Collibra

USE CASE

Your inspection window is 14 days.

Your data is spread across dozens of systems.

Using Collibra MCP to respond to FDA and EMA regulatory inquiries in life sciences

For most pharmaceutical and life sciences companies, assembling supporting evidence is a manual process. Clinical data lineage is reconstructed from memory and email threads. GxP asset ownership is called up system by system. ALCOA+ evidence is often compiled by hand from disconnected repositories. The result is a response that is slow, incomplete and difficult to defend.

- In 2025, FDA drug quality inspections reached 1,248, a 56% increase over the prior year. Warning letters issued by the Center for Drug Evaluation and Research (CDER) alone rose approximately 50% year over year
- Data integrity has remained the most cited category in FDA Form 483 observations for six consecutive years, now reaching 91 observations per 100 pharmaceutical inspections

The response timer starts immediately when a Form 483 arrives, an FDA reviewer submits a data clarification request, or a pre-approval inspection is announced. Sponsors typically have just 14 business days to respond. Evidence must be available, complete, traceable and defensible.

Collibra MCP changes the equation.

The Collibra Model Context Protocol server connects third-party LLMs including Claude, ChatGPT, Microsoft Copilot and others directly to governed metadata, data lineage, classification records and business context in the Collibra Platform. Regulatory teams can ask questions in plain language and receive audit-ready evidence assembled from the catalog automatically, rather than reconstructing system by system under deadline pressure.

Collibra © 2026 Collibra. All rights reserved. | collibra.com

A governed catalog does not just organize your data. It prepares your defense before the inspector arrives. Life sciences regulatory teams face many risks, but the following inspection risks are the most challenging:

- **Data integrity citations dominate inspections:** Sixty to eighty percent of FDA Warning Letter citations relate to data integrity in a broad sense. The most frequent observations include failure to maintain complete original records, attributability failures from shared credentials, disabled audit trails on laboratory instruments and re-testing to pass without documented rationale. Each of these is a governance failure
- **Form 483 response windows are unforgiving:** Sponsors have 14 business days to respond to a Form 483 observation. Without a governed infrastructure, reconstructing clinical data lineage from raw eCRF to final ADaM datasets manually across multiple trial sites, within the established timeframe, is operationally unsustainable
- **GxP asset ownership is inconsistently documented:** Pre-approval inspections require ALCOA+ evidence demonstrating that every GMP manufacturing data asset has a documented owner, complete lineage to a validated source system and a data contract referencing the applicable validation protocol
- **PHI classification is a weeks-long manual exercise:** IRB and Privacy Office requirements mandate that each of the 18 HIPAA PHI identifiers is classified per column across clinical datasets before research use. For a 50-million-row patient events dataset, manual classification is both slow and error-prone
- **FDA has adopted AI-assisted inspection targeting:** In 2025, FDA launched its Enterprise Language Support Assistant (ELSA), an AI system that analyzes internal FDA data to target inspections. Sponsors who cannot match the pace and precision of AI-assisted inspection with AI-assisted response are at a structural disadvantage

Collibra MCP for Life Sciences establishes four regulatory workflows:

PHI classification sweep across clinical data lakes: Before a research team can use a 50-million-row patient events dataset, the IRB and Privacy Office require a complete HIPAA PHI map, with all 18 PHI identifiers classified per column. Manual classification takes weeks.

With Collibra MCP, a privacy analyst instructs a connected LLM in plain language. The MCP server calls to enumerate every column for PHI assessment, retrieves the HIPAA PHI identifiers. The dataset is ready for the Institutional Review Board (IRB) review, with a complete, governed classification record, in hours rather than weeks.

21 CFR part 11 audit trail reconstruction: The FDA issues a Form 483 observation questioning an ADaM ADRS dataset. The sponsor must reconstruct complete clinical data lineage from raw eCRF to final dataset within 14 business days.

With Collibra MCP, a Data Engineer locates the cited dataset to confirm the CDISC version and validation status, and calls to trace the complete derivation path from raw eCRF to the final dataset. The Form 483, or 21 CFR Part 11, response package is assembled from governed records, not from reconstructed memory.

 © 2026 Collibra. All rights reserved. | collibra.com

GxP data integrity compliance for pre-approval inspection: An FDA pre-approval inspection requires ALCOA+ evidence for all GMP manufacturing data systems. The quality team must demonstrate that every data asset has documented ownership (Attributable), complete lineage (Legible and Original) and validated controls (Accurate), within 30 days of inspection notice.

With Collibra, the PAI evidence package is assembled from the catalog automatically.

Example prompt: "We have an FDA PAI inspection in 30 days. Audit all GMP manufacturing data assets against ALCOA+ criteria: confirm each has an owner, complete lineage trace, and validated data contract. Flag gaps."

EMA and FDA adverse event data lineage for NDA submission: An FDA reviewer issues a data clarification request on ISS adverse event incidence rates in an NDA submission. The sponsor must trace each ADAE row back to the source eCRF entry across all clinical trial sites, within the agency response window.

With Collibra MCP, a Data Engineer locates the ADAE dataset, confirms CDISC version and TLG reference through asset details and lineage to trace it. The Privacy Officer confirms the clinical data sharing agreement covers all sites and includes an audit trail provision. The FDA DCR response package is complete, with an AE traceability matrix generated from governed records.

Key outcomes with Collibra for Life Sciences

Regulatory response at the speed regulators now expect.

Collibra MCP transforms regulatory inspection response from a manual, deadline-driven scramble into a governed, repeatable, AI-assisted workflow. Clinical data lineage is already traced. PHI is already classified. GxP asset ownership is already documented. When a Form 483 arrives or a DCR lands, the evidence package is assembled from a governed catalog, in hours, not days. That is not just faster compliance. It is the regulatory readiness that an AI-powered FDA inspection environment now requires.

Collibra MCP and third-party LLMs

The Collibra MCP server provides a standardized access layer connecting third-party LLMs including Claude, ChatGPT, Microsoft Copilot and agents built on Databricks and Azure AI Foundry to governed metadata, lineage, classification records and business context in the Collibra platform. Natural language questions from regulatory, quality and data teams return governed, audit-ready answers drawn from the catalog, without requiring technical query skills or manual system-by-system searches.

Learn more:
collibra.com/products/mcp-server

Collibra © 2026 Collibra. All rights reserved. | collibra.com

Sources

1. IntuitionLabs, "FDA Form 483 and Warning Letter Statistics FY2016-FY2026." FDA conducted 1,248 drug quality assurance inspections in FY2025, a 56% increase over FY2024. https://intuitionlabs.ai/articles/fda-form-483-warning-letter-statistics

2. APSinco, "Common FDA 483 Observations for Data Integrity: Examples and Remediation in 2026." Data integrity has remained the most cited category for FDA citations for six consecutive years, reaching 91 observations per 100 pharmaceutical inspections. Remediation and production shutdowns for a mid-size manufacturer range from $3.3 million to $9.8 million. https://www.apsinco.com/post/common-fda-483-observations-for-data-integrity-examples-and-remediation-in-2026

3. BLA Regulatory, "FDA Inspection Observations and Form 483: Practical Insights." Form 483 observations include inconsistent data integrity, incomplete documentation, and procedural gaps among top-cited issues. FDA 483 Response: A Guide to Best Practices & Compliance | IntuitionLabshttps://bla-regulatory.com/pharmafocus-2025-fda-form-483/

4. Collibra Product Innovations, "Model Context Protocol (MCP) Server." Generally available November 2025. https://www.collibra.com/products/innovations

5. Collibra Marketplace, "Databricks MCP Marketplace: Collibra MCP Server." The MCP server addresses inconsistent context retrieval, fragmented integration workflows, and lack of standardized communication between Collibra and other information sources. https://marketplace.collibra.com/listings/databricks-mcp-marketplace-collibra-mcp-server/

6. PharmaDossier, "FDA Warning Letters by the Numbers 2026." FDA implemented a stricter escalation policy automatically escalating Form 483 data integrity findings to Warning Letters without extended grace periods for CAPA plans. https://pharmadossier.com/blog/fda-warning-letters-by-the-numbers-2026

Collibra © 2026 Collibra. All rights reserved. | collibra.com