Enterprise Data Platform Audit

Internal Audit Report – # **20188**

Bobby Mukundan, SVP Enterprise Technology Services
Kathleen Kadziolka, VP IT Enterprise Products
Navaneet Barathalwar, VP Data Engineering and Strategy

October 5, 2020

I. AUDIT BACKGROUND

The Enterprise Data Platform (EDP) environment is a critical repository of business intelligence information, which is leveraged across the enterprise using Hadoop technology and the latest tools to accelerate data analysis and reporting. The EDP combines various data sources to deliver a unified view of customer and business data. The environment, operating since 2014, is currently storing 19+ petabytes of data from more than 100 internal and external sources, roughly a 36% increase in data storage since the first audit over this environment in 2016. The environment is an integral source of stored data used in the Enterprise's "Big Data" processes and is utilized by various stakeholders in support of data science initiatives and operational reporting used to execute downstream business decisions.

The EDP team works with various business stakeholders to provide deeper data insights to allow for strategic business decisions to be made. Examples include providing advanced analytic solutions to promote consumer engagement, improve health outcomes, and drive revenue growth by solving complex business problems, while utilizing the "Big Data" concept. Data content is modeled, and analytics are performed using multiple structured and unstructured data sources to gain valuable business insights. Reports and dashboards are created for each project by utilizing business intelligence (BI) and reporting tools such as Tableau, which supports interactive visualization for in-depth analysis.

In Q4 of 2019, the company developed an initiative to migrate CVS Health source data from various data warehouses into the EDP. A separate environment was created for CVS data utilizing the Hadoop Infrastructure Platform and will contain several classifications of data including PII, PHI or HIPAA, and SOX relevant data. A logical firewall has been established between the lines of business (Retail and Long-Term Care (LTC), Pharmacy Benefit Manager (PBM), and Insurance) to prevent unauthorized access to confidential, proprietary, or otherwise sensitive information. As this asset is a single source of insight, it is important to ensure that the data contained within is accurate, and that the information is stored in a secure manner.

II. OBJECTIVE & SCOPE

The objective of this audit was to provide reasonable assurance that management has effective controls in place to ensure the design and operating effectiveness of the Enterprise Data Platform. This audit was performed for the period November 2019 through May 2020 and the scope included the following:

A) Governing controls are in place to monitor the ongoing effectiveness of the end to end data platform environment and to ensure the data lifecycle effectiveness.
B) Evaluate EDP patch management procedures for policy changes or version upgrades to ensure alignment with company standards.
C) Evaluate EDP change management procedures for policy changes or version upgrades to ensure alignment with company standards.
D) Access and logical firewall controls are in place to ensure access is appropriately managed by authorized individuals.
E) Data integrity controls are in place to ensure the data entering the EDP is complete and accurate.
F) Data protection controls are in place to ensure data is secured in transit and at rest.

III. CONCLUSION & FINDINGS SUMMARY

☐	Effective	☒	Mostly Effective	☐	Improvement Needed	☐	Ineffective

Based on the procedures performed for this audit, except for the issues noted, controls in place over Enterprise Data Platform provide reasonable assurance that the business risks reviewed are adequately mitigated.

Description	Business Area	Rating	Ref. to Scope Item
Findings			
1. Off-shore User Access	Business Intelligence Solution / Human Resources IT (HR IT)	High	D
2. Periodic Monitoring Process	Enterprise Products	Medium	D
3. Minimum Security Baseline	Enterprise Technology Services (ETS) / Global Security	Medium	C
4. Policy and Procedures	Enterprise Products	Low	A
Observations			
1. Decryption Tickets	Enterprise Products	MO	E
2. Access De-provisioning	Enterprise Products	MO	D
3. Data Monitoring	Enterprise Products	MO	E
4. Steering Committee	Enterprise Products	MO	A

IV. FINDINGS & MANAGEMENT ACTION PLANS

1. Off-Shore User Access

The EDP environment hosts data from across the enterprise including data classified as Restricted Client Data requiring that no off-shore personnel may obtain access to view or process any such data. Access is granted by the various group owners that own the data and is their responsibility to follow policies (Restricted Data Client Handling Standard). Access was provisioned by two approvers, first approval is the user's immediate supervisor, second approver is from the analytics group where the data is owned.

Currently, off-shore contract users are not accurately distinguished within the HR systems. Without the identification or flag of off-shore contract users prior to administering network group access, various business units are not able to systematically determine a user's location to perform valid access reviews for restricting access not made available to off-shore personnel. Internal Audit performed an analysis for all users who were provisioned access to the platform and concluded that nine unique users were identified with offshore/offsite location ID, across eight network groups. Upon further analysis, eight users had the wrong code and were updated in the system, and one user was identified as offshore with inappropriate access to production PBM data, which contains restricted client data. As a result, CVS Health may be in breach of regulatory and/or contractual obligations.

Rating: High

Management Action Plan:

HR IT:

HR Technology changed the generic location data for contingent workers that are not on-site on 7/17/2020. Ultipro was coded to assign the default location code based on the existing location type of On-Shore, Off-Shore, or On-Site, and the location codes have been updated for all existing contingent workers. This now allows network group administrators or business units to clearly distinguish offshore users. (Completed)

Business Intelligence Solution:

1. Partner with the EDP, HR teams and Info Governance/Legal to understand the policy (restricted data clients – Caremark) & process to identify offshore resources
2. Train all managers / approvers on the policy
3. Institute a monthly internal review of approved contractors including revoking access to onsite contractors who have moved offshore

Responsible Party: Sanjay Rajagopal, Director, Analytics
Mahesh Madadi, Sr. Director, Data Engineering
Debra Osborn, Executive Director, IT HR

Target Completion Date: 10/30/2020

2. Periodic Monitoring Process

Prior to this audit, a periodic monitoring process was not in place. While the responsibly of user provisioning (requests, approvals, and establishment) is owned by the individual technology business units, the Enterprise Product Team is also responsible for governing the access granted to users across their platform.

IA identified one off-shore user who was not appropriately provisioned to user network groups granting access to Restricted Client Data information. Without a periodic monitoring process in place to review off-shore users and their access within the Enterprise Data Platform, individuals may not be properly authorized to view restricted or otherwise sensitive data contained within the environment. As a compensating control to user provisioning processes, IA recommends that the Enterprise Product Team perform regular access reviews of the users granted access to the Enterprise Data Platform.

Rating: Medium

Management Action Plan:

The flag to identify user location is now accessible when reviewing access requests. This will streamline the manual process and allow approvers to respond more efficiently. Enterprise Products will perform a semiannual periodic recertification exercise for all users in production support groups, in addition to assessing all users migrating from the 2.6x environment to ensure all users' access remains appropriate within the UDP environment. (10/16/20)

In addition, the Enterprise Products team sent out a communication to all network group owners with access to the UDP environment illustrating the user provisioning procedures, which includes a review of user locations to ensure offshore users are not granted access to sensitive information. (Completed)

Responsible Party: Carmen Malangone
Executive Director, Enterprise Products

Target Completion Date: 10/16/2020

3. Minimum Security Baseline (MSB)

MSBs are the minimum information security configuration standards and processes, utilized within the Company's IT infrastructure, that ensures new servers and other equipment are initially installed with the appropriate level of security necessary to keep the Company's systems safe from hacking and data breaches.

MSB scans were performed for all 3 sampled production EDP edge node servers (1 CVS server 'rri2hdpepl10p' and 2 Aetna servers 'xhadrtlpiam1p' and 'xhadpbmpiam1p'). Identified failed results are tracked however they are not prioritized or assigned for remediation. Configurations that failed include password complexities, account expirations, default login settings, account timeouts, and IP redirects.

Failure to ensure systems are compliant with CVS Health Information Security requirements, could result in the exploitation of security weaknesses to gain access to sensitive data leading to disclosure of privileged information.

Rating: Medium

Management Action Plan:

Global Security (GS):
GS will perform the following actions to address this issue:
a. Review/validate/publish new minimum-security baseline configurations for RHEL 7 servers across the Hadoop Infrastructure with Engineering teams (October 16, 2020).
b. Re-run scans based on validation reviews (October 16, 2020).

Enterprise Technology Services (ETS):
There is an effort underway to revise the Enterprise model for MSBs, as such ETS will perform the following actions to address this issue:
a. Based on new re-run scans, work with engineering leads to follow the governance model for remediation activities outlined in ACS-4763 and create targeted plans (by operating systems) remediation plans based on priority of impact to system availability (October 16, 2020).
b. Complete Remediation Rollout (January 29, 2021).

Responsible Party: Razi Uddin, Sr. Director Systems Engineering
Hal Yost, Sr. Director Systems Engineering
Karolyn Maloney, Sr. Director, Information Security

Target Completion Date: January 29, 2021

4. Policy and Procedures

Enterprise Data Platform policies and procedures were reviewed to ensure alignment with Enterprise Standards. Policies and procedures related to EDP are stored within the Confluence site to ensure users have the proper guidance when navigating the EDP environments, and their associated processes/procedures. The following EDP policies and procedures were not updated to reflect the required annual review:

a. Access Control Policy
b. Data Ingestion Procedure
c. Data Protection Policy
d. Data Use Policy
e. Decryption Procedure
f. Encryption Procedure
g. Environment Oversight Procedure
h. External Release Policy
i. External Release Procedure
j. Tenancy Policy
k. Technical Metadata Procedure
l. UDF Management Policy
m. Ingestion Production Support Procedure

In addition to the annual review, proper ownership over the policies and procedures should be evaluated. Of the 13 EDP related policies and procedures reviewed, 11 displayed ownership of individuals who had changed roles or are no longer responsible for the processes.

Failure to ensure policies and procedures are kept up to date including appropriate ownership, could result in modifications of procedures not being reflected, leading to potential misalignment with Enterprise standards.

Rating: Low

Management Action Plan:

The EDP Production Support team reviewed the policies in the current confluence site and conducted a review on those documents to ensure appropriate ownership, assessed for any process changes, and the review dates now represents the annual review by July 10, 2020.

Responsible Party: Carmen Malangone,
Executive Director, App Software Deliver

Target Completion Date: Completed

V. OBSERVATIONS

Internal Audit (IA) identified the following opportunities to enhance management's processes.

1. Decryption Tickets

The current workflow is outdated and does not allow for updated required process attributes to be recorded within the ticket. The team is relaying on manual inputs of comments to track approvals and hive policy creations, which are policies that apply to the EDP environment. IA recommends updating the JIRA tickets to appropriately represent the required information needed to support the decryption process approval for all EDP decryption requests.

2. Access De-provisioning

Enterprise policy ATP 009 – Access Control Policy requires access to system resources is granted on a minimum basis and revoked when no longer required. IA identified four users who still have access and are no longer with the organization. The EDP environments utilize a Single Sign On (SSO) authentication method to access the platform. While the four network ID's have been disabled, IA recommends removing all inactive users or users no longer requiring permissions from groups that grant access to the EDP environments

3. Data Monitoring

In response to a previous audit finding, management implemented a data monitoring tool called DataGuise. DataGuise is a solution that detects, audits, protects, and monitors sensitive data assets in real time over big data environments. Management determined the DataGuise tool was ineffective in accurately identifying unencrypted confidential information on the platform. Management relies on the compensating encryption process that utilizes data dictionaries to identify PHI and PII data elements during ingestion process, to ensure confidential data is encrypted and is not displayed in plain text. Enterprise Products plans to migrate the data from the previous 2.0 platform to the 3.0 platform and will re-perform the ingestion process over all the data within the environment to ensure proper encryption is applied.

4. Steering Committee Agenda

Steering Committees help to steer a project from start to completion and are made up of representatives and stakeholders who partner together to ensure the project is correctly meeting its target. The role of the Steering Committee is to provide advice, ensure delivery, and achievement of project outputs. Lack of establishing the steering committee and meeting on a frequent basis could result in risks not being properly identified or tracked and could impact the overall project deliverables.

The EDP management team is currently in the process of formalizing a Data Platform Steering Committee made up of representatives from the Chief Technology Office, Infrastructure, and the Data Office, however these three stakeholders are currently holding their own individual meetings. IA recommends consolidating the three independent business line meetings and create a Steering Committee focused on Data Strategy that meets on frequent intervals.

VI. STANDARD TERMINOLOGY

Overall Control Environment Opinion:

Effective	Overall, controls are appropriately designed and functioning as intended. Control weaknesses, if noted, do not threaten the effectiveness of the process reviewed.
Mostly Effective	Except for the issues noted, controls in place provide reasonable assurance that business risks are adequately mitigated.
Improvement Needed	One or more significant control weaknesses exist that require prompt action to prevent the process from becoming ineffective.
Ineffective	Control weaknesses are pervasive or one weakness is so severe that it impacts the entire operation under review. Immediate management attention is needed to remediate the finding identified.

Rating:

Low	The identified risk does not warrant immediate attention; however, there should be an agreed-upon action plan for ultimate resolution.

Medium	The identified risk requires the near-term attention of the responsible manager. There should be an agreed-upon action plan for its resolution.
High	The identified risk requires the immediate attention of department and senior management to prevent the process from becoming ineffective, and an agreed-upon action plan for resolution is needed.

Other:

Target Completion Date	Reflects the time required for Management to complete the agreed upon action plan, as well as time for IA to complete the associated validation procedures to ensure the action plan has been implemented effectively.

VII. DISTRIBUTION LIST

Business Planning Committee
Larry Merlo
Eva Boratto
Troy Brennan
Tom Moriarty
Roshan Navagamuwa
Jon Roberts

Compliance
David Falkowski

Legal
Betsy Ferguson
Anna Shimanek

Enterprise Products
John Laguna
Carmen Malangone

Business Intelligence Solution
Bob Darin
Karthik Kirubakaran

Global Security
Sofia Bayne
Karolyn Maloney

Data Analytics
Ali Keshavarz
Shyam Munjuluri

Enterprise Technology Services
Nancy Gorczyca
Christopher Millheim
Razi Uddin
Hal Yost

Internal Audit
Lynn Atkin
Michael Bavasso
Dan Benner
Bryan Nazworth
Shawn Sousa
Sol Vazquez

External Audit
Allison Capprini
Tom Derkacs
Mike Fischer
Ryan Murphy