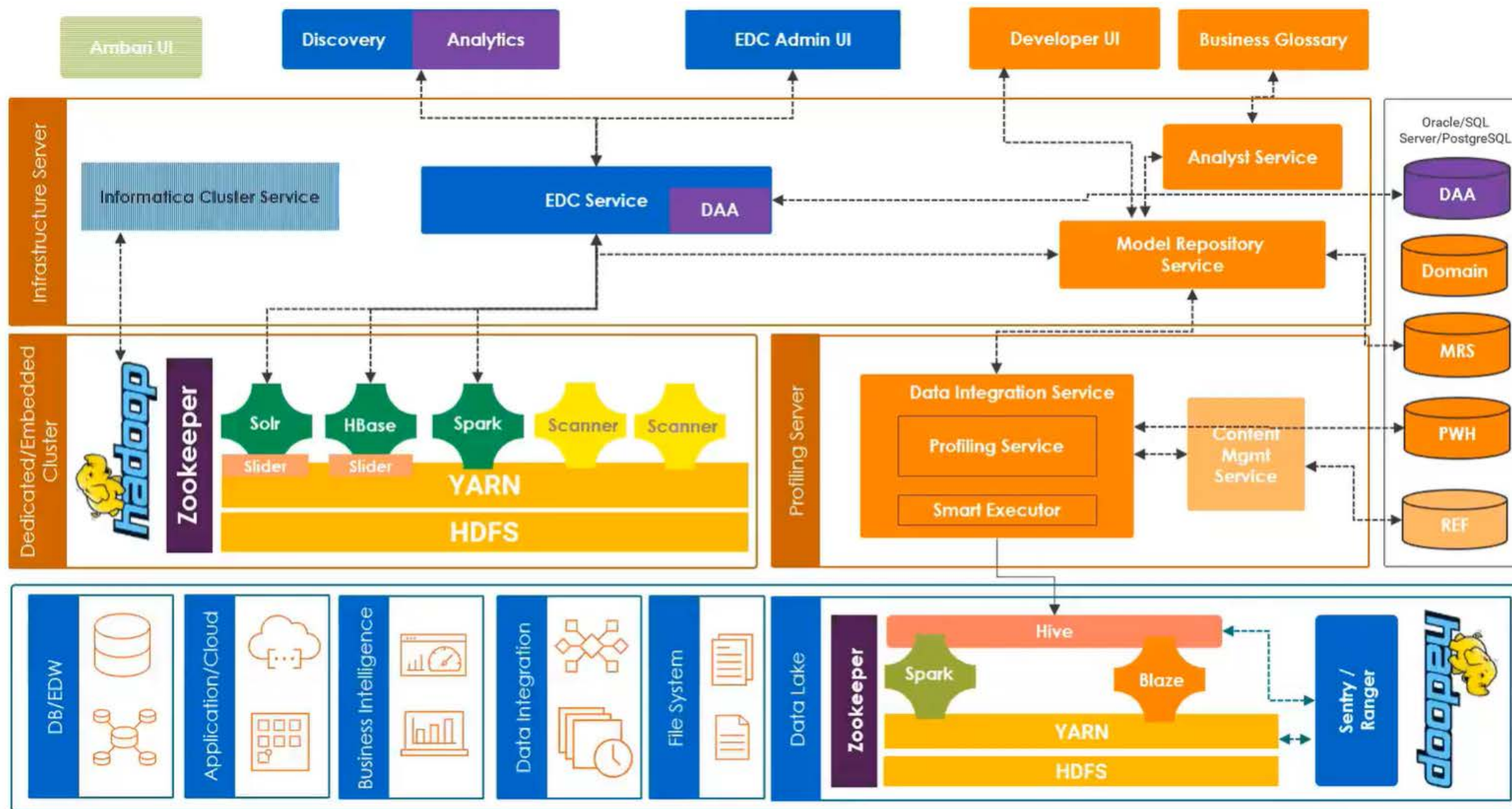
▶ Capture Mis-use



## Answers to Questions

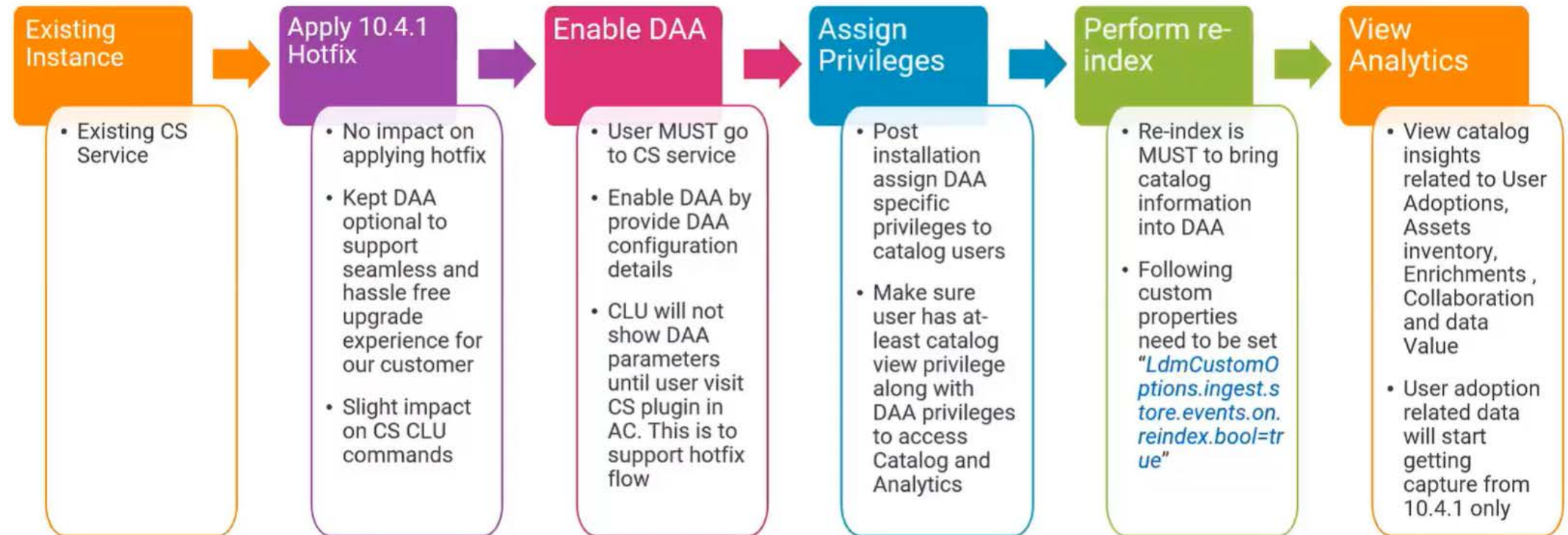
- Who is using, what asset, what feature, and when?
- What is the catalog adoption rate?
- What are the top searches, and average searches per user?
- What is the most access catalog module or assets?
- Who is doing DG and curating asset?
- What is % of objects with lineage, owner, description, bus. term?
- Who are the top contributors and SME inferred from collaboration activities?
- Which user groups, datasets need more attention and why?
- Who modified and scanned sensitive data sources?
- How to grow data asset value?

# High Level Architecture – Catalog Service with DAA



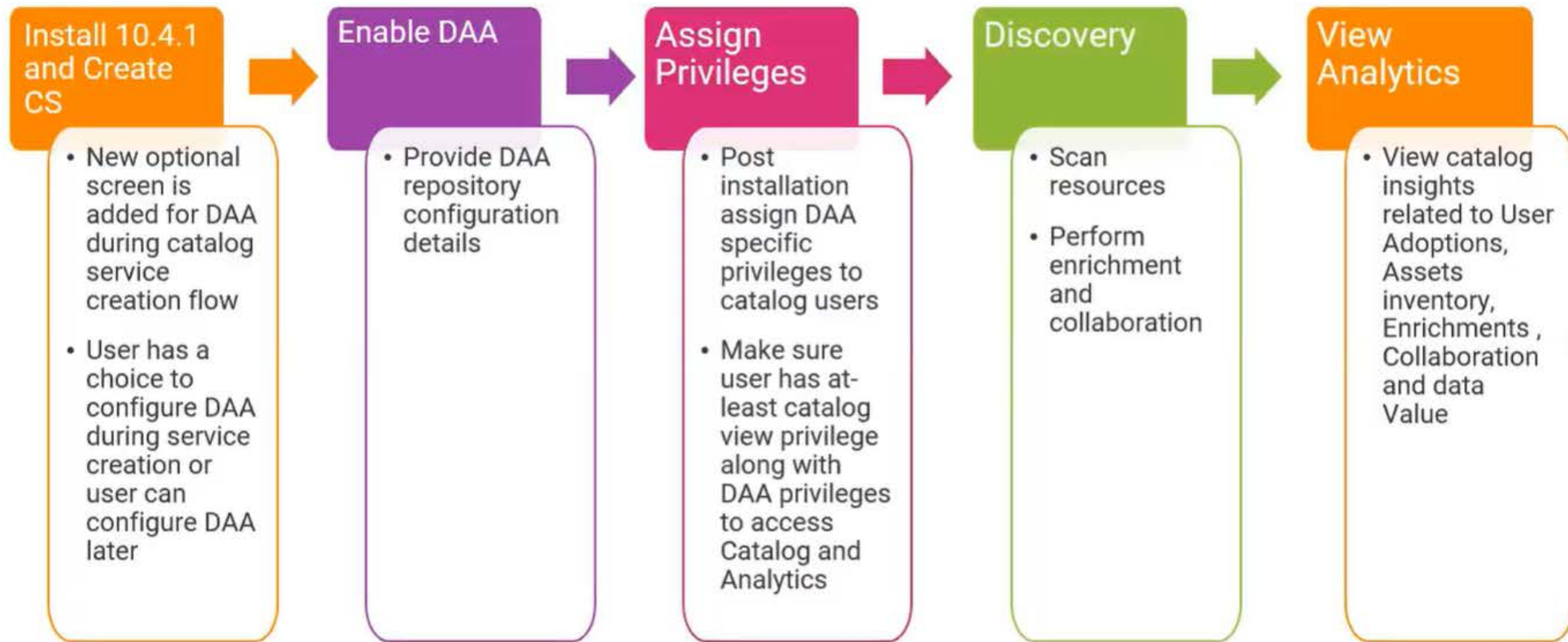
# Installation/Configuration

## On-Boarding 10.4.0 Customers to 10.4.1



# Installation/Configuration

On-Boarding New Customers to 10.4.1



# Configuration Parameters

**Edit Data Asset Analytics Configuration Properties** ✕

Fields marked with an asterisk (\*) are required.

☒ Enable Data Asset Analytics

Database Type to Select Database \*

ORACLE

Username \*

ORACLE

Password

SQLSERVER

POSTGRESQL

☐ Modify Password

Database Connection String \*

Secure JDBC Parameters

\*\*\*\*\*

☐ Modify Secure JDBC Parameters

Database Driver

com.informatica.jdbc.oracle.OracleDriver

?

OK

Cancel



# Administration

## Supported Databases and Privileges

Oracle	SQL Server	PostgreSQL
<ul style="list-style-type: none"><li>GRANT CONNECT TO &lt;DAA_SCHEMA&gt;</li><li>GRANT RESOURCE TO &lt;DAA_SCHEMA&gt;</li><li>GRANT CREATE VIEW TO &lt;DAA_SCHEMA&gt;</li><li>GRANT CREATE MATERIALIZED VIEW TO &lt;DAA_SCHEMA&gt;</li><li>GRANT UNLIMITED TABLESPACE TO &lt;DAA_SCHEMA&gt;</li></ul> or <ul style="list-style-type: none"><li>ALTER USER &lt;username&gt; QUOTA &lt;SIZE&gt; ON &lt;tablespace name&gt;</li></ul>	<ul style="list-style-type: none"><li>GRANT CREATE TABLE TO &lt;NEW USER NAME&gt;</li><li>GRANT CREATE VIEW TO &lt;NEW USER NAME&gt;</li><li>GRANT SELECT, ALTER, INSERT, DELETE, UPDATE ON SCHEMA::DBO TO &lt;NEW USER NAME&gt;</li></ul>	No explicit privileges are required

Command Line -- ./infacmd.sh ldm

- Create Service Option
- List Service Option
- Update Service Option

# Administration – Privileges

- Three new privileges are added into existing CS privileges

Privileges	What user can perform
Report Management : View and Download Raw Data	<ul style="list-style-type: none"><li>Will have access to Chart Dashboard</li><li>Should be able to download raw data from report dashboard</li><li>Should be able to download chart data from chart dashboard</li><li>Should have access to Monitoring tab</li><li>No access to Application Configuration setting page</li><li>Not allowed to edit value calculator</li></ul>
Dashboard Management : Visualize	<ul style="list-style-type: none"><li>Only View access to chart dashboard</li><li>No access to:<ul style="list-style-type: none"><li>Report dashboard</li><li>Monitoring dashboard</li><li>Application Configuration setting page</li><li>Cannot download chart or report data</li><li>Not allowed to edit value calculator</li></ul></li></ul>
Admin Management : Configuration	<ul style="list-style-type: none"><li>Can perform all actions in Analytics</li></ul>

# Application Configuration

## Application Settings

The screenshot displays the 'Application Configuration' interface within the Informatica Enterprise Data Catalog. The top navigation bar includes 'Informatica Enterprise Data Catalog', 'Discovery', and 'Analytics'. A left sidebar contains 'Dashboard', 'Reports', and 'Monitoring'. The main content area is titled 'Application Configuration' and has two tabs: 'Application' (selected) and 'Server Settings'. The 'Application' tab is divided into three sections: 'General', 'Data Value', and 'Degree of Collaboration'. The 'General' section includes 'View Analytics From' (set to 01/01/2010) and 'Reports Staging Location in the Server' (set to /home/ilmsatsqe/Informatica/10.4.1/server/info\_shared/reports/). The 'Data Value' section includes 'Value Calculator' (Min: 0, Max: 100). The 'Degree of Collaboration' section includes 'Low' (Min: 1, Max: 15), 'Medium' (Min: 16, Max: 75), and 'High (>)' (Min: 75). A 'Save' button is located in the top right corner of the configuration area.

Informatica Enterprise Data Catalog Discovery | Analytics

Dashboard Reports Monitoring

Application Configuration

Application Server Settings

**General**

View Analytics From: \* 01/01/2010

Reports Staging Location in the Server: \* /home/ilmsatsqe/Informatica/10.4.1/server/info\_shared/reports/

**Data Value**

Value Calculator: \* Min: 0 Max: 100

**Degree of Collaboration**

Low: \* Min: 1 Max: 15

Medium: \* Min: 16 Max: 75

High (>): \* Min: 75

Save Cancel



# Application Configuration

## Server Settings

The screenshot shows the Informatica Enterprise Data Catalog interface. The top navigation bar includes the Informatica logo, 'Enterprise Data Catalog', and tabs for 'Discovery' and 'Analytics'. The left sidebar contains 'Dashboard', 'Reports', and 'Monitoring'. The main content area is titled 'Application Configuration' and has two tabs: 'Application' and 'Server Settings' (which is active). The 'Server Settings' tab contains two sections: 'Event Load Settings' and 'Performance Cache Settings'. In the 'Event Load Settings' section, 'User Account Synchronization Frequency (in minutes):' is set to 1440, 'Enable Event Purge:' is checked, and 'Daily Event Purge Time:' is set to 00:00. In the 'Performance Cache Settings' section, 'Cache Size (MB):' is set to 20 and 'Clear Expire Interval (in milliseconds):' is set to 600000. A file path 'Documents - root@asjsridhar00502 - WinSCP' is visible in the background of the settings area. At the top right of the configuration area are 'Save' and 'Cancel' buttons.

Informatica Enterprise Data Catalog Discovery | Analytics

Dashboard Reports Monitoring

Application Configuration

Application Server Settings

**Event Load Settings**

User Account Synchronization Frequency (in minutes): \* ⓘ 1440

Enable Event Purge: \* ⓘ ☒

Daily Event Purge Time: \* ⓘ 00:00

**Performance Cache Settings**

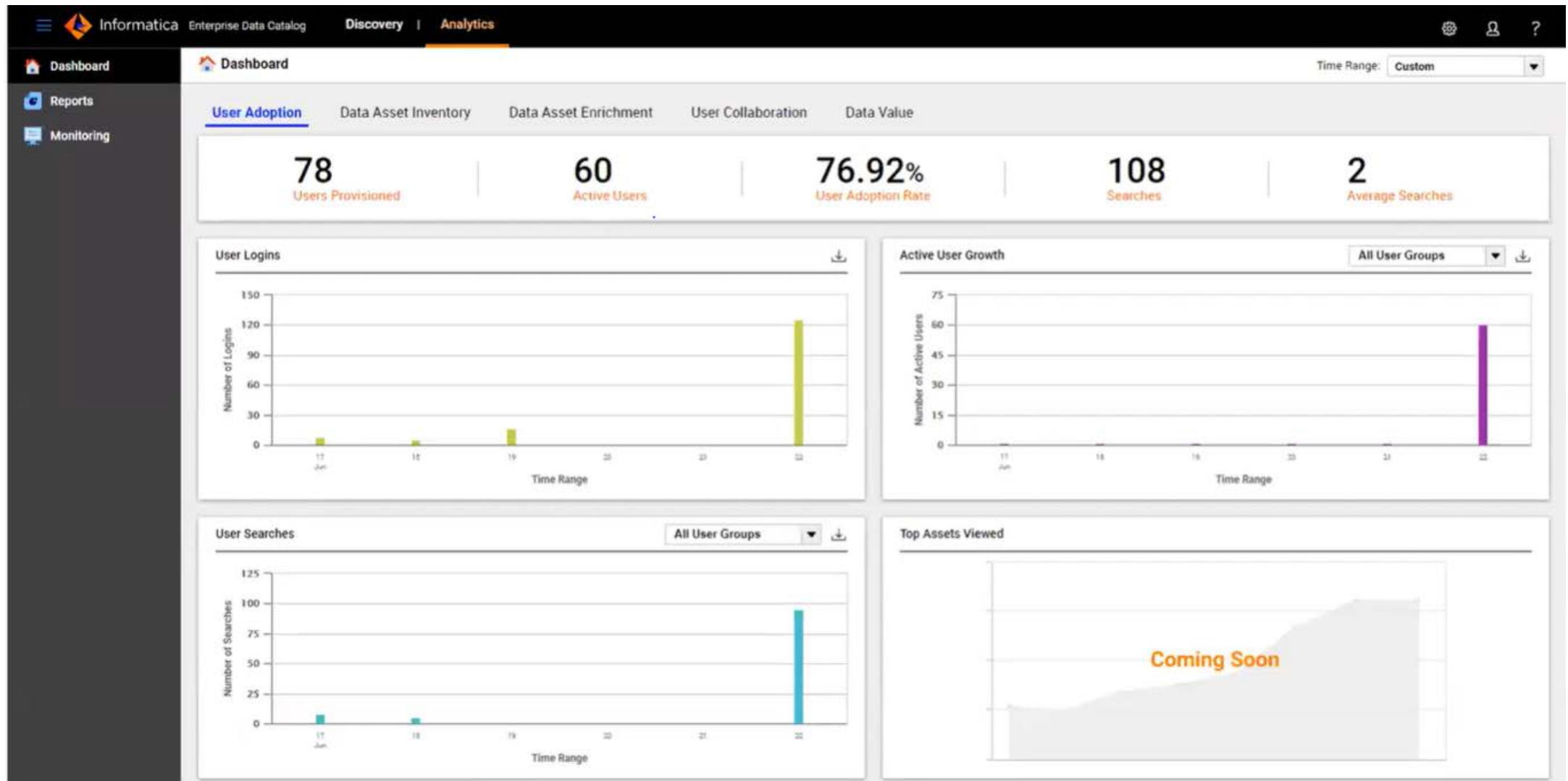
Cache Size (MB): \* ⓘ 20

Clear Expire Interval (in milliseconds): \* ⓘ 600000

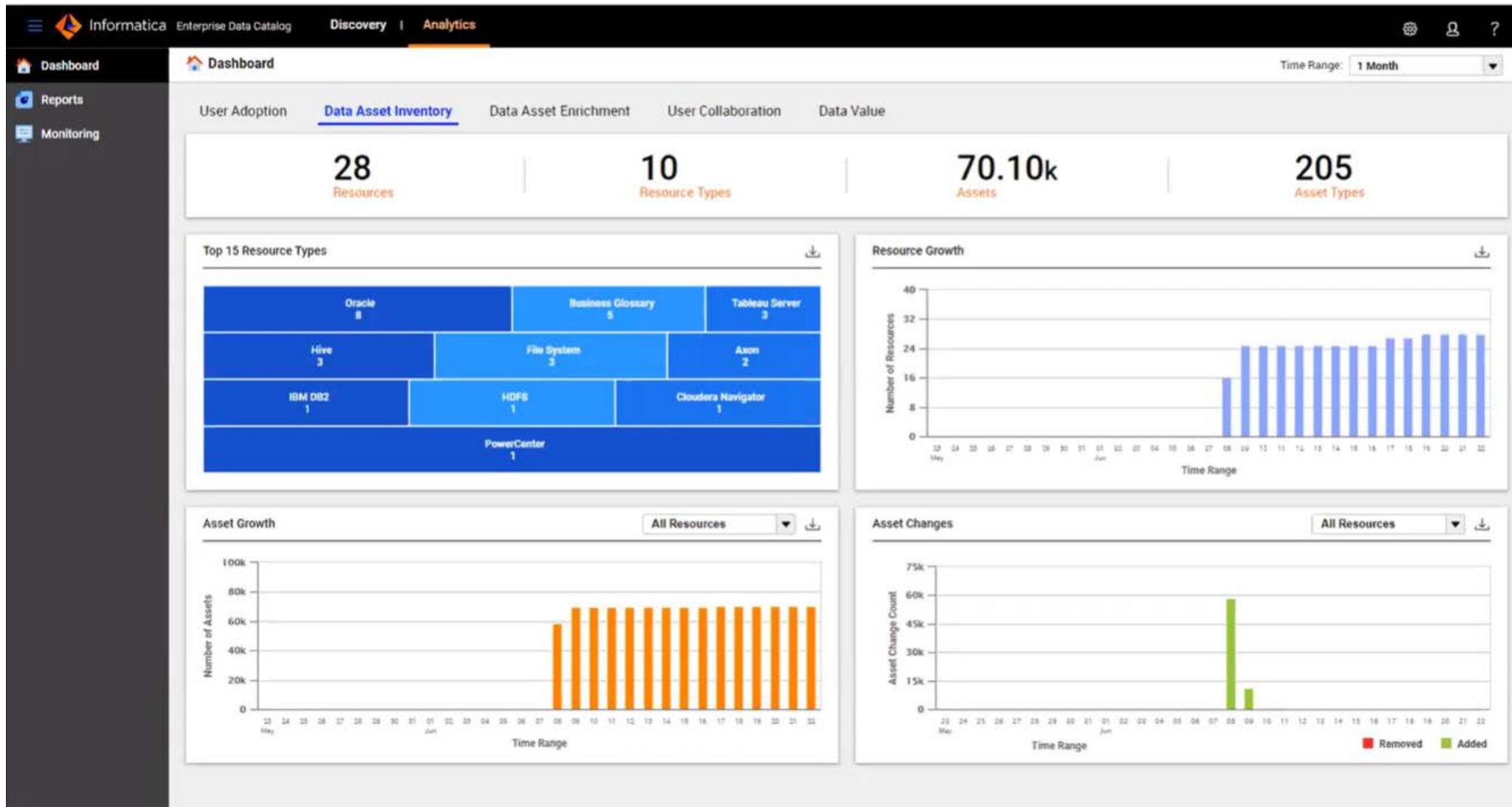
Documents - root@asjsridhar00502 - WinSCP

Save Cancel

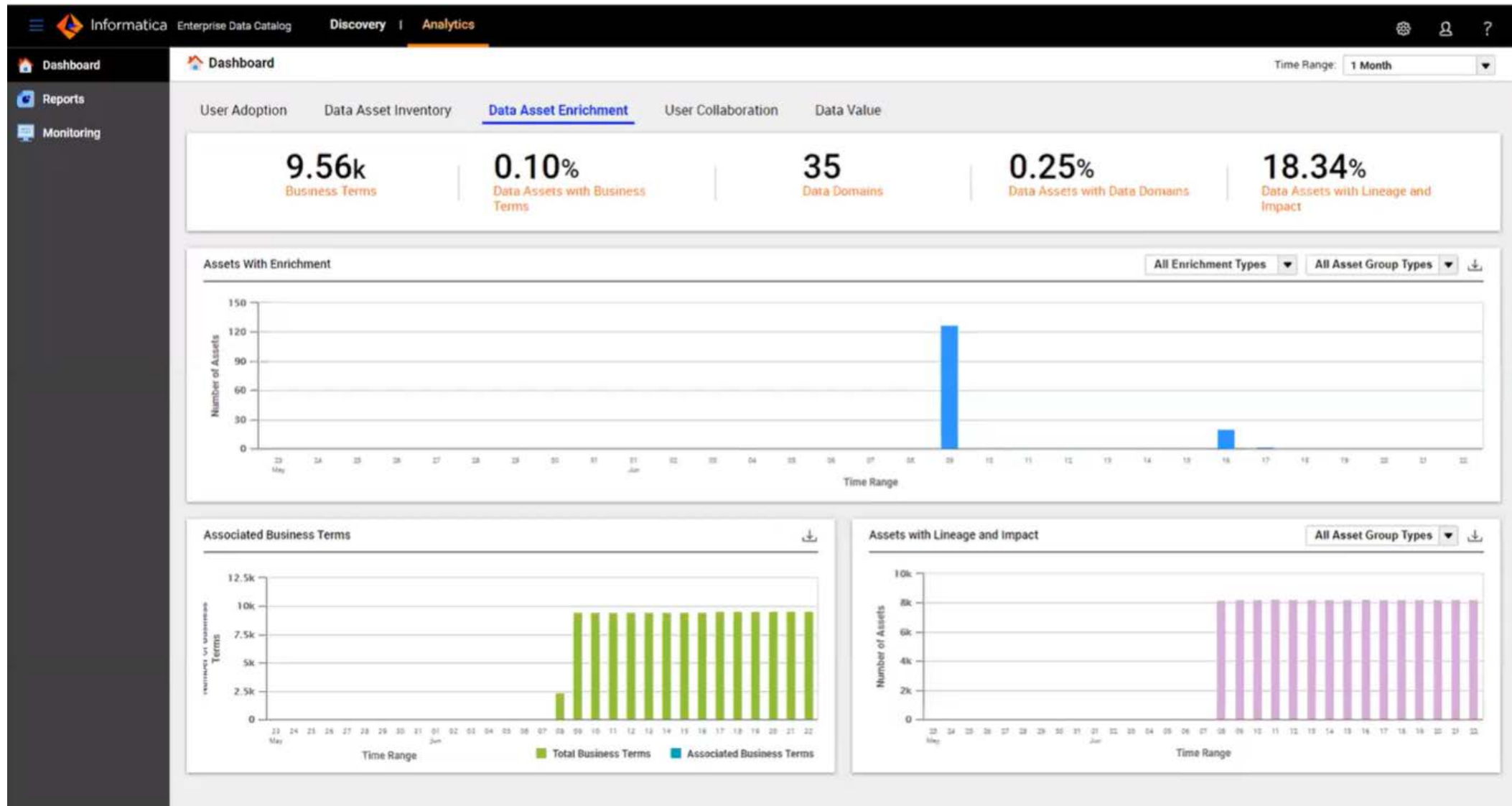
# Analytics: User Adoption Charts



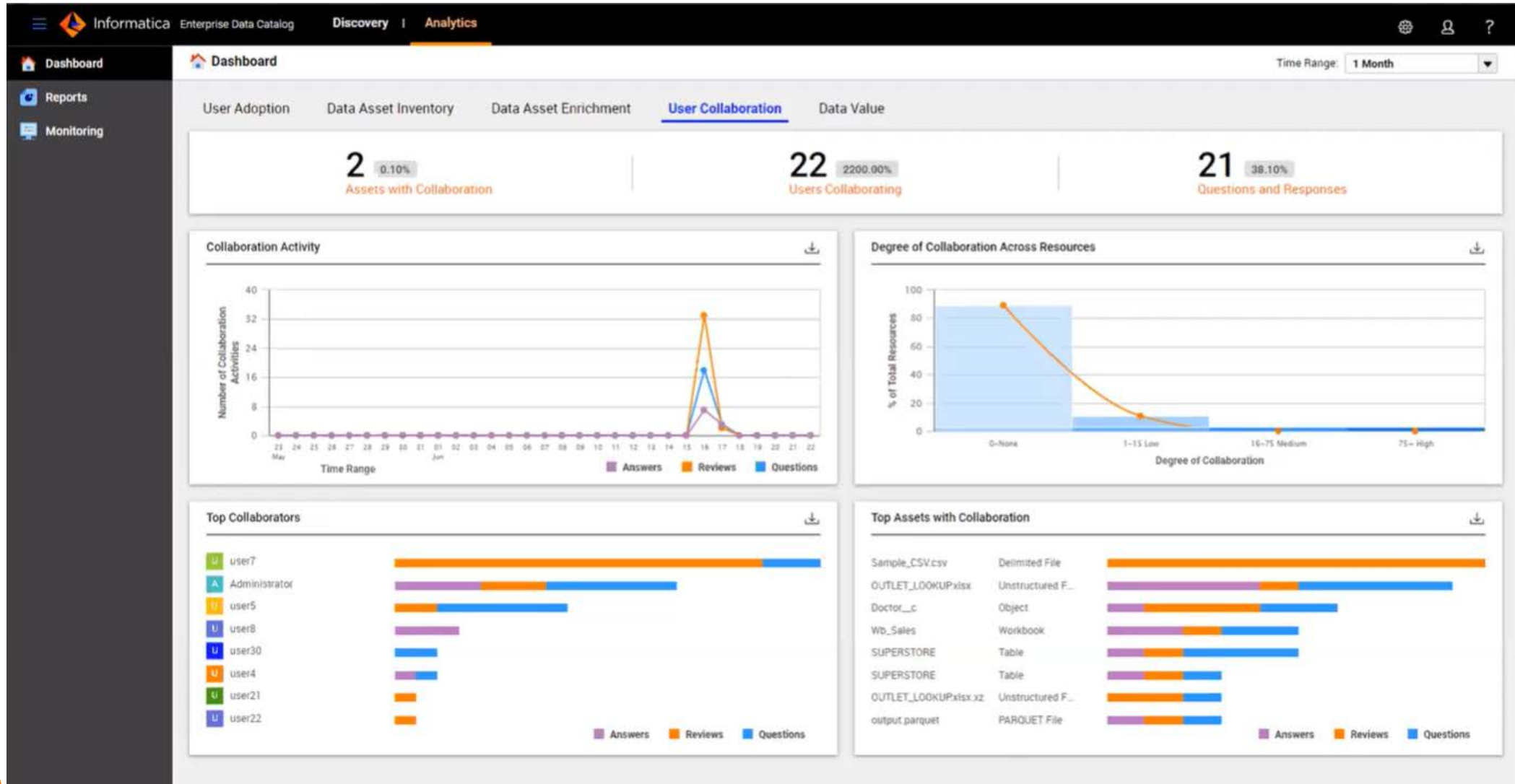
# Analytics: Data Asset Inventory



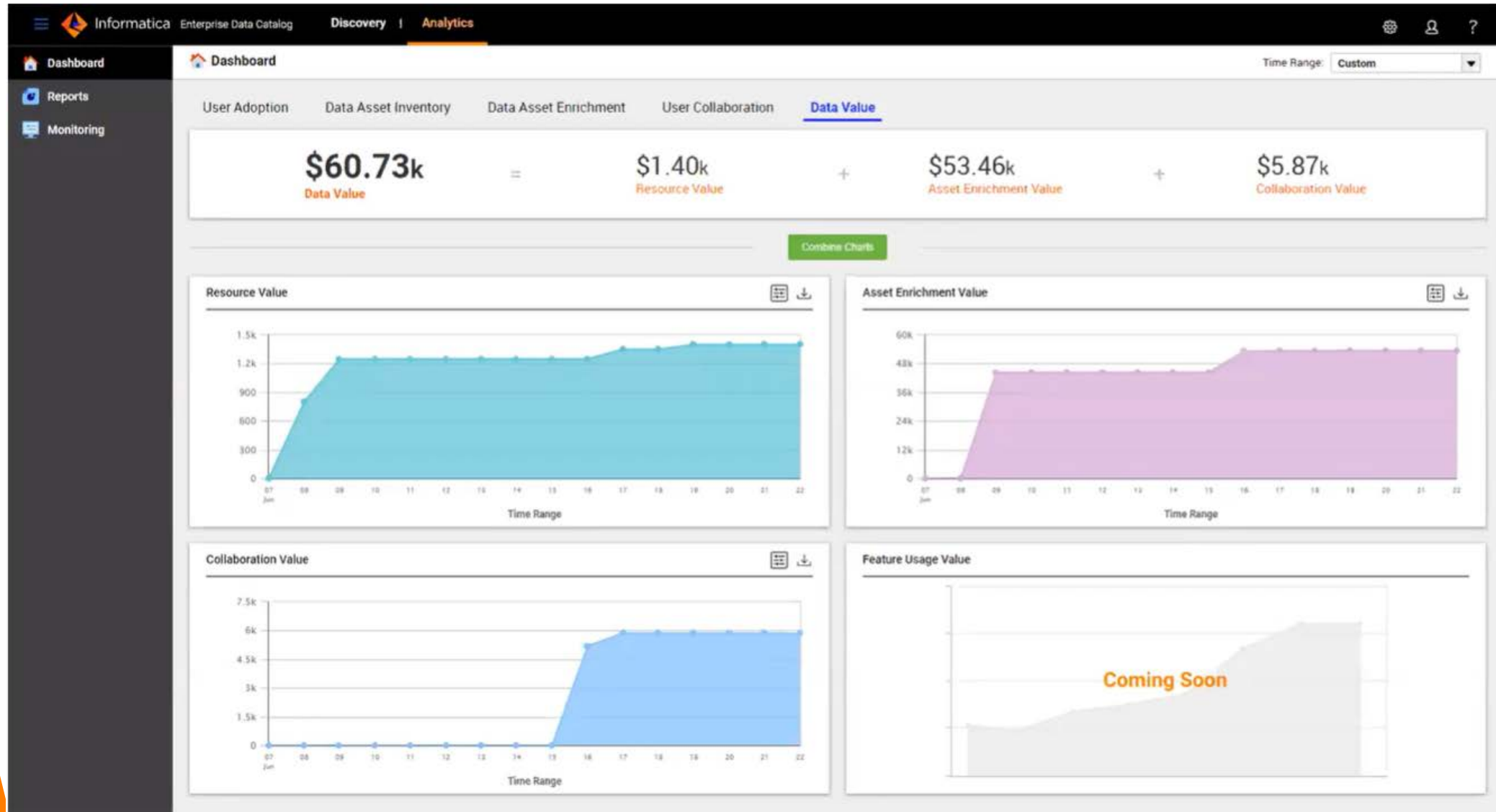
# Analytics: Data Asset Enrichments



# Analytics : User collaboration

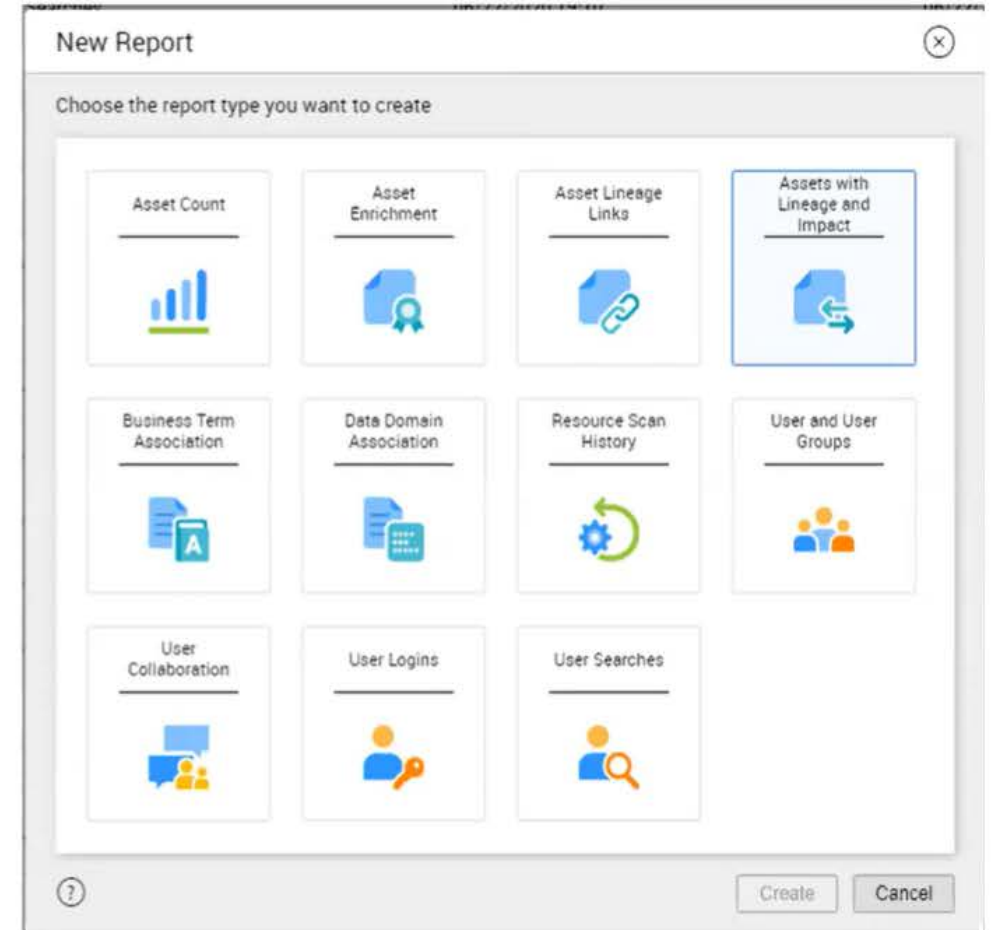


# Analytics: Data Value



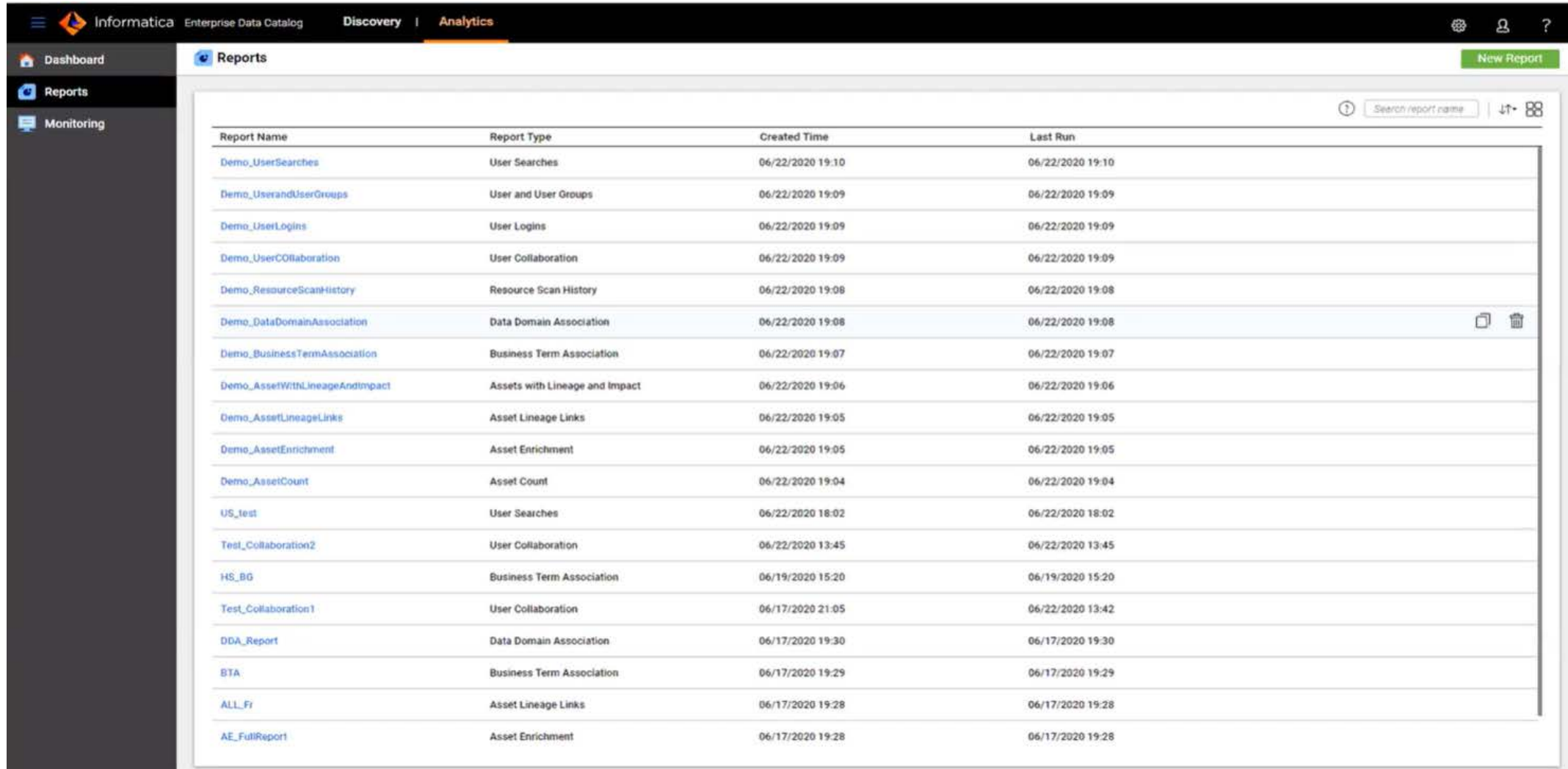
# Analytics: Report Type Details

Report Types	Report Description
Asset Count	Data asset count by asset types of resources in the catalog
Asset Lineage Links	Data asset lineage and/or impact links to other data assets
Asset with Enrichment	Data assets and associated enrichments types, such as description, business glossary terms and data owner
Asset with Lineage and Impact	Data assets and an indication of whether lineage and/or impact exists
Business Term Association	Business Terms and data asset association details
Data Domain Association	Data Domains and data asset association details
Resource Scan History	History of resource scans and scan settings
User and User Groups	Users who has access to the catalog and user group association details
User Collaboration	User collaboration activities on data assets
User Logins	User login timestamps
User Searches	Search strings entered by users





# Analytics: Reports



Report Name	Report Type	Created Time	Last Run
<a href="#">Demo_UserSearches</a>	User Searches	06/22/2020 19:10	06/22/2020 19:10
<a href="#">Demo_UserandUserGroups</a>	User and User Groups	06/22/2020 19:09	06/22/2020 19:09
<a href="#">Demo_UserLogins</a>	User Logins	06/22/2020 19:09	06/22/2020 19:09
<a href="#">Demo_UserCollaboration</a>	User Collaboration	06/22/2020 19:09	06/22/2020 19:09
<a href="#">Demo_ResourceScanHistory</a>	Resource Scan History	06/22/2020 19:08	06/22/2020 19:08
<a href="#">Demo_DataDomainAssociation</a>	Data Domain Association	06/22/2020 19:08	06/22/2020 19:08
<a href="#">Demo_BusinessTermAssociation</a>	Business Term Association	06/22/2020 19:07	06/22/2020 19:07
<a href="#">Demo_AssetWithLineageAndImpact</a>	Assets with Lineage and Impact	06/22/2020 19:06	06/22/2020 19:06
<a href="#">Demo_AssetLineageLinks</a>	Asset Lineage Links	06/22/2020 19:05	06/22/2020 19:05
<a href="#">Demo_AssetEnrichment</a>	Asset Enrichment	06/22/2020 19:05	06/22/2020 19:05
<a href="#">Demo_AssetCount</a>	Asset Count	06/22/2020 19:04	06/22/2020 19:04
<a href="#">US_test</a>	User Searches	06/22/2020 18:02	06/22/2020 18:02
<a href="#">Test_Collaboration2</a>	User Collaboration	06/22/2020 13:45	06/22/2020 13:45
<a href="#">HS_BG</a>	Business Term Association	06/19/2020 15:20	06/19/2020 15:20
<a href="#">Test_Collaboration1</a>	User Collaboration	06/17/2020 21:05	06/22/2020 13:42
<a href="#">DDA_Report</a>	Data Domain Association	06/17/2020 19:30	06/17/2020 19:30
<a href="#">BTA</a>	Business Term Association	06/17/2020 19:29	06/17/2020 19:29
<a href="#">ALL_Fr</a>	Asset Lineage Links	06/17/2020 19:28	06/17/2020 19:28
<a href="#">AE_FullReport</a>	Asset Enrichment	06/17/2020 19:28	06/17/2020 19:28



# Analytics: Reports Monitoring

The screenshot displays the Informatica Enterprise Data Catalog interface, specifically the Analytics section under Reports Monitoring. The left sidebar contains navigation links for Dashboard, Reports, and Monitoring. The main content area is divided into two sections: Report Runs and Report Preview.

**Report Runs**

Status	Report Name	Report Type	Start Time	End Time	Size
✓	Demo_UserSearches	User Searches	06/22/2020 19:10	06/22/2020 19:10	0.017 MB
✓	Demo_UserLogins	User Logins	06/22/2020 19:09	06/22/2020 19:09	0.03 MB
✓	Demo_UserCollaboration	User Collaboration	06/22/2020 19:09	06/22/2020 19:09	0.008 MB
✓	Demo_UserandUserGroups	User and User Groups	06/22/2020 19:09	06/22/2020 19:09	0.015 MB
✓	Demo_ResourceScanHistory	Resource Scan History	06/22/2020 19:08	06/22/2020 19:08	0.024 MB
✓	Demo_DataDomainAssociation	Data Domain Association	06/22/2020 19:08	06/22/2020 19:08	0.029 MB
✓	Demo_BusinessTermAssociation	Business Term Association	06/22/2020 19:07	06/22/2020 19:07	0.053 MB
✓	Demo_AssetWithLineageAndImpact	Assets with Lineage and Impact	06/22/2020 19:06	06/22/2020 19:06	1.753 MB
✓	Demo_AssetLineageLinks	Asset Lineage Links	06/22/2020 19:05	06/22/2020 19:05	3.038 MB

**Report Preview**

User Login Name	User Group Name	Search String	Search Date	User Group Association Created on	User Group Association Deleted on
Administrator	Everyone	*	06/17/2020	06/17/2020	
Administrator	Administrator	*	06/17/2020	06/17/2020	
DAA_Test1	Everyone	*	06/22/2020	06/22/2020	
DAA_Test2	Everyone	*	06/22/2020	06/22/2020	
DAA_Test2	Everyone	*	06/22/2020	06/22/2020	
Aaron_Dean	Human_Resources	Aaron_Dean	06/22/2020	06/22/2020	
Aaron_Dean	Everyone	Aaron_Dean	06/22/2020	06/22/2020	
ragilla	Everyone	*	06/22/2020	06/22/2020	

# FAQ/Troubleshooting

Ask	Response
Where to look for DAA log file ?	DAA log will be available under {INFA_HOME}/logs/{node_name}/services/CatalogService/{CS_service_name}/daa.log
How to enable DAA with debug log ?	Change log level to DEBUG in \$INFA_HOME/services/CatalogService/daa/WEB-INF/classes/log4j.properties
What are the supported databases for DAA repository ?	Check out Installation for Enterprise Data Catalog guide for more details. We support Oracle, SQL Server and PostgreSQL. Look for "Data Asset Analytics Database Requirements" section and PAM for supported versions
What are Database privileges required for DAA repository ?	Check out Installation for Enterprise Data Catalog guide for more details. Look for "Data Asset Analytics Database Requirements" section
DAA is optional during upgrade process ?	Yes. It is optional during upgrade process and user can also configure DAA post upgrade by going to CS service
What all application level privileges are available for DAA ?	There are three new high level privileges added for DAA <ul style="list-style-type: none"><li>• Report Management : View and Download Raw Data</li><li>• Dashboard Management : Visualize</li><li>• Admin Management : Configuration</li></ul>
On click of Analytics Tab, getting error. What could be the reason ?	Could be due to following reasons: <ul style="list-style-type: none"><li>• Not configured</li><li>• Manually disabled by the Administrator</li><li>• DB repository unreachable</li><li>• DB configuration is no longer valid</li><li>• Content creation failed during startup</li></ul>
How to assure that content creation is successful?	Once DAA repository has a table called DAA_SCHEMA_VERSION and contains a row having schema version and release version details

# Contd....

Ask	Response
When to reduce Cache to 0 ?	NEVER but If the load on system is low and also if the EDC activity is very less, we can go for it. Also, if there are symptoms of any memory issues on server we can try reduce this but please note this may impact the server response time
When to increase the Cache size?	If the DAA server response time is getting slow it would be suggested to increase chunks of 5MB. If the response time does not change you can revert it back.
What happens I update Cache Settings in Config page	<p>The new cache size and/or interval will be validated first. If they are valid, the entire Cache will be invalidated and restarts with new settings.</p> <p><b>Since this is an important configuration change, frequent updates are not allowed. Meaning, if you try update within 30 minutes of last update, DAA will not update but logs an error asking to retry after 30 minutes window.</b></p>
Any information will be lost if DAA is disabled and enable after sometime time or days ?	User need to trigger re-index to bring the newly added resource, assets and collaboration and enrichment activities but user adoption related information like logins and searches will be lost and can not be retrieved
How to interpret User creation date in DAA ?	User creation date is the date when user has been brought to DAA. This is "By Design" as Admin console not keeping track of user creation date.
I created some data in catalog few mins back but in DAA dashboard, the charts are not reflecting the newly created data. What could have gone wrong ?	It is because of the cache interval defined in the DAA( application configuration) wait for cache to get expired.