

GenAlz Whitepaper: AI-Powered Holistic Quality Assurance and Control

A step change in pharmaceutical and biotechnology batch production and release

Executive Summary

Good Practice (GxP) leaders in Life Sciences need to manage information and insights across complex, complicated processes. They have spent billions of dollars on Knowledge Management Systems, trying to organize petabytes of structured and unstructured data into information that can improve the commercialization success of new therapies. Yet the success rate for too many of those therapies remains stubbornly low.

Enter the next wave of Life Sciences transformation, AI-powered Holistic Quality Assurance and Quality Control. Life Sciences organizations can exit the hamster wheel of data and knowledge unification, with an intelligence technology layer that starts with outcomes, not data, going to where the data already resides. Holistic insights are generated without sacrificing compliance, security, or data sovereignty.

Three key innovations are driving this transformation: AI Orchestration and Data Mesh, which keep data local while protecting security and privacy; Ethical human-in-the-loop AI for GxP processes that are as accessible, explainable and traceable; and Computer System Validation (CSV) adapted for AI and Life Sciences standards and regulations.

Customers have seen outstanding results from early implementations, including 12-14x faster GxP processes that are up to 100% accurate, saving over 50% on costs.

The GxP Data Challenge

Good Practice (GxP) process leaders, whether in Pharmaceuticals, Biotechnology or CDMOs, need to manage information and insights across complex, complicated processes:

- Tech Transfer

- Process Validation
- Manufacturing
- Quality Control
- Quality Assurance
- Regulatory Submission
- Batch Release
- Periodic Review
- Audits and more.

These processes create petabytes of data¹, an estimated 90% of life sciences data being unstructured². Most organizations have most (or all) of these types of data:

- Current process parameters
- Equipment data
- Validation data
- Risk assessments
- Regulatory context
- Product specifications
- Supplier data
- Process maps
- KPIs across all PQS elements
- Audit results
- Deviations
- Complaints
- Change control metrics
- Training compliance
- Supplier performance
- Deviation logs
- Batch records
- Equipment logs
- Training records
- Stability data

¹ AWS, Dell and Bio-IT World estimate the top 20 global pharma companies each could have anywhere from 10 to 100 petabytes of data.

² The "90% unstructured data" statement is repeated often in the industry, including in "Life Sciences Tech Trends" (Deloitte), "Life Sciences Insights" (IDC), and The Pistoia Alliance & Bio-IT World Case Studies.

- Environmental monitoring
- Specifications & methods
- Process parameters (CPPs, CQAs)
- Release testing results.

This data is further complicated when used to develop reporting and documentation, which creates additional data and meta data. PQE Group estimates the typical Life Sciences organization has almost all of these activities in their operations:

- Development reports
- Change control requests
- Audit reports
- Impact assessments
- Risk assessments
- Validation impact analyses
- Updated specifications and Standard Operating Procedures (SOP)
- Change implementation plans
- Mitigation plans
- Change closure reports
- Management review packages
- Quality metrics reports
- Site performance summaries
- Risk summaries
- Compliance dashboards
- Batch performance summaries
- Out-of-specification (OOS) summaries
- Knowledge summaries
- Process understanding documents
- Tech transfer packages
- Lessons learned reports
- Product knowledge file
- Annual Product Quality Review (APQR/PQR)
- Continued Process Verification (CPV) reports
- Stability trend reports
- Quality metrics dashboards
- Product quality trend analyses
- Deviation reports

- Investigation reports
- Root cause analysis summaries
- Corrective and Preventive Action (CAPA) plans
- CAPA effectiveness checks
- Trending reports for recurring issues.

All this data lives in multiple, often disconnected, systems:

"In many cases, batch data is still spread across Manufacturing Execution Systems (MES), Laboratory Information Management Systems (LIMS), and standalone Excel trackers, requiring manual stitching before release... Deviations still begin with a scramble for scattered evidence."³

"[Pharma 4.0] has brought data everywhere, but that doesn't mean people are better informed. Operational knowledge remains fragmented across SOPs, spreadsheets, and tribal habits that defy easy standardization."⁴

"Traditional facilities were 'often dependent on manual processes, isolated machines, and fragmented data systems.' These practices, while compliant, introduced inefficiencies, delays, and risk of human error."⁵

Most Life Sciences organizations have spent billions of dollars and significant time to implement and operate Knowledge Management Systems (KMS) tools across multiple processes. In one estimate, 2025 spending by Life Sciences on KMS and related systems was over \$5.6 billion⁶, growing at over 19% CAGR⁷. That was at least ¼ of the total amount spent on software by Life Sciences organizations⁸. Fortune estimates that the global quality control and compliance software market alone was \$10.8B in 2025, growing at 8.9% CAGR⁹. This includes customization of horizontal applications like document management (e.g. Microsoft SharePoint) or Enterprise Resource Planning (ERP), to vertical- or process-specific applications including material transport, sampling, quality control (QC), packaging,

³ Source: "Digital Factory Maturity in Pharma", Supply Chain Wizard.

⁴ "The Hidden Cost of Process Variability in Pharma", DeepHow.

⁵ Source: "AI Pharma Manufacturing: Autonomous GMP & Quality Control", Intuition Labs.

⁶ TechSciResearch, "Global Pharma Knowledge Management Software Market", January 2026.

⁷ Source: "Knowledge Management Software Market (2026-2035)", Market Research Future, June 30 2026.

⁸ Source: "Life Sciences Software Market Size and Companies – 2026-2035", Towards Healthcare, March 13 2026.

⁹ Source: "Pharma Quality Control and Compliance Market", Fortune Business Insights, June 22 2026.

upstream and downstream process control¹⁰, Process Analytical Technology (PAT) systems, Electronic Batch Record (EBR) systems, MES, Supervisory Control and Data Acquisition (SCADA) and Distributed Control Systems (DCS), and regulatory documentation and automation systems¹¹.

Layered on top are additional tools such as Learning Management Systems (LMS), Risk Management, LIMS, Quality Management Systems (QMS); and more. No surprise, these systems and the processes they help manage also create their own additional data, meta data and documentation as they feed into downstream Pharmaceutical Quality System elements:

- Process Performance & Product Quality Monitoring System,
- Management Review of Process Performance and Product Quality
- Change Management System
- CAPA

These knowledge assets, therefore, encompass a wide range of formats, including unstructured digital documents such as spreadsheets, word processing documents, presentations, scanned documents and images, videos, emails, meeting recordings, voicemails, and webinars; and analog documents such as laboratory notebooks, diagnostic images, certificates of analysis, and paper invoices.

Yet even after all that cost and effort most Life Science organizations struggle to digitize and unify knowledge so they can derive, integrate and manage meaning, predict the next best action, then safely automate those actions in a complex and shifting regulatory and compliance landscape. They consume enormous resources (budget, headcount, time), spending on average \$5B per approved drug¹², implying a continued large overhead of manual and inefficient processes. These reduced margins and commercialization success divert funds away from researching the next blockbuster therapy, resulting in “orphan drugs” that may have enormous clinical benefit to targeted populations but are too costly or niche to commercialize.

¹⁰ Source: "Robots in Biomanufacturing: A Road Map for Automation of Biopharmaceutical Operations", BioProcess International, June 23 2022.

¹¹ Source: "Implementing Lab Workflow Automation in the Pharma Industry", Janea Systems, May 7, 2025.

¹² Source: Nature Reviews Drug Discovery, August 6 2024

Statistics point to where part of the problem lies: inefficient, manual processes and limited success in digitization of those processes. For example, 4.5% of Life Sciences revenues are going to IT spend, of which 22.5% is just for software. An additional 33% of those revenues are spent on payroll. That's 7x what was spent on software. It is also estimated that 12% of those workforce tasks can be automated by AI.

Unlike prior software that forced people and processes into unnatural acts of digitization, AI meets people – and their data -- where they are: in day-to-day workflows are rely on complex and nuanced language, evidence and traceability.

How AI and Intelligent Automation can leapfrog Digital Transformation

By understanding an organization's digital knowledge management path to proficiency, we can begin to grasp how Life Sciences organizations got to the current state, and where they can go from here.

1. **Wave 0: "manual"**. Most data are not digital or structured, and reports are manually generated and recorded.
2. **Wave 1.0: "digitization"**. Digitize what can be digitized, so that information is more easily managed, and (hopefully) accessible. Structured databases may speed some processes, but most processes remain essentially unchanged, with one major exception: new digital tasks are tacked onto existing processes, increasing the effort of front-line staff without significantly improving their productivity (the "digital productivity gap"¹³).
3. **Wave 2.0: "help users find information"**. More advanced (and expensive) tools such as KMS and data lakes make it easier to digitize, structure, integrate, report and analyze data. Older data sets and processes are transformed where possible, and newer data sets and processes are born "digital first". Yet the digital productivity gap remains while underlying processes remain essentially unchanged.
4. **Wave 3.0: "help professionals decide what matters"**. Wave 3.0 shifts the focus from knowledge to insights. From a "digital knowledge user" paradigm to an "smarter professional" paradigm that leads with human outcomes.

¹³ ["Workforce Utilization Benchmarks Report, 2H 2025"](#), Activtrak 2025.

The potential for Wave 3.0 is clear. Pharmaceutical companies could bring therapies to market more than 500 days faster and reduce costs of development by 25 percent by implementing a comprehensive set of levers, including—crucially—automation.”¹⁴ Moreover, Process Manufacturers expect an average of 12% in total annual plant operating cost savings from digital and AI initiatives within the next three years.¹⁵

By looking at the impact of automation and AI on R&D, we can intuit the potential for pharmaceutical and biotechnology manufacturing as well: 53% of surveyed organizations saw increased laboratory throughput, 45% saw a reduction in human error, 30% achieved greater cost efficiencies, and 27% noted faster therapy discovery, when applying tools such as AI and automation.¹⁶

AIHQ also holds great promise to overcome the historically poor ROI of AI projects in the enterprise. Gartner estimates that 60% of AI Projects will be abandoned due to lack of AI-ready data¹⁷, while almost half of AI projects are scrapped between proof of concept and broad adoption¹⁸. These cross-industry failure rates may be even higher in Life Sciences, when one layers on GxP requirements, FDA expectations, and EU AI Act compliance.¹⁹

There are three interconnected technology innovations, which we call AI-powered Holistic Quality Assurance and Control (“AIHQ”), that are driving this third wave:

1. **AI Orchestration (Data Mesh) for Life Sciences.** AI providers to the enterprise can now include data federation and orchestration between AI models (regardless of where they come from, internally or externally) that can respect data security and privacy rules and avoid data delocalization. By not prompting 3rd-party models directly, data mesh and orchestration together dramatically increase the power of insights and automations, while simultaneously protecting results and reducing data and IP risk. In addition, orchestration allows simultaneous use of models via workflow management, making the entire process more efficient than single-threaded processes using multi-purpose models. For example, Large Language Models (LLM) may not be the best model to use for calculation or data retrieval.

¹⁴ Source: "From Bench to Bedside: Transforming R&D Labs Through Automation", McKinsey, March 8 2023.

¹⁵ Source: "State of digital in process manufacturing: AI exploration rises as companies target 12% savings", IoT Analytics, June 3 2026.

¹⁶ Source: "[Pharma's R&D Lab of the Future](#)", Deloitte Center for Health Solutions, July 29 2025.

¹⁷ Source: Gartner, Press Release, February 26 2025.

¹⁸ Source: S&P Global Market Intelligence, 2025.

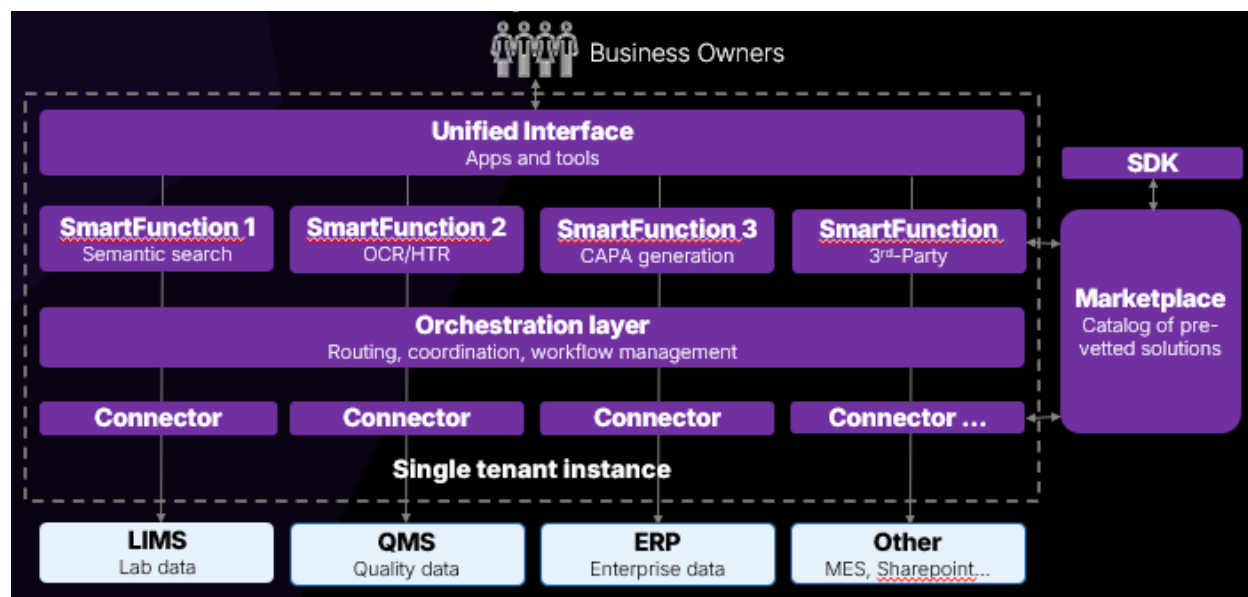
¹⁹ Source: USDM, "Why AI Adoption Is Failing in Life Sciences", December 30 2025.

2. **Ethical AI.** In GxP processes, “ethical” means auditability and human-centered. Using “white-box first” designs, AI can be as accessible, explainable and traceable as other software used in Life Sciences. This Ethical AI also puts people first, moving from a “software user” lens to an empowered Human-in-the-Loop that assures AI performs as intended and keeps final decision authority with trained professionals.
3. **Computer System Validation (CSV) for Life Sciences.** CSV frameworks have now been adapted for AI to operate safely within Life Sciences regulations and best practices. This ensures that every component, including data pipelines, AI enabled functions, and report generation logic, is validated, monitored, and maintained in a state of control. With embedded data and model governance, human-in-the-loop oversight, and continuous monitoring for drift or anomalies, the solution delivers both operational efficiency and regulatory defensibility.

The Intelligence Layer for the future of Life Sciences

Figure 1 below provides an overview of the architecture needed to create a unified intelligence layer for the future of Life Sciences: one that is ethical and compliant to the exacting standards and regulations of the industry.

Figure 1: Reference platform for AIHQ



Any Data Source: AIHQ should be able to connect with data from any source, in its source format. Data should never leave the confines of the organization’s protected systems to avoid risk of leaking valuable IP. Data prep or cleansing

should be kept to a minimum, in order to reduce cost, complexity, bias, and opaqueness.

Platform as Intermediary: Containerized, orchestrated AI-powered functions bring together protected data with intelligent automation solutions that address specific use cases. These functions live in a single tenant architecture, which is designed to work within an organization's existing GxP processes and data security, privacy and sovereignty frameworks. Every step is auditable, and compliant with the most rigorous industry standards. And tools such as a Software Development Kit (SDK) ensure the platform is flexible and extensible to meet new and unexpected challenges.

Intelligent Automation Applications: Models, whether developed internally or sourced from third-parties, can be applied to the organizations' data to identify patterns and next-best actions, automate appropriate activities, and accelerate GxP processes. The platform and the business professionals together determine appropriate next steps, minimizing the significant risks inherent in autonomous agentic AI²⁰.

A smart AIHQ architecture relies on off-the-shelf, predefined combinations of connectors, AI models ("smart functions"), process triggers and notifications that target common user workflows. Rather than risky Agentic AI, Intelligent Automation Applications start with Containerized AI and analytics capabilities packaged as common objects. This could include Natural Language Processing (NLP), Open Information Extraction, Neural Networks, Machine Learning, Reinforced Learning, Statistics and Clustering models, and Custom Semantic Search Engine – essentially any AI-powered action that can meet the data where it stands, and how it is.

These common objects and processes can be further enhanced with an SDK to address new, novel or unexpected situations. Customers, content creators, and third parties can directly combine and customize smart functions themselves, ensuring solution scalability.

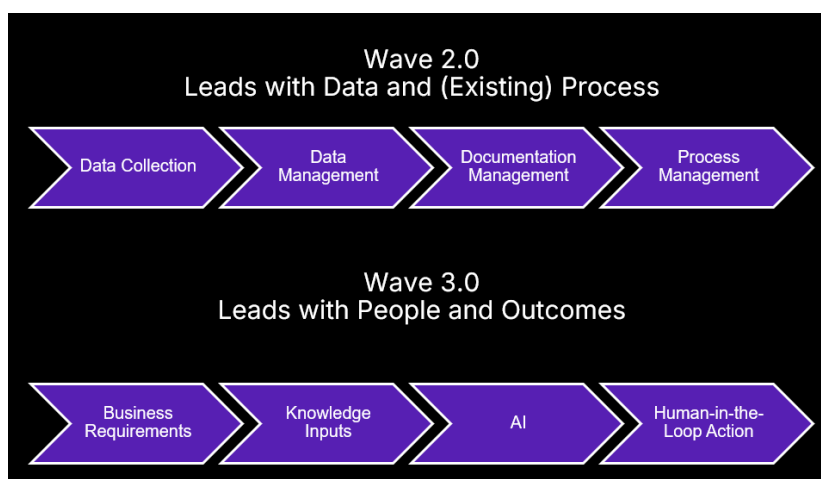
The process triggers and notifications are also critical. These "human in the loop" capabilities ensure the AI never goes rogue: both require a knowledgeable professional, who has experience operating in, or leading, the relevant business processes, to authorize action before it can be taken.

²⁰ ["Can Autonomous AI Agents Go Rogue? Why Goal Misalignment is a Real Security Threat"](#), NOMA, December 3 2025.

Holistic Insights and Actions in the real world: sample use cases

The paradigm shift of AIHQ naturally leads to a different way to look at the problem of data, intelligence, insights and actions in GxP processes:

Figure 2: AIHQ Paradigm Shift



Wave 2.0 leads with a data and process lens. The very first activity is data collection. Whether in paper-based templates or software user interfaces, the paradigm dehumanizes professionals into “users” who need to adapt to those input forms. That leads to a “management” mindset, where data and the documents have to be managed before the process itself can be managed. Note the emphasis on “management” vs. creative problem-solving.

Wave 3.0 starts with a business- and people-first mindset: what is the business problem we are trying to solve? (e.g.: annual or periodic quality reviews, fix recurring deviations, validate our systems, etc.). This question drives the knowledge inputs that will be needed to research and investigate potential solutions, including not just data but the systems and documents that may hold the clues we seek. That investigative process naturally leads to the true power of AI to extract knowledge from those sources, make comparisons, identify patterns, summarize what it finds, harmonize findings across siloes, and even predict the right actions to take. Which then leads to AI-enhanced human decision-making and action-taking, where professionals are not replaced by mindless agents, but augmented in their capacity to move the business forward.

Table 1 provides just a few examples of where Life Sciences organizations can apply AIHQ:

Table 1: AIHQ Use Cases for AI Transformation of Life Sciences Lifecycle Stages

Lifecycle Stage	Requirement	Knowledge Input	AI-powered Use Case
Pharmaceutical Development	Common technical document redaction	Design and development data	Literature Review: Sematic, contextual search filters essential research from thousands of scientific papers with neural network matching tuned for biomedical purposes.
Technology Transfer	Tech transfer package build	Scale-up activities data	Tech Transfer: Ensure all appropriate documents have been prepared and are complete and accurate.
Commercial Manufacturing	Metrics & trend analysis, QMR, periodic review, audit	Drug substance, drug product, equipment and facilities established conditions	Periodic Review: Automate drafting of APQR documents for accelerated and accurate repetitive reporting of already QC'ed data.
Product Discontinuation	Metrics & trend analysis, periodic review	Market data	Pharmacovigilance: Accelerate PV related investigation process with risk assessment and traceability.

Table 2: AIHQ Use Cases for Life Sciences Quality System Transformation

Quality System Element	Requirement	Knowledge Input	Use Case
Change Management	Science & risk assessment	Process, equipment, product, validation, supplier data	Change Management: Standardize impact assessment with improved accuracy and preserved knowledge.
Process Performance & Product Quality Monitoring	Parameter and attribute control analysis, identification of sources of variability	Process, equipment, product, validation, supplier data	Regulatory Submission: Consolidate and verify scattered data for country-specific completeness.
Management Review of Process Performance & Product Quality	Continuous improvement indications, program effectiveness checks	Product quality and stability data, process performance data	QMR: Accurate, compliant, and efficient tracking of anomalies and exceptions.
CAPA	Science & risk assessment, action identification	Quality events, equipment, product, process, supplier data	CAPA: Automate deviation and CAPA management, standardize investigations, improve root cause accuracy, and preserve knowledge

Table 3: AIHQ Use Cases for AI Transformation of Life Science Quality Enablers

Quality System Enabler	Requirement	Knowledge Input	Use Case
Knowledge Management	Metrics & trend analysis reports, performance summaries, periodic reviews	Quality events, equipment, product, process, validation, training, supplier data	Validation: Holistic approach to validation studies with a clear view of established conditions and accurate risk-based approach.
Quality Risk Management	Quality system performance review package, remediation & audit action identification	Quality events, equipment, product, process, validation, training, supplier data	APRQ Report Automation (see below)

Let's dive deeper into one specific example: **Annual Product Quality Review (APQR) report automation.**

Because AIHQ starts with people and outcomes, the APQR-enhanced process begins with the desired outcomes of the annual product quality review process, not (just) generating a report. The desired outcomes could be reducing persistent quality issues in a particular batch process. That outcome dictates the knowledge inputs needed to successfully execute on that requirement. Business analysts and an intelligent platform then translate those requirements into the data Connectors needed to securely and compliantly gather the data from multiple siloed systems. The AI data mesh pulls in the right data, and orchestration then pulls in specific Intelligent Automation Applications that standardize and reconcile these inputs, perform trend analysis, and generate a complete APQR report with full traceability and audit ready documentation. The human-in-the-loop AI then recommends next steps for publishing the report, as well as taking the next best actions identified that will improve the persistent quality issue.

In short, with AIHQ, APQR moves from a “we need to generate a report” problem to “how do I solve this persistent batch quality” problem. This human- and business-first reframing was not possible before AI, when terabytes of structured and unstructured data across multiple systems and repositories required dozens of people dozens of weeks to just comb through the information.

AIHQ takes the data where it is, and how it is, and eliminates the manual burden of extracting, cleaning, and reconciling it across disparate systems and formats. It ensures consistent application of quality criteria, standardized reporting across products and sites, and seamless integration with existing QMS and data sources. The result is not just a faster, more reliable APQR process that reduces cycle times, minimizes errors and omissions, but pushes the business forward to enhance quality and compliance, reduce operational risk, and free up highly competent (and paid) teams to focus on higher value activities.

Next Steps

AIHQ is not just a new tool or set of tools to improve existing Life Sciences quality. It is a paradigm shift that can fundamentally change the ability of life sciences organizations to turn quality from a compliance exercise to a strategic advantage.

Knowledge management shifts away from mundane tasks to business outcomes. Professionals and teams focus on creatively solving important business challenges and identify opportunities that can truly move the business forward: improve efficiency, accuracy, and compliance, enhance product quality and patient safety, and more successfully commercialize promising therapies.

Forward-thinking Life Sciences organizations can start today to take full advantage of the benefits of AIHQ.:

They can start by reconsidering AI not (just) as an efficiency gain, but as a powerful augmentation tool that can enhance human judgment and creativity to solve problems.

They can then shift from manual data and knowledge management tasks to a flexible, orchestrated AI architecture that takes advantage of three key innovations:

1. AI Orchestration and Data Mesh that begins with, not bolts on, a deep respect and understanding of your data security and privacy rules.
2. Ethical AI that is human-centered, auditable, accessible, explainable and traceable.
3. AI that embraces Computer System Validation (CSV) for Life Sciences.

And lastly, they should work with a trusted partner who understands how to apply AI within the context of the complex and ever-shifting regulations, practices and policies of pharmaceutical, biotechnology and related Life Sciences manufacturing environments.