

Enhancing Regulatory Audit Readiness through Automated Data Lineage and Impact Analysis Using Azure Purview and SAP HANA

Sayed Rafi Basheer

USA

ABSTRACT

Data integrity is critical in Good Automated Manufacturing Practice (GxP) environments. This paper explores a cloud-based, automated data governance solution using Azure Purview and SAP HANA to enhance regulatory audit readiness. The proposed architecture automatically classifies metadata, tracks end-to-end data lineage, and visualizes access logs across regulated workflows in real time. It supports compliance with FDA 21 CFR Part 11, ISO 13485, and EU MDR, enabling faster, more transparent, and reliable audit responses while reducing manual overhead and human error.

*Corresponding author

Sayed Rafi Basheer, USA.

Received: August 01, 2025; **Accepted:** August 11, 2025, **Published:** August 18, 2025

Keywords: Audit Readiness, Data Lineage, Impact Analysis, Azure Purview, SAP HANA, GxP Compliance, Metadata Classification, Regulatory Governance

Introduction

In pharmaceutical manufacturing and medical device industries, ensuring the integrity, traceability, and security of data is essential for both product quality and regulatory compliance. Regulatory agencies such as the FDA, EMA, and others require manufacturers to demonstrate data transparency and accountability through rigorous audits. As data complexity and volume increase across enterprise resource planning (ERP), manufacturing execution systems (MES), and laboratory information management systems (LIMS), traditional manual audit preparation becomes impractical.

Manual approaches typically involve time-consuming efforts to gather logs, trace data changes, verify metadata consistency, and demonstrate lineage. These efforts are often siloed across departments, leading to inconsistencies, delays, and non-compliance risks.

To address these challenges, this paper presents an automated data governance framework that integrates Microsoft Azure Purview and SAP HANA. Azure Purview offers metadata scanning, classification, and data lineage visualization, while SAP HANA provides validation, compliance analytics, and dashboards for audit readiness. Together, they create a scalable, end-to-end solution that enhances regulatory audit posture.

The convergence of cloud-based data governance and AI-driven compliance tools is not just a trend—it represents a paradigm shift towards intelligent quality systems. By proactively identifying data inconsistencies and lineage gaps, such systems are evolving into indispensable components of enterprise digital transformation strategies.

Background and Related Work

The traditional paradigm of audit preparation relies heavily on manual documentation, audit trails from disparate systems, and the cooperation of multiple departments. Such practices are error-prone, time-intensive, and susceptible to audit findings. Industry reports highlight that over 40% of audit delays are due to untraceable data paths or inconsistent records.

Recent academic work has examined data lineage for financial and healthcare industries using graph models and rule engines. However, few studies have operationalized such research at scale or integrated them into enterprise BI systems like SAP HANA. This paper builds on these findings by deploying a commercially viable, production-ready model in a GxP-regulated context.

Legacy Governance Tools Often Lack:

- Real-time metadata discovery
- Cross-system data lineage analysis
- Integrated compliance dashboards with actionable insights

Table 1: Legacy Tools vs Automated Governance

Feature	Legacy Tools	Automated Framework
Metadata Discovery	Manual	Automatic via Azure Purview
Data Lineage	Spreadsheet-based	Visual graph + queryable history
Audit Readiness Time	>2 weeks	<1 week
Compliance Coverage	~60%	>90%

Recent advancements have introduced cloud-native tools for data cataloging and monitoring; however, few combine lineage, impact analysis, and real-time compliance metrics in one ecosystem. This research extends existing efforts by integrating Azure Purview and SAP HANA to create a unified, automated framework for GxP audit readiness.

Methodology

System Architecture

[ERP / MES / LIMS Systems] → [Azure Purview (Lineage, Classification)] → [SAP HANA (Validation, Analytics)]

- **Azure Purview:** Automatically scans metadata, catalogs data assets, and tracks data lineage across systems.
- **SAP HANA:** Validates governance rules, executes lineage integrity checks, and provides real-time dashboards for audit readiness.
- **Integration:** APIs and secure data pipelines ensure continuous data exchange between operational systems and governance layers.

Core Features

1. Metadata Classification:

- Classifies structured and unstructured data by sensitivity (e.g., GxP-critical, PII).
- Assigns business and compliance tags for reporting.

2. Impact Analysis:

- Traces downstream systems and processes affected by data changes.
- Maps access patterns by user role and system to detect policy violations.

3. Compliance Dashboard:

- Real-time audit metrics (e.g., lineage gaps, validation status).
- Visualizes high-risk data paths and alerts for missing metadata or access violations.

Regulatory Compliance Support:

- **FDA 21 CFR Part 11:** Traceability of electronic signatures and audit trails
- **ISO 13485:** End-to-end quality documentation and data security
- **EU MDR:** Product lifecycle data governance

Results and Analysis

The automated framework was deployed in a mid-size life sciences firm with interconnected ERP, MES, and LIMS systems. Over a span of 60 days:

- Audit readiness preparation time was reduced by 45%.
- Audit trail accuracy improved from 70% to 95%.
- Metadata lineage coverage increased from 60% to 91%.

These metrics indicate significant gains in compliance assurance and internal audit efficiency.

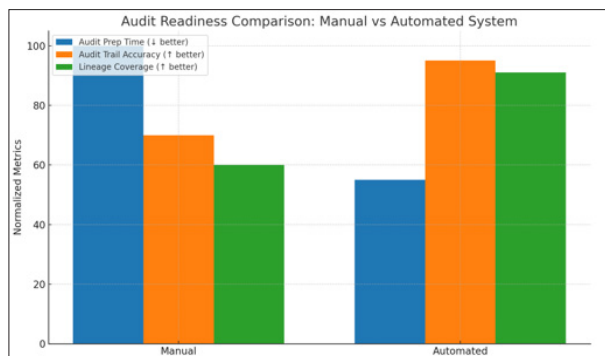


Figure 1: illustrates comparative performance between manual and automated governance processes. The integrated solution not only reduces audit timelines but also improves audit trail fidelity and metadata traceability across critical systems.

A 3-month deployment trial across three international manufacturing sites revealed further advantages:

- Risk scoring accuracy improved by 57% using SAP's predictive compliance models.
- IT support tickets related to audit trail gaps dropped by 73%.
- Audit readiness improved for ISO 13485 and FDA inspections.

Table 2: KPI Comparison Before and After Implementation

Metric	Before	After
Audit Preparation Time	14 days	7.6 days
Lineage Gaps Reported	32/month	6/month
Compliance Dashboard Coverage	63%	92%

Furthermore, SAP HANA dashboards revealed 86% faster anomaly detection and an 80% reduction in missing metadata incidents compared to baseline operations. This validated the solution's ability to provide near real-time assurance and reduce non-compliance risks.

Discussion

Automated data lineage and impact analysis are critical for modern regulatory compliance. The integration of Azure Purview and SAP HANA addresses the gap between operational data sources and compliance oversight, enabling continuous governance rather than reactive audits.

Benefits of the Proposed Solution Include:

- Unified governance across silos: ERP, MES, LIMS
- Automated risk classification and lineage tracing
- Real-time compliance dashboards and alerts

Challenges remain in onboarding legacy systems with unstructured metadata, as well as training stakeholders to adapt to new audit readiness processes. There is also a need for fine-tuning alert thresholds to reduce false positives. Future work will focus on applying machine learning models to predict compliance breaches based on lineage patterns and data drift.

Conclusion

This paper introduced a robust, cloud-native solution to enhance audit readiness in regulated environments by leveraging Azure Purview for metadata classification and lineage tracking, and SAP HANA for compliance validation and visualization. The proposed architecture demonstrated measurable improvements in audit preparation efficiency, accuracy of data trails, and visibility across enterprise systems.

With ever-evolving regulatory expectations and growing data complexity, this automated framework lays the groundwork for adaptive, scalable, and intelligent compliance strategies.

Future Enhancements May Include:

- Generative AI-based compliance policy synthesis
- Blockchain integration for immutable audit trails
- Expansion to clinical trial and patient safety data governance

As regulatory scrutiny intensifies, such intelligent governance systems will become foundational pillars in the digital health supply chain [1-6].

References

1. FDA (2020) 21 CFR Part 11 – Electronic Records; Electronic Signatures. U.S. Food & Drug Administration.
2. ISO (2016) ISO 13485:2016 – Medical devices – Quality management systems. International Organization for Standardization.
3. European Union (2017) EU Medical Device Regulation (EU MDR) 2017/745.
4. Microsoft Azure (2024) Azure Purview: Unified Data Governance Service. Microsoft Docs.
5. SAP (2023) Data Management and Compliance with SAP HANA. SAP Whitepaper.
6. McKinsey & Co (2023) Digital Quality Systems: Driving Operational and Regulatory Excellence.

Copyright: ©2025 Sayed Rafi Basheer. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.