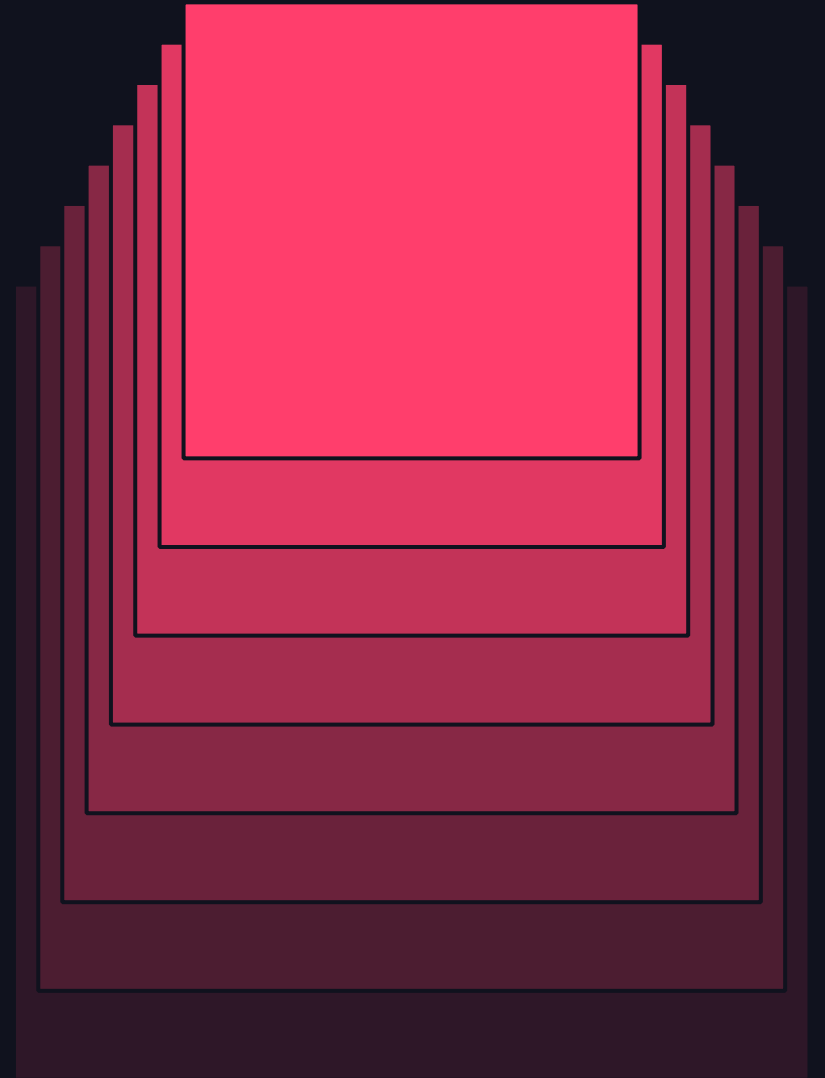


DRIVING DATA ACCESS GOVERNANCE and SECURITY



Navindra Yadav (Theom) and Zafer Bilaloglu (Databricks)
June 2024

Speaker Introduction



Navindra Yadav
Co-founder & CEO, Theom



Zafer Bilaloglu
Lead Solutions Architect, Databricks

Databricks Unity Catalog

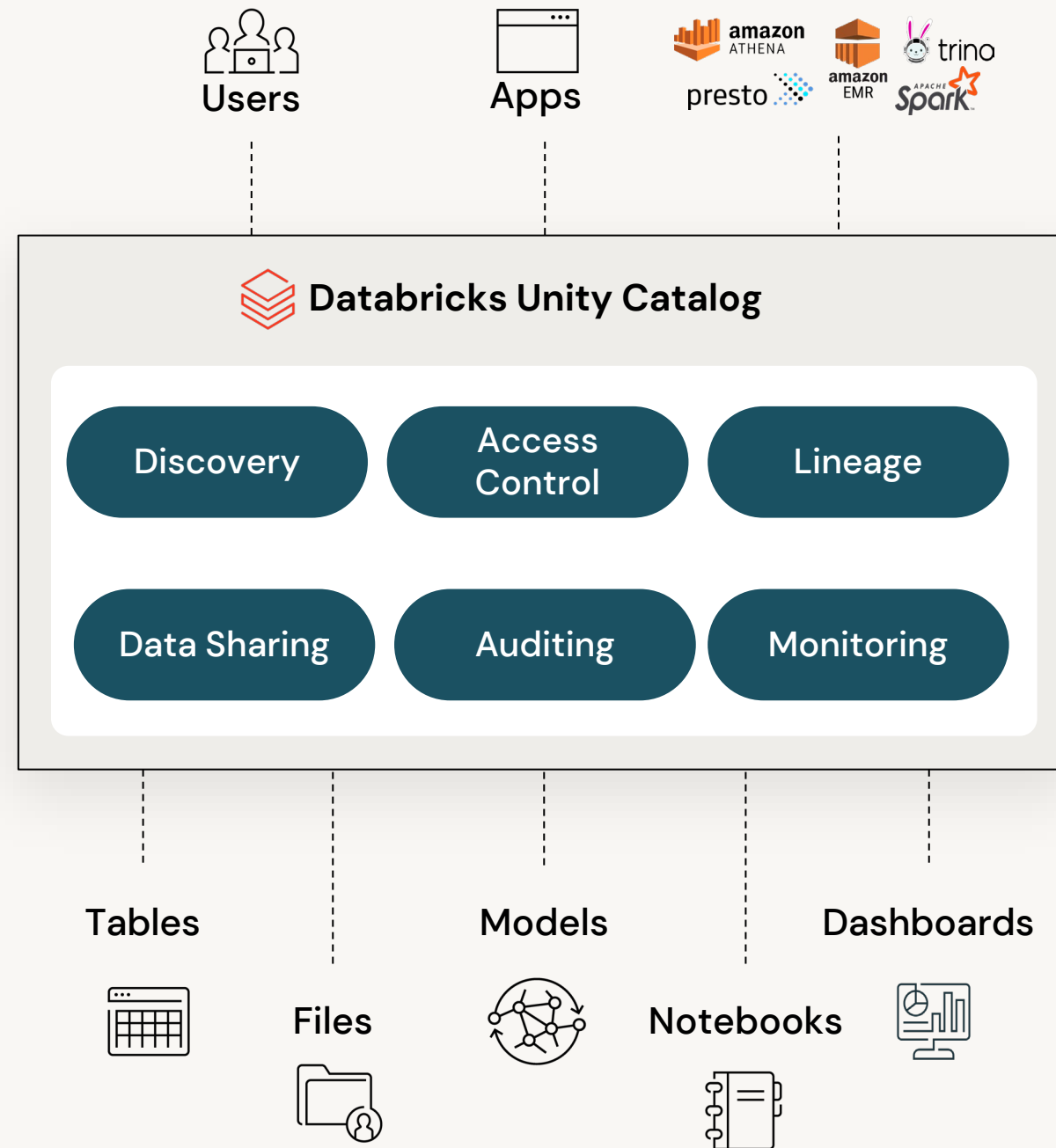
Unified governance for data & AI

Unified visibility into data and AI

Simple permission model for data and AI

AI-powered monitoring and observability

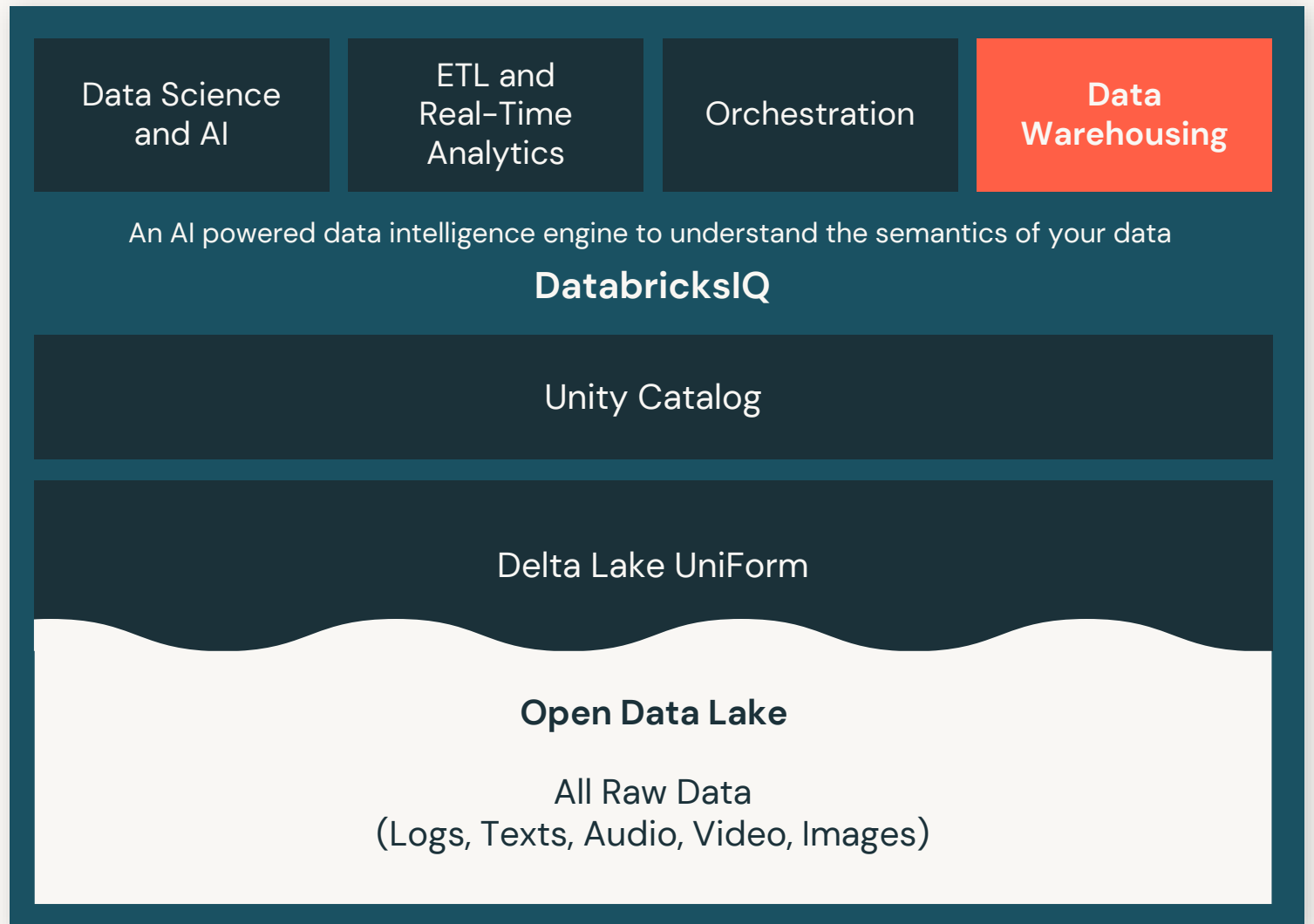
Open data sharing





Databricks SQL

intelligent data
warehousing on the
Data Intelligence Platform



What is Theom?



Discoverability / Active Metadata

DSPM+

Data Access Governance

Insider Risk/Threat Det. / DLP

Data Privacy / Lineage

Data Mesh / Data Contract Gov

Data Exch. Gov

AI Sec / Gov

Data Control Plane (Real time Observability & Control) Theom auto establish 5 Ws: What, Who, When, Where, Why

Data



AI Platform



Theom protects the data foundation to power AI from within the data store

1

Data Access Governance

Principle of Least Privilege Enforcement;
Data Access Compliance
Ideal State Discovery & Remediation

Discover Data, Policies & Compliance Issues

Automate Remediation

2

Detect & Stop Data Breaches

Insider Threat Detection & Data Leak Prevention

Detect Insider Risks & Active Data Attacks

Stop Data Attacks (via ITSM/SIEM/SOAR)

THEOM FOUNDATION
5 W'S:



WHAT

What Data?



WHERE

Where does it flow? Data Topology?



WHO

Who has access? Do they need access?



WHEN

When was it accessed?

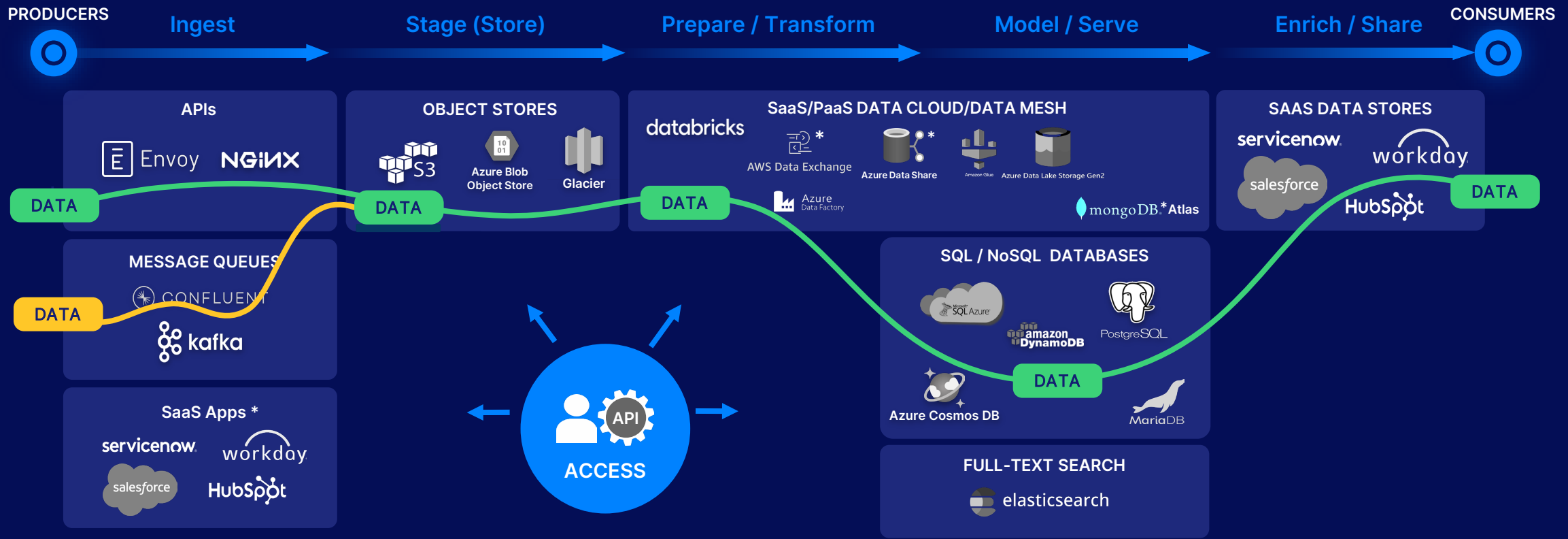


WHY

What's the purpose of access?

Federated SuperVision: Wide coverage

(Private/Public Multi Cloud + Data Cloud/Lake + OnPrem*)



Identity + Data + Flow of Data

- Who has Access to What Data?
- What are they doing with the Data?
- Data Exchange/Flow

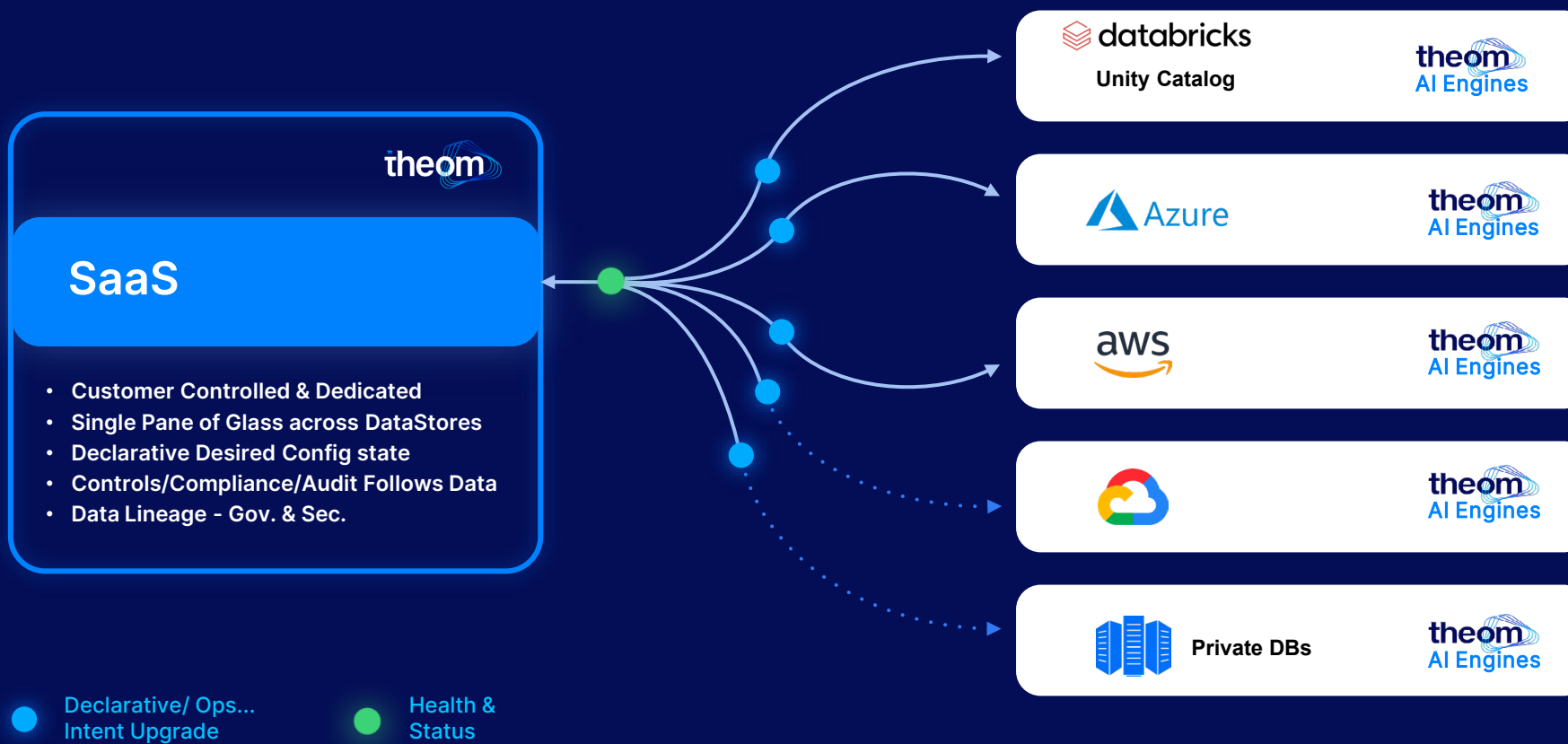


Automate

- AI & Data Governance & Compliance
- Breach Detection
- Data Contracts Compliance

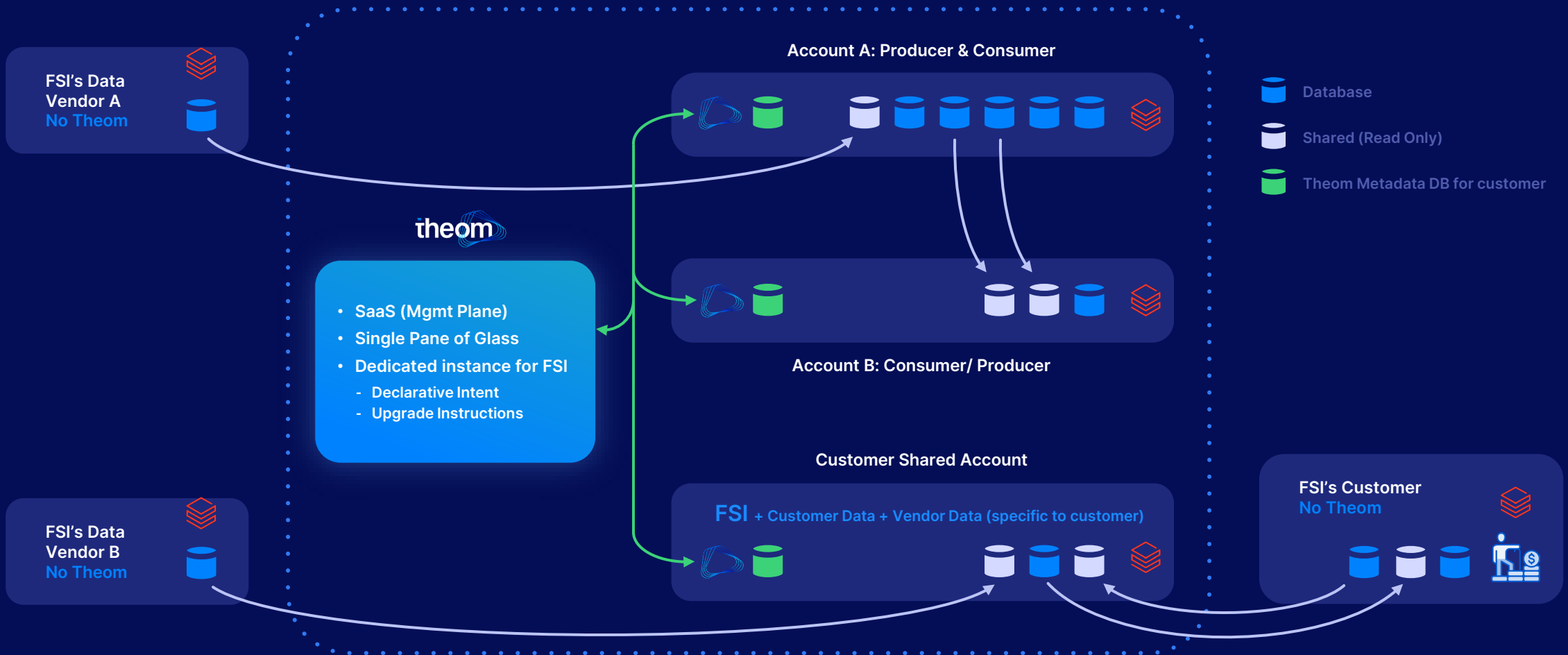
* On Contract

Theom Architecture: Best of both worlds (SaaS and On Premise)

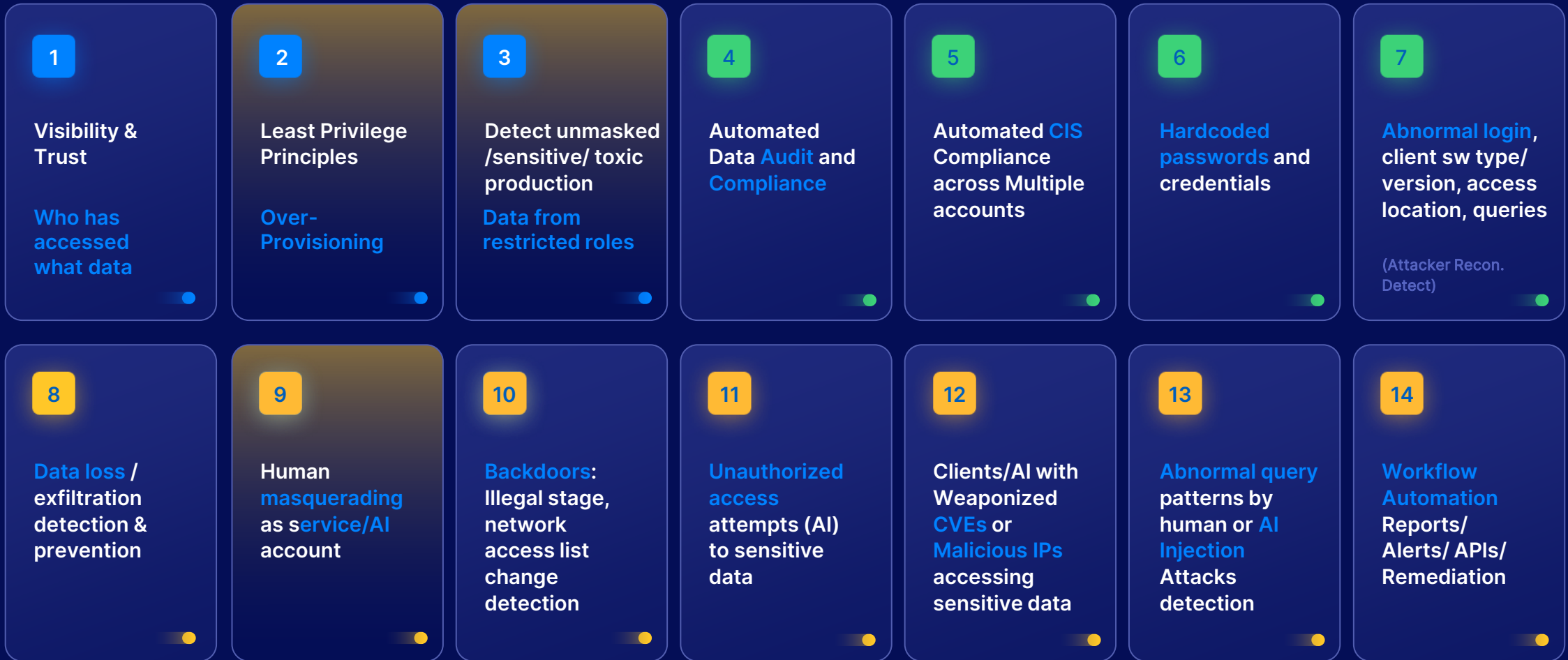


Novell
architecture


Theom with Multiple DB Accounts



Real world customer use cases: DAG & Security(Stop BreachesDLP/Insider Risk)

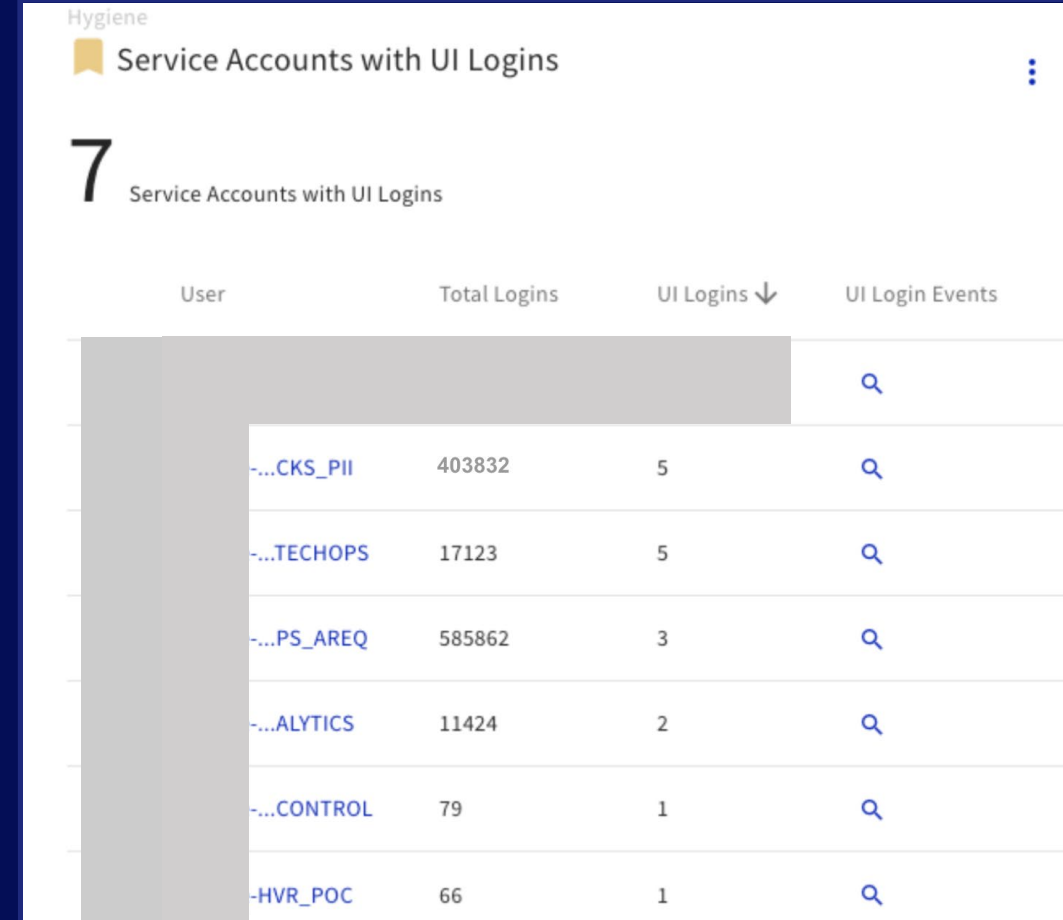


Humans Masquerading as Service Accounts

 **Theom detects and alerts when humans masquerade service accounts**

Prioritized by Recency and Prevalence

 **“Theom also identifies the human user behind the masquerading attempt!!”**



Hygiene

Service Accounts with UI Logins

7 Service Accounts with UI Logins

User	Total Logins	UI Logins ↓	UI Login Events
...			
...CKS_PII	403832	5	
...TECHOPS	17123	5	
...PS_AREQ	585862	3	
...ALYTICS	11424	2	
...CONTROL	79	1	
-HVR_POC	66	1	

#1 User

#2 Masquerading as Service

#3 Accessed Data

Theom's AI auto prioritizes interesting events by Recency, Prevalence, data Criticality

CLIENTIP	HUMAN USER(?)	QUERYTYPE	ROWSPRODU	QUERYTEXT	USERNAME	ROLENAME	OBJECTNAME
249	CHERIC@SF64859@DV44802@	SELECT	184158	SELECT DISTINCT vc.first_name, vc.last_name<trimmed> and p.scheduled_ date = current_date - 2	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.VIP_CUSTOMERS
249	CHERIC@SF64859@DV44802@	SELECT	10	select top 10 * from analytics.vip_customer_ customers AS vc	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.VIP_CUSTOMERS
249	CHERIC@SF64859@DV44802@	SELECT	10	select top 10 * from analytics.check_ins_rt.p ; AS vc	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.CHECK_INS_RT.
249	CHERIC@SF64859@DV44802@	SELECT	6	SELECT DISTINCT vc.first_name, vc.last_name<trimmed>e_date between current_date - 20 and current_date;	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.VIP_CUSTOMERS
249	CHERIC@SF64859@DV44802@	SELECT	1	select top 1 document_full_first_name, fr<trimmed> analytics.check_ins_rt.f documents AS td	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.CHECK_INS_RT. MENTS
249	CHERIC@SF64859@DV44802@	SELECT	1	select top 1 document_full_first_name, fr<trimmed> analytics.check_ins_rt.f ocuments AS td	-ANALYST_SVC_ITBI_DATABRICKS_PII-SVC_ITBI_DATABRICKS_PII	-ANALYST_SVC_ITBI_DATABRICKS_PII	ANALYTICS.CHECK_INS_RT. MENTS

IDENTIFIER	CRITICALITY	UI_LOGINS	OTHER_LOGINS	UI_LOGINS_LAST_7_DAYS	UI_LOGINS_LAST_90_DAYS	RECENTY	PREVALENCE
-SVC_ITBI_DATABRICKS_PII	HIGH	5	145	4	5	80	3.3
-HVR_	HIGH	1	17	1	1	100	5.6
-SVC_ S_CONTROL	HIGH	1	71	0	1	0	1.4
-SVC_	HIGH	78	15451	30	78	38.5	0.5
-SVC_	HIGH	2	10143	0	2	0	0
-SVC_	HIGH	3	510954	0	3	0	0

Detect: Unmasked Sensitive Data to unauthorized Roles



theom.ai/insights/datalake/...-ANALYTICS.SURVEYS.POST_DIGITAL?theomid=

Masking Policy Details

Table Column Masking Policy

Column	Masking Policy	Masking Policy Details
CUSTOMER_SATISFACTION	×	🔍
METADATA_INSERT_TIMESTAMP	×	🔍
POST_TRAVEL_NEED	×	🔍
POST_TRAVEL_NEED_OTHER	×	🔍
CUSTOMER_PHONE	×	🔍

Rows per page: 5 11-15 of 44 < >

Table Columns Usage Summary

Table column usage by role

SUMMARY ALL

By Role: ...-TRANSFORMER_PRD

Column	Accessed	Modified
CUSTOMER_SATISFACTION	✓	✓
METADATA_INSERT_TIMESTAMP	✓	✓
POST_TRAVEL_NEED	✓	✓
POST_TRAVEL_NEED_OTHER	✓	✓
CUSTOMER_PHONE	✓	✓

Rows per page: 5 11-15 of 44 < >

● Different workflows and Alert levels based on Sensitivity and Hygiene vs Misuse

Identify Over Entitled Datastores (Table/View/Object/...)



Overprovisioned Datastores 113

Monitors for overprovisioned tables in a data lake, identifying tables with excessive access. This detects potential governance issues and access control inefficiencies.

[View column explanation](#) ⓘ

Account	Datastore	Criticality	Financial Value	Overprovisioned ↑
[Redacted]	SACREDSOULHOSPITAL.PUBLIC.PATIENT	HIGH	\$ 6.8 K	<div style="width: 75%;"></div> 75 % ⓘ
	SACREDSOULHOSPITAL.PUBLIC.NURSE	HIGH	\$ 1.2 K	<div style="width: 84%;"></div> 84 % ⓘ

● Theom sorts based on Criticality or value of Data inside the Datastore (Table/View/Object/...)

Remediate over Entitled Roles by Datastore (Table/View/Object)



SACREDSOULHOSPITAL.PUBLIC.PATIENT

REMEDIATE ALL

Role Permissions/Privileges Analysis

Role	Permissions/Privileges	Remediation
DBAS_2024_03_05_05_41_AM	0% UPDATE 🔍 SELECT 🔍 0% INSERT 🔍	PENDING ⓘ
DBAS_2024_04_17_07_19_PM	0% UPDATE 🔍 0% INSERT 🔍 SELECT 🔍	PENDING ⓘ

● Theom automatically identifies over entitled roles and drives workflows to automatically/manually remediate

Remediate over Entitled Users by Datastore (Table/View/Object)



SACREDSOULHOSPITAL.PUBLIC.PATIENT

User Roles Analysis

Search for user and role

User

Roles

● PERRYC

0% HEADS 🔍

0% PHYSICIANS 🔍

● BOBK

0% HEADS 🔍

0% PHYSICIANASSISTANTS 🔍

0% PHYSICIANS 🔍

0% EPIDEMIOLOGISTS 🔍

- Theom automatically identifies over entitled users and drives workflows to automatically/manually remediate

Remediate over Entitled Groups/Roles/Users by Datastore (Table/View/Object)



Remediation Status

Details of the remediation query status

Query ID	Query	Queued At	Completed At	Status
01b3ff18-0002-6075-0019-f38700911156	REVOKE UPDATE ON ..._03_05_05_41_AM;	9 minutes ago	7 minutes ago	Completed
01b3ff18-0002-6075-0019-f3870091115a	REVOKE INSERT ON ..._03_05_05_41_AM;	9 minutes ago	7 minutes ago	Completed

- Theom can be integrated with the Entitlement system to drive fix up of Entitlements

Remediations & Remediation Workflows

Role	Permissions/Privileges	Remediation
ACCOUNTADMIN	0% INSERT 🔍 12% OWNERSHIP 🔍 12% SELECT 🔍 0% UPDATE 🔍	🔑
BUSINESS_ANALYST_INTERNS_2022	SELECT 🔍	🔑
BUSINESS_ANALYST_INTERNS_2023	SELECT 🔍	🔑
BUSINESS_ANALYSTS	50% INSERT 🔍 SELECT 🔍 50% UPDATE 🔍	🔑
DBAS	0% INSERT 🔍 50% SELECT 🔍 0% UPDATE 🔍	🔑
DBAS_2023_03_10_07_38_PM	0% INSERT 🔍 0% SELECT 🔍 0% UPDATE 🔍	COMPLETE ⓘ
DBAS_2023_03_27_08_58_PM	0% INSERT 🔍 0% SELECT 🔍 0% UPDATE 🔍	PENDING ⓘ

- Theom has built in Remediation workflows OR can integrate with external workflows (eg SOARs) or playbooks

Thank You! / Appendix

