

Precision in every page

AI-powered clinical document validation on AWS



Summary

A global life sciences technology company needed to streamline the quality validation of clinical trial documents before their archival in an electronic Trial Master File. Sonata designed and implemented an AI-powered document validation solution on AWS using Amazon Bedrock and Amazon Bedrock AgentCore. The solution automates signature, pagination, image-quality, and document-completeness checks, reducing document review time by 95% while achieving more than 99% accuracy in detecting pagination errors.

Customer overview

A global life sciences technology company that provides digital platforms for clinical trial management and regulated document workflows.

Pressure points

Clinical trial documents must pass stringent quality checks before they can be archived in an electronic Trial Master File

Clinical trial documents must pass stringent quality checks before they can be archived in an electronic Trial Master File

The customer’s eTMF platform processed more than 100,000 pages across 500 documents each month

Reviewers spent more than 900 hours monthly checking signatures, pagination, document completeness, and image quality

This manual process created operational bottlenecks, introduced inconsistent results, and contributed to pagination-related errors. The customer needed a scalable way to automate repetitive quality-control activities while strengthening audit readiness.

Solution highlights

Sonata designed and implemented an AI-powered document validation solution on AWS using Amazon Bedrock and Amazon Bedrock AgentCore. The solution applies multimodal AI, agentic orchestration, and rules-based validation to evaluate clinical documents across four key areas:

Signature validation
Identifies missing signatures, initials, digital approval markings, and incomplete signatory requirements

Pagination validation
Analyzes “Page X of Y” conventions, continuous numbering, missing pages, and out-of-sequence content

Image-quality assessment
Detects blurred, skewed, unreadable, or poor-quality document scans that may not meet archival standards

Document-completeness validation
Uses Amazon Bedrock and AgentCore to orchestrate validation agents that evaluate document structure, identify missing pages or content, consolidate validation results, and flag exceptions before documents enter the regulated archive.

Enterprise manageability
Improved deployment, updates, governance, and supportability through a modern add-in architecture aligned with Microsoft 365.

Documents are securely processed through an event-driven AWS architecture. Files are stored in Amazon S3, prepared for analysis using AWS Lambda and Amazon ECS on AWS Fargate, and evaluated using foundation models available through Amazon Bedrock. Amazon AgentCore manages and coordinates the AI-driven validation workflows. Validation results and audit records are retained in Amazon DynamoDB, while Amazon CloudWatch and AWS X-Ray provide monitoring and traceability.

Rather than requiring reviewers to inspect every page manually, the solution flags exceptions for review, allowing quality teams to focus on documents requiring attention. Following a successful two-week proof of concept, Sonata delivered the production-ready solution through a 12-week implementation.

Results that speak volumes

95% reduction in document review time

More than 99% accuracy in pagination-error detection

Reduced manual effort across repetitive quality-control activities

Scalable processing capacity of approximately 5,000 documents per day

Support for PDF, image, and Microsoft Office document formats

Improved audit readiness and consistency of document validation