



INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

ISPE GAMP AI Guide: Validation Framework for GxP Systems

Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-02-08 · ispe gamp · ai validation · gxp compliance · machine learning · computer software assurance · data integrity · regulatory affairs · pharma ai · risk management · csa

Original article: <https://intuitionlabs.ai/articles/ispe-gamp-ai-validation-guide-gxp>



In this article

1. Executive Summary
 2. Introduction
 3. ISPE GAMP® Guide: Artificial Intelligence — Overview and Background
 4. Regulatory and Standards Landscape for AI in GxP
 5. Challenges in Validating AI-Enabled Computerized Systems
 6. Key Concepts in AI System Validation for GxP
 7. Case Studies and Examples
 8. Implications and Future Directions
 9. Conclusion
-

Executive Summary

The International Society for Pharmaceutical Engineering (ISPE) has released the **GAMP® Guide: Artificial Intelligence** (July 2025) to address a critical gap in industry guidance: how to *validate AI- and machine learning-enabled computerized systems in GxP-regulated environments*. This new 290-page guide provides a **holistic, risk-based framework** for developing, implementing, and overseeing AI systems such that patient safety, product quality, and data integrity remain paramount (ispe.org) (ispe.org). It explicitly builds on ISPE's established GAMP 5 principles and data integrity guidelines, extending them to AI's unique characteristics (ispe.org). The guide tackles known industry challenges (e.g. pilot-phase failures, lack of common language, fragmented data, evolving regs) by prescribing best practices at each life-cycle stage — from ideation through operation — and by defining roles, data and model governance, risk management, and change control specific to AI (ispe.org) (ispe.org).

In practice, **validating AI in GxP** includes assessing data quality, model development, intended use, and performance over time. Distinguish frozen models from models that learn in production: a frozen model does not adapt during use, while changing deployment data can make performance monitoring important ([pharmacovigilance validation paper](#); [EMA reflection paper](#)). Determine and document Part 11 applicability for specific records from the relevant predicate rules, actual reliance on electronic records, and electronic submission requirements; AI use alone does not establish applicability ([FDA scope guidance](#)). Retain GxP-relevant AI evidence according to its purpose, applicable requirements, and documented risk. Internal data-integrity, version-control, and logging measures can support assurance without making every prompt or output a legally regulated record or requiring a signature for every interaction. FDA's final February 2026 CSA guidance addresses software used in medical-device production or quality management systems ([FDA CSA guidance](#)). A proportionate assurance approach may inform pharmaceutical practice, but applicable validation and documentation duties remain. Recommended planning includes defining the AI's **context of use**, performance criteria, testing, supplier assessment, [change-control](#) and monitoring.

This report reviews the ISPE GAMP AI guidance in context: it traces its development, outlines the evolving regulatory landscape ([FDA draft guidance on AI](#), EU Annex 11/AI Act, ICH Q9, etc.), and examines how the new GAMP framework aligns with existing standards. We analyze challenges with AI validation (e.g. black-box models, data drift, bias), and contrast traditional CSV approaches with AI-appropriate methods. Example use

cases (from industry surveys, expert articles, and vendor case studies) illustrate practical application of AI in manufacturing, clinical trials, and quality systems — highlighting both productivity gains and compliance concerns. Finally, we discuss implications for organizations: the need for new governance structures, staff training, continuous monitoring, and future regulatory developments. Throughout, all key points are supported by recent industry and regulatory sources (www.pwc.com) (ispe.org).

Introduction

The life sciences industry is undergoing a **digital transformation** driven by artificial intelligence (AI) and machine learning (ML). AI applications now span drug discovery, process development, manufacturing, and quality systems. Generative AI pilots and scaling efforts feature in surveys of selected industry executives. Bain's February 12, 2024 release reported that nearly 60% of surveyed pharmaceutical executives had moved beyond ideation to building use cases ([Bain survey release](#)). Separately, McKinsey surveyed more than 100 pharma and medtech leaders responsible for their organizations' generative AI efforts in late summer 2024; all respondents had experimented with the technology, and 32% reported taking steps to scale it ([McKinsey survey](#)). These surveys measure different activities in different respondent populations; they do not establish adoption rates for all pharmaceutical companies.

While AI promises greater efficiency and insight, its integration into Good Manufacturing Practice (GMP, or "GxP") environments introduces new **validation challenges**. Traditional computerized systems were typically rule-based and static: once programmed and validated via installation/operational qualifications, they produced reproducible outputs under strict controls (www.ey.com) (www.bioprocessonline.com). By contrast, modern AI systems — especially models like neural networks or large language models (LLMs) — learn during training; frozen models do not adapt in production, whereas dynamic models can update through a defined learning process ([validation paper](#)). Their behavior can be probabilistic or opaque ("black box"), and small changes in input can lead to markedly different outputs (intuitionlabs.ai) (www.ey.com). It follows that **validation approaches need to address the particular model and deployment mode** as AI becomes embedded in GxP processes.

Regulatory expectations are accordingly evolving. In the U.S., 21 CFR Part 11 applicability depends on the specific predicate-rule record, electronic reliance, and submission requirements ([FDA scope guidance](#)). EU GMP Annex 11 distinguishes application validation from IT infrastructure qualification and calls for documented risk-based controls ([Annex 11](#)). FDA's final February 2026 **Computer Software Assurance (CSA)** guidance concerns medical-device production and quality management system software; it does not mandate replacing pharmaceutical CSV ([FDA CSA guidance](#)). Similar risk-based thinking underlies ICH Q9(R1) on Quality Risk Management. EMA's AI reflection paper was adopted by CHMP and CVMP in September 2024, following the 2023 consultation draft ([EMA adoption information](#); [final reflection paper](#)). It emphasizes a human-centric approach within existing legal requirements. Other standards, such as ISO/IEC 42001:2023, and the enacted EU AI Act further underscore the need for **governance of AI technologies**; the Act entered into force on August 1, 2024, with phased application dates ([European Commission timeline](#)).

The culmination of these trends is the ISPE GAMP® AI Guide, which provides **the first industry-wide best practice framework** tailored to AI in GxP. The guide emerges at a time when many pharmaceutical AI projects are "stuck in pilot" or failing to reach full implementation (ispe.org). By formally defining how to design,

develop, test, and maintain AI systems under GxP, ISPE seeks to turn AI innovation into a reliable, compliant reality (ispe.org) (ispe.org). In the sections below, we examine how this guidance fits into the broader data-integrity and regulatory landscape, the specific validation approaches it recommends, and the implications for organizations seeking to harness AI without compromising compliance.

ISPE GAMP® Guide: Artificial Intelligence — Overview and Background

Scope and Development of the GAMP AI Guide

ISPE's GAMP guidance has long provided an industry consensus on computerized system validation. The new **GAMP® Guide: Artificial Intelligence** (published July 2025) is the *first comprehensive GAMP guidance focused exclusively on AI applications* in regulated life sciences (ispe.org) (ispe.org). Conceived by an international team of over 20 industry and academic experts (co-leads Brandi Stockton, Eric Staib, Martin Heitmann, and others), the Guide was developed under ISPE's GAMP Software Automation & AI Special Interest Group and AI Community of Practice. It draws on member experiences and peer review, reflecting the collective learning of early adopters who had inundated the GAMP CoP with questions about AI validation (ispe.org) (ispe.org).

The Guide explicitly **builds on prior GAMP materials and data-integrity guides**. It leverages the foundational second edition of *ISPE GAMP 5: A Risk-Based Approach to Compliant GxP Computerized Systems* and its Appendix D11 (which first introduced AI/ML concepts) (ispe.org). It also incorporates key principles from ISPE's **Records and Data Integrity** guidance. In fact, GAMP AI is positioned as an “industry response” to accelerating regulatory discussion papers on AI from FDA and EU authorities (ispe.org). It even references the novel ISO/IEC 42001:2023 “AI management systems” standard, indicating its alignment with the very latest frameworks (ispe.org).

In practice, the ISPE AI Guide serves as a **common reference** for all stakeholders — regulated companies, contract organizations, software vendors, and regulators themselves. Its stated goal is to facilitate **effective and efficient AI use while safeguarding patient safety, product quality, and data integrity** (ispe.org) (ispe.org). Phrased differently, the guide seeks to allow innovation *with compliance*. The framework is designed to be scalable and risk-based, so that high-risk AI applications (e.g. impacting patient safety) get the most scrutiny, while lower-risk projects can proceed with leaner controls. Key themes include partnering between sponsor and supplier, establishing a common language for AI projects, strengthening knowledge management (to build AI literacy), and ensuring inspection readiness even for advanced systems (ispe.org) (ispe.org).

Key Themes and Concepts in the GAMP AI Guide

According to ISPE's summary, the GAMP AI Guide introduces a number of **new concepts and extensions** to traditional GAMP thinking (ispe.org) (ispe.org):

- **Enhanced Risk Management:** As always in GAMP, *Quality Risk Management* (ICH Q9) is central. The AI Guide specifies AI-specific risk considerations (e.g. data bias, algorithmic errors, model drift) and outlines how to identify and mitigate them. It emphasizes a science-based QRM approach that continuously assesses the AI's influence on patient safety and product quality (ispe.org).

- **Scalable Lifecycle Activities:** The Guide adapts the GAMP “V-model” life cycle to AI projects. It encourages a **risk-tiered, fit-for-purpose approach**: for example, a mature company with AI experience might use streamlined templates, while a less-experienced team might need more extensive documentation. Importantly, the Guide covers all phases – concept, design, development, operation, and retirement – with attention to AI-specific tasks (see next bullets) (www.scribd.com) (www.scribd.com).
- **Roles and Responsibilities:** New or augmented roles are defined for AI projects. Besides traditional roles, the Guide highlights the need for AI data scientists, ML engineers, and AI risk managers to collaborate with quality units. It also introduces roles like “Data Steward” or “Model Owner” whose responsibility is to ensure data/model governance. The concept of a Quality Assurance Unit (QAU) is preserved and expanded: the QAU must ensure oversight of AI decisions and sign-off on AI validation outcomes (ispe.org).
- **Data and Model Governance:** A central focus is on data quality and fit-for-purpose data. The Guide insists that *data are the foundation* (“backbone”) of AI systems (ispe.org). It introduces a broad data governance framework (Appendix M7) that covers data collection, cleaning, augmentation, and provenance. Equally important is **model governance**: version control, performance metrics, and documentation requirements for model development and deployment are addressed. Transparency (explainability) and change-control for models are emphasized (ispe.org).
- **Knowledge Management:** Appendix M5 addresses “AI literacy” and training. Industry practitioners noted a *knowledge gap*: to use AI tools responsibly, operators must understand AI concepts. The Guide therefore gives strategies for training, documentation of “know-how,” and means of capturing lessons learned.
- **Explainable and Trustworthy AI:** The Guide promotes *explainable AI* (XAI) techniques so that users can interpret model outputs in GxP contexts. Likewise, it advocates the principles of “Trustworthy AI” (transparency, fairness, human oversight) as key to corporate governance. These concepts reflect broader societal standards being integrated into GxP teams.
- **Appendices for AI-specific Guidance:** Beyond concepts, the Guide provides **detailed appendices** on technical areas: life cycle scenarios, AI model testing methods, continuous monitoring and CAPA management for AI incidents, cybersecurity in AI contexts (e.g. adversarial attacks on models), and even fit-for-purpose computational infrastructure (cloud security, AI pipelines). This level of granularity far exceeds existing CSV guidance. In effect, each appendix (M1–M8, etc.) expands a portion of the life cycle with AI-centric content. For example, Appendix M7 is entirely on data/model management; M8 on IT infrastructure; M6 on data quality.
- **Alignment with Regulations:** The Guide does not replace regulations but explains how to meet them with AI. It cross-references EU and FDA initiatives (e.g. FDA’s draft AI Device guidance, EMA AI reflection paper) and ties them to GAMP practices (ispe.org). The intent is that compliance officials and auditors can use the Guide to interpret how GxP rules apply in AI cases.

In summary, the GAMP AI Guide provides both **philosophy and practical detail**. It reiterates classic GxP tenets (most importantly, risk-based thinking and patient safety) while integrating new layers for AI (data, algorithms, adaptivity, ethics). As one of the guide's co-authors stated, it "serves as a reference point for stakeholders...on best practices in developing, implementing, and using AI-enabled systems" (ispe.org), enabling innovation in a controlled way.

Regulatory and Standards Landscape for AI in GxP

Existing GxP Regulations and Guidelines

AI systems used in regulated pharmaceutical and related environments are subject to **all standard GxP requirements**. Key among these are:

- **21 CFR Part 11 (FDA)** – Determine and document whether each relevant AI-related record falls within Part 11 by assessing predicate-rule requirements, whether electronic records are relied on instead of or alongside paper, and applicable electronic submission requirements ([FDA scope guidance](#)). Records merely used to generate a submission are not automatically Part 11 records. Apply required record and signature controls to the records and signings that are in scope; retaining prompts, datasets, model versions, and outputs can also be an internal, risk-based assurance measure. FDA recommends basing audit-trail decisions and record retention on predicate rules and documented risk, rather than recording every AI interaction as an automatic legal duty. FDA's final February 2026 CSA guidance addresses medical-device production and quality management system software ([FDA CSA guidance](#)); its scope should not be generalized to all pharmaceutical GxP.
- **EU GMP Annex 11 (EU)** – Applies to computerized systems used in GMP activities. The application should be validated and IT infrastructure qualified. The extent of validation and data-integrity controls should follow a justified, documented risk assessment, with lifecycle evidence supporting the selected methods and acceptance criteria ([Annex 11](#)). IQ/OQ/PQ can be an implementation strategy, but Annex 11 does not prescribe that sequence for every computerized system. The Commission's consultation on proposed revisions and a new Annex 22 opened on July 7, 2025 and closed on October 7, 2025; as checked on October 5, 2026, its page records the response period as ended ([Commission consultation](#)). The [2025 draft Annex 22](#) proposes a narrower scope than the wider ISPE framework: static AI/ML models with deterministic outputs in critical GMP applications. It excludes dynamic models, probabilistic-output models, generative AI and LLMs, and states that these should not be used in critical GMP applications. These are provisions of the consultation draft, not a statement that that draft is applicable law.
- **ISO/IEC 17025 and ISO 13485:2016 — assess applicability** – [ISO/IEC 17025](#) supports organizations performing testing, sampling or calibration in obtaining reliable results. Assess the organization's applicable regulatory, accreditation and contractual obligations before selecting conformity requirements. For medical-device quality management, assess the applicable ISO 13485:2016 requirements under [FDA's QMSR](#), including the manufacturer's and product's scope and any

exemptions. FDA identifies certain device types as exempt from CGMP requirements. Distinguish organizational laboratory competence and quality-system conformity from validation of the AI software for its intended use.

- **ICH Q9(R1) – Quality Risk Management** – This international guideline (adopted by FDA, EMA, others) sets the expectations for risk management across the product lifecycle. It does not mention AI explicitly, but its core idea – that risk assessment should focus on patient safety and product quality – is the foundation of all computerized system validation. ISPE’s GAMP AI guidance explicitly integrates QRM: it promotes **integrating AI-specific risks** into the QRM process, and using “science-based” risk management at every phase ([ispe.org](https://www.ispe.org)). For example, human safety systems, patient data analysis, or decisions that affect dose must be identified as high-risk and controlled accordingly.
- **Data Integrity Guidelines (ALCOA+)** – Regulatory agencies (FDA, EMA, MHRA, PIC/S) have underscored the “ALCOA+” principles: records must be **Attributable, Legible, Contemporaneous, Original, Accurate**, plus Complete, Consistent, Enduring, and Available. For AI systems, data lineage, transformations, model versions and retained outputs warrant controls proportionate to their GxP purpose and risk. EMA recommends extending governance SOPs to AI/ML data, models and algorithms in high regulatory-impact or high patient-risk uses ([EMA reflection paper](#)). This assurance practice is distinct from assigning Part 11 status or a signature requirement to every prompt or output ([FDA scope guidance](#)).
- **EU AI Act (enacted)** – Regulation (EU) 2024/1689 entered into force on August 1, 2024. The [Commission’s current timeline](#) distinguishes that date from phased application: initial prohibited practices and AI literacy obligations from February 2, 2025; governance and general-purpose AI obligations from August 2, 2025; and general applicability from August 2, 2026, including transparency rules. Following the 2026 AI Omnibus changes, Annex III high-risk rules apply from December 2, 2027, and high-risk systems embedded in Annex I regulated products from August 2, 2028. Assess the particular use case, risk classification, and provider/deployer role; pharmaceutical use alone does not establish high-risk status.
- **Other regulatory guidances** – Several FDA documents, though not legally binding, point to how AI should be approached:
 - *FDA AI/ML-Based Software as a Medical Device Action Plan (January 12, 2021)* – An action plan proposing further regulatory work, not a draft guidance ([FDA announcement](#)). For subsequent device change-control recommendations, see *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions*, final guidance issued August 18, 2025 ([FDA final guidance](#)). It addresses AI-enabled device software functions that manufacturers intend to modify over time.
 - *FDA Discussion Paper on AI/ML in Drug/Biologics (2023)* – Addresses AI-enabled processes in drug development, endorsing practices like transparency and post-market monitoring.
 - *EMA reflection paper on AI in the medicinal product lifecycle* – The final paper was adopted by CHMP and CVMP in September 2024 after the 2023 consultation draft ([EMA adoption information](#); [final paper](#)). It discusses data and algorithm governance and a human-centric approach within existing legal

requirements.

- *MHRA GxP Data Integrity (UK)* – The UK's authority has clarified ALCOA+ requirements, which apply equally to AI. These guidance documents underscore many points now in GAMP AI. For instance, the reflection paper emphasizes a human-centric approach and compliance with existing laws even for AI development ([EMA final reflection paper](#)), echoing GAMP's instruction to integrate AI into the standard GxP framework.

In sum, **AI use does not remove applicable GxP duties**. Determine which rules on validation, data integrity, and quality apply to the regulated activity. What has changed is how one meets those rules in the face of AI's novel features. The new GAMP Guide helps bridge that gap by translating general principles (e.g. risk management, validation life cycle, ALCOA+) into an AI context.

Regulatory Expectations on AI Validation

Regulators have signaled that AI's novelty requires enhanced scrutiny on certain points. Key expectations include:

- **Contextual Risk Assessment:** Authorities expect that companies **define the intended use of the AI system** and the associated risk to patients or product quality. The concept of "Context of Use" (COU) has been emphasized at FDA and in industry guidance ([intuitionlabs.ai](#)). This means, for each AI tool, document exactly what it will do and how errors could impact safety or efficacy. For example, an AI that flags out-of-spec batches bears high risk, while one that suggests suppliers may be lower risk. This COU analysis then drives how rigorously the tool must be validated ([www.bioprocessonline.com](#)).
- **Data Quality and Management:** Establish risk-based governance for AI data, models and algorithms, especially for high patient-risk or high regulatory-impact applications ([EMA reflection paper](#)). Document data provenance and relevant transformations. Determine separately which retained records and electronic submissions are within Part 11, rather than classifying dynamically generated AI outputs as regulated records merely because they are AI-generated ([FDA scope guidance](#)).
- **Transparency and Explainability:** Regulatory bodies — particularly EMA — have noted that AI's "black-box" nature can obscure critical information ([www.ema.europa.eu](#)). They expect, where possible, that the decision logic be interpretable by experts. GAMP AI therefore emphasizes documenting algorithms, choice of hyperparameters, and intended outputs so that validation evidence is understandable. Explaining how an AI reached a decision (even partially) can be important for audit readiness.
- **Change Control and Monitoring:** Distinguish frozen models, controlled retraining, and online learning. A frozen model does not update its parameters from routine inference data; retraining or an enabled learning process changes the model. Performance can nevertheless degrade under changing deployment conditions. Monitor performance and define acceptable thresholds and reassessment triggers ([pharmacovigilance validation paper](#); [EMA reflection paper](#)).

- **Vendor and Supply Chain Oversight:** When AI tools are acquired (e.g. third-party algorithms, cloud AI services), regulators expect users to rigorously qualify these suppliers. The GAMP Guide underscores that *ultimate responsibility lies with the regulated company*, even if an AI vendor provides the model. Therefore, quality agreements, audits, and technical evaluations of vendor validation practices are expected (intuitionlabs.ai).

Table 1 below summarizes major regulations and how they relate to AI validation in GxP:

Table 1: Key Regulatory Standards and Their AI Validation Focus

Regulation / Guidance	Key Requirements (AI Perspective)	Notes / References
FDA 21 CFR Part 11	Determine record-specific applicability from predicate rules, electronic reliance and eligible electronic submissions. AI inputs/outputs are not automatically Part 11 records; signing requirements must be assessed separately.	FDA scope guidance .
FDA Computer Software Assurance (CSA)	Risk-based assurance for software used in medical-device production or quality management systems.	Final guidance, February 2026; not a universal pharmaceutical CSV replacement (FDA guidance).
EU GMP Annex 11	Validate applications; qualify IT infrastructure. Justify lifecycle evidence and the extent of controls by documented risk assessment.	No universal IQ/OQ/PQ sequence prescribed (Annex 11).
EU consultation draft Annex 22 (AI/ML)	Proposed scope: static, deterministic models in critical GMP applications; excludes dynamic and probabilistic models, generative AI and LLMs, which the draft says should not be used in critical GMP applications.	Draft text ; the 2025 consultation has ended. Distinguish these proposed provisions from applicable GMP requirements and the wider ISPE framework.
ICH Q9	Quality risk management for all processes, emphasizing patient safety and product quality.	Basis for risk-based approach in AI validation (ispe.org).
ALCOA+ (MHRA/FDA Data Integrity)	Records must be Attributable, Legible, Contemporaneous, Original, Accurate (plus Complete, Consistent, Enduring, Available).	Apply governance according to GxP purpose and risk; separately assess legal record and signing obligations (FDA scope guidance ; EMA reflection paper).
EU AI Act (enacted)	Risk-based law in force since August 1, 2024, with phased application. AI literacy/initial prohibitions: February 2025; GPAl/governance: August 2025; general applicability/transparency: August 2026.	Current high-risk timeline: Annex III, December 2, 2027; Annex I product systems, August 2, 2028. Applicability depends on the use case and role (Commission timeline).
ISO 13485:2016 / medical-device QMS	Assess the manufacturer's applicable quality-system obligations and exemptions; plan validation of AI software for its intended use separately.	Assess applicability under FDA's QMSR .
Software consensus standards	Consensus standards are voluntary for FDA medical-device premarket submissions unless incorporated by reference into a regulation.	Assess the selected standard's product-specific scope before claiming conformity (FDA standards guidance).

These overall requirements create a **demand for a new validation framework**. In the next sections, we explore how the ISPE GAMP AI Guide and related thinking address these needs.

Challenges in Validating AI-Enabled Computerized Systems

Before detailing guidance approaches, we first review *why* validating AI systems poses unique difficulties beyond traditional systems. Several interrelated challenges have been identified by industry and regulators:

- **Black-Box and Complexity:** Many AI models (deep neural networks, ensemble models) do not expose transparent logic like rule-based software. This opaqueness makes specific code-path testing impossible; instead, validation must focus on statistical performance and behavior under varied scenarios. It also complicates root-cause analysis of errors (a key part of CSV) if something goes wrong.
- **Non-Deterministic Outputs:** Unlike deterministic code, AI outputs may vary due to random seeds, training samples, or non-linear interactions. For example, an LLM might generate different valid sentences on each run for the same question. This means validation cannot expect a single “correct” answer; instead, users must define acceptable performance ranges and use large test sets. Regulatory commentary notes that reproducibility expectations differ: outputs from static, validated datasets should be consistent, whereas AI systems using open data can yield more unpredictable results (www.ey.com). GxP validation must adapt to this variability.
- **Bias and Data Quality:** If training data are incomplete, biased, or flawed, the AI model will inherit those issues. In GxP, this can threaten patient safety (e.g. if a diagnostic AI was only trained on one demographic). Traditional validation seldom addressed dataset bias; for AI, handling such risks is crucial. The GAMP AI Guide explicitly calls out data bias as a risk to manage.
- **Data Volume and Preprocessing:** AI often requires large amounts of data, which may come from multiple sources (lab instruments, electronic records, external databases). Ensuring the integrity of large-scale data ingestion pipelines is nontrivial. It necessitates robust data governance — versioning datasets, verifying data provenance, and recording all preprocessing steps. This is far beyond typical CSV concerns, which usually treat data generation as given.
- **Drift, Retraining, and Online Learning:** Frozen models do not adapt in production. Controlled retraining produces a changed model version, whereas online learning uses new production data to update the model through a defined process ([pharmacovigilance validation paper](#)). Changing deployment data can require performance reassessment without the frozen model learning from those inputs. Monitor for drift/degradation and define thresholds and investigation triggers ([EMA reflection paper](#)). Tailor change control to the actual deployment mode. The [draft Annex 22](#) excludes dynamic models and says they should not be used in critical GMP applications; discussing online learning in the wider validation framework does not authorize that use.
- **Toolchain and Infrastructure:** Implementing AI often relies on complex toolchains (data lakes, notebooks, cloud services). Standard CSV focuses on packaged software or configuration changes; AI toolchains can be more fluid and distributed (e.g., pipelines in the cloud, many open-source components). This raises questions about infrastructure qualification and cybersecurity. The GAMP Guide accordingly devotes appendices to IT infrastructure, noting new threats (e.g. adversarial attacks on ML) and the need for secure automated pipelines.

- **Skill and Knowledge Gaps:** Many quality professionals have little familiarity with AI/ML. As one guide co-author remarked, terms and processes were poorly understood by many in the industry ([ispe.org](https://www.ispe.org)). This lack of competency itself is a risk: validation requires critical thinking about AI methods. Hence, the GAMP AI Guide emphasizes upskilling and knowledge management as part of the validation strategy.

These challenges underscore that *the essence of validation doesn't change*, but the **methods do**. We still must ensure “systems are of high-quality, effective, fit for their intended use, and compliant” ([ispe.org](https://www.ispe.org)) – but this now entails verifying model performance, data pipeline integrity, and new lifecycle activities. The next section lays out how to approach these needs systematically.

Key Concepts in AI System Validation for GxP

Building on existing GAMP 5 risk-based CSV, validation of AI in GxP includes several additional layers. The ISPE GAMP AI Guide and related expert analyses highlight the following key considerations:

1. **Context of Use and Risk Analysis:** The first step is to clearly define *what the AI system is intended to do* in the process and what could go wrong. Every AI function must be tied to a risk assessment. For example, if an AI algorithm suggests batch release, a faulty suggestion could directly affect patient safety. Therefore, such a use-case is high-risk and demands robust validation. Conversely, if AI only helps prioritize routine tasks, the risk is lower. Explicitly mapping each AI function to the levels of risk ensures validation effort is *tiered and proportionate* (www.bioprocessonline.com) (intuitionlabs.ai). In practice, one should document the “use case” and patient QTPP (quality target product profile) impact, and rank validations accordingly.
2. **Data as the Foundation:** Establish governance for the data, models and algorithms relevant to the AI's GxP use; EMA recommends extending governance SOPs in cases of high regulatory impact or high patient risk ([EMA reflection paper](#)). Document data lineage and relevant transformations. **Assess Part 11 applicability record by record** using predicate rules, electronic reliance, and submission requirements ([FDA scope guidance](#)). Retain training, test, prompt and output evidence according to its purpose, applicable retention requirements, and documented risk. An internal decision to retain AI evidence does not itself make every interaction a Part 11 record. Determine separately which approvals or other signings are required by the applicable predicate rules; do not assign a signature obligation to every prompt or output.
3. **Performance and Validation Criteria:** AI models can have deterministic or probabilistic outputs; thus the validation goal is to demonstrate *adequate performance*, not perfection. Stakeholders must **a priori** define what success looks like: target accuracy, sensitivity, or error tolerances relevant to the use case (www.bioprocessonline.com). For instance, an AI that identifies anomalous QC lab results might be expected to meet a predefined sensitivity target for true anomalies while a higher false-positive rate might be allowable. These acceptance criteria should be justified by risk and clinical context. During validation testing, one should use test datasets representative of real-world conditions (including edge cases) and measure the AI's results against the criterion. Importantly, review boards should *agree*

upfront on acceptable error rates, rather than rationalize them *ex post* (www.bioprocessonline.com). The GAMP AI Guide recommends building in performance metrics and validation endpoints into the design and documentation (e.g. in the User Requirements Specification).

4. **Testing Strategies:** AI system testing extends beyond code execution. In addition to functional tests, one must perform **model-specific testing**. This can include:
 - *Robustness testing*: e.g. verify the model's performance when input data contain noise or novel conditions.
 - *Bias testing*: check that the model performs equally well across expected subpopulations (dosage forms, patient subgroups).
 - *Stress/Adversarial tests*: evaluate how minor perturbations in input might affect outputs (especially for image or clinical algorithms).
 - *End-to-End tests*: use the model in the intended process workflow (including data collection and decision execution) to confirm overall fit-for-purpose. Documenting these tests parallels CSV but with an emphasis on empirical performance. The Guide suggests test plans include human reviewer judgments where applicable, and the use of large, labeled datasets for quantitative validation (ispe.org). Finally, any test must itself be controlled; test datasets should be version-controlled and segregated from training data (see item 5 below).
5. **Training Discipline and Model Versioning:** Model development is a critical validation activity. One expert recommends treating the **train/test split** as part of validation: once the partitions are fixed, they must not be re-used or updated to avoid data leakage (www.bioprocessonline.com). All training runs should be reproducible: record the code version, training parameters, and configuration so results can be traced. When a model is retrained (due to new data or improvements), the new model version should be documented, and the rationale and data differences should be validated. The GAMP Guide emphasizes strict version control of both data and model artifacts. Essentially, model development logs (notebooks, scripts, parameters) become part of the validation record.
6. **Classic Quality Controls (Systems & Process Controls):** Despite AI's novelty, traditional GxP controls still apply. This includes:
 - **Access Controls:** Only authorized personnel (or validated automated processes) can interact with the AI system. In practice, this may mean restricting who can initiate training, deploy a model, or make predictions.
 - **Audit Trails:** Define the records, changes and events to log from applicable predicate-rule requirements and a documented risk assessment of product quality, safety and record integrity ([FDA scope guidance](#)). Model-version and configuration logs can support internal assurance; do not treat logging every interaction as an automatic Part 11 duty.
 - **Backup/Restore:** As with any computerized system, validated processes should exist to back up AI assets (datasets, models, code) and restore them in the event of failures. This is crucial since an AI model might be considered a digital "asset" requiring protection.

- **Segregation of Duties:** Ensure that the people approving an AI model are independent from those developing or operating it, when feasible, to avoid conflicts of interest in validation. These controls ensure that, even though a machine learning model is doing data-driven work, it remains within the standard GxP quality framework (www.bioprocessonline.com).
7. **Change Management and Monitoring:** For AI systems, **change is expected**, not exceptional. Therefore, a robust change control process is vital. The GAMP AI guidance calls for predefined **revalidation triggers**. Triggers might include: significant updates to training data, retraining of the model, detection of performance drift beyond acceptable limits, major software updates in the AI (e.g., new algorithm), or even hardware/platform changes affecting outputs. When such an event occurs, the change management process should require impact assessment and potentially a partial or full revalidation. Additionally, an ongoing monitoring plan should be in place with metrics and alerts. Ideally, the organization has dashboards tracking model performance (e.g. error rates, Q2I trends) against thresholds, enabling early detection of degradation (www.bioprocessonline.com). This is analogous to normal post-market surveillance for devices, only automated and continuous.
8. **Vendor and Supply Chain Considerations:** Many AI solutions are sourced from vendors. Regulatory bodies expect firms to **rigorously qualify AI vendors**. This can include the vendor's quality management processes, their own validation documentation, and contractual quality agreements. The GAMP AI Guide stresses that companies must hold external AI providers to GxP standards; after all, the ultimate compliance responsibility lies with the user company (intuitionlabs.ai). Supplier assessments should include reviews of the vendor's data sources (are they GMP-compliant?), model development practices, and update policies. In summary, third-party AI tools are treated similarly to any critical equipment or software supplier in the QMS.

In essence, *validating an AI system in GxP means addressing all the above points in a structured, risk-based way*. To help operationalize this, experts have proposed step-by-step frameworks. For example, one recent framework (focused on 21 CFR Part 11 compliance) outlines stages like: (1) **Context-of-Use Definition** and risk mapping; (2) **Data Governance**, documenting record applicability and risk-based data-integrity controls; (3) **Performance Requirements** (defining accuracy/error thresholds); (4) **Testing and Validation** (risk-based test cases appropriate to the applicable regulated activity); (5) **Audit Trails and Documentation** (logging relevant records and events according to applicable duties and documented risk); (6) **Vendor Qualification** (ensuring third-party tools meet quality standards); (7) **Change Management** (formal process for model updates); and (8) **Periodic Review** (analogous to pharmacovigilance for continuous monitoring) (intuitionlabs.ai). (Table 2 below summarizes these core steps in a validation framework.)

Table 2: Core Steps in a Risk-Based Validation Framework for AI in GxP

Step	Description
1. Context of Use (COU) & Risk	Define intended use and patient/product/data risk. Document Part 11 applicability by predicate rule, electronic reliance and eligible submission requirements; assess signing duties separately (FDA scope guidance).
2. Data Governance (ALCOA+ controls)	Define data-integrity controls and retained AI evidence according to GxP purpose, applicable record duties and documented risk. Prompts and outputs are not automatically Part 11 records (FDA scope guidance).

Step	Description
3. Performance Requirements	Set clear acceptance criteria for model accuracy, error rates, bias tolerances a priori (www.bioprocessonline.com). Align these with clinical/regulatory expectations.
4. Validation Testing	Execute risk-based tests appropriate to intended use, including relevant edge cases; document results. FDA's final CSA guidance concerns medical-device production/QMS software (FDA CSA guidance); pharmaceutical validation duties require their own assessment.
5. Audit Trails & Documentation	Define logging and retention from predicate rules and documented risk. Include relevant AI lifecycle evidence and SOPs without imposing a legal duty to log or sign every interaction (FDA scope guidance).
6. Vendor Qualification	Audit or assess third-party AI tools/suppliers; verify their quality systems, data provenance, and update controls (intuitionlabs.ai).
7. Change Management	Establish governance for model retraining/updates. Define triggers for revalidation (e.g. drift detection) and version control strategies (intuitionlabs.ai).
8. Ongoing Monitoring/Review	Continuously track model performance metrics and data integrity. Implement dashboards and alerts for deviations; plan periodic reassessment (analogous to post-market) (www.bioprocessonline.com).

(Sources: adapted from ISPE GAMP AI Guide concepts and published frameworks (intuitionlabs.ai) (www.bioprocessonline.com))

This framework aligns traditional CSV elements with AI-specific needs. Steps 2–4 and 8 are especially novel for AI, focusing on data stewardship and continuous oversight which were not explicitly part of older guidance. By following such a risk-based blueprint, organizations can systematically demonstrate that “the system consistently meets requirements and is fit for its intended use” – in this case, using AI to support GxP processes (www.bioprocessonline.com).

Case Studies and Examples

While AI in fully regulated production is still emerging, the following published example, potential applications and explicitly hypothetical scenario illustrate issues to consider. Reported vendor results and illustrative controls should not be read as regulatory confirmation of a deployment.

- **Generative AI in Validation (PwC example):** A PwC quality transformation team reports using a GPT-based solution to automate parts of the validation documentation process. For a data analytics dashboard in manufacturing, the traditional approach was to manually script tests for each view. The AI-powered solution “automatically generated baseline scripts” for tests, filling in navigation paths and data logic, under human review. PwC reported a 40% reduction in drafting time for baseline dashboard scripts and improved standardization (www.pwc.com). This example highlights how generative AI can accelerate GxP tasks (drafting test cases, documentation), but also underscores the need for human oversight to ensure compliance. PwC describes baseline-script generation with human reviewers focusing on context-specific inputs (www.pwc.com).
- **AI in Quality Management Systems (potential application):** Stephanie Gaulding of Pharmatech Associates discusses AI-assisted quality management as a potential application, including presenting analytical findings to QA staff for further investigation ([Gaulding's commentary](#)). This is commentary on

possible capabilities, not evidence that identified pilot companies achieved fewer batch deviations. An organization considering this use would need to evaluate performance in its own workflow before relying on it.

- **Vision Systems for Manufacturing QC (hypothetical scenario):** Consider a static, deterministic AI camera proposed for blister inspection. This is an illustrative validation scenario, not a report of a vendor, site, achieved accuracy, or replacement of human inspection. Before relying on the system, the team would define the intended use, representative test conditions, suitable performance metrics, and risk-dependent acceptance criteria. The [draft Annex 22](#) proposes defining acceptance criteria for model test metrics according to intended use. Numerical thresholds would require justification for the particular product and inspection task; no achieved sensitivity, specificity or compliance result is claimed here.
- **AI in Pharmacovigilance:** A 2021 expert paper discussed AI-assisted adverse event case processing ([pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/35111111/)). For example, some firms use AI/ML to field-parse case narratives or code MedDRA terms. One category (“AI-based static systems”) entails fixed models that do not learn in production. For these, the authors recommend validating under GAMP risk-based concepts: document the algorithm, validate against known case sets, and maintain version logs. For more dynamic systems (continuously learning), they propose even more sophisticated monitoring. This pharmacovigilance context parallels GxP: patient safety is paramount, so the same validation rigor (data integrity, testing with known cases, audit trails) applies.

The PwC example reports a 40% reduction in baseline dashboard-script drafting time ([PwC account](#)). It does not establish a 40% reduction in test execution or the full validation process. The other potential and hypothetical applications illustrate questions for evaluation, not achieved site results or compliance guarantees. Relevant considerations include data infrastructure, justified performance criteria, human review and ongoing oversight.

Implications and Future Directions

The ISPE GAMP AI Guide and corresponding practices signal important shifts in how the regulated industry will handle AI going forward:

- **Risk-Based Validation and Assurance:** Organizations should choose proportionate validation activities while preserving the duties applicable to their regulated processes. FDA’s final February 2026 CSA guidance provides recommendations for medical-device production and quality management system software ([FDA CSA guidance](#)); it is not a mandate for all pharmaceutical GxP organizations to replace CSV. AI-assisted test generation can be evaluated as a tool, but tool use does not itself demonstrate fitness for intended use or regulatory compliance.
- **Evolving Roles and Skills:** Quality and validation roles will expand to include data scientists and AI engineers. ISPE explicitly calls for building AI literacy and incorporating specialized expertise into QA units ([www.scribd.com](https://www.scribd.com/document/58111111/ISPE-GAMP-AI-Guide)) (www.bioprocessonline.com). Regulatory personnel, historically focused on 21 CFR/Annex 11 compliance, will need familiarity with AI terms. Training initiatives (internal or via organizations like PDA, ISPE) are likely to multiply. Over time, one might see formal certification or competency frameworks for “AI in GxP” roles.

- **Continuous Monitoring and MLOps Integration:** The operational mindset may increasingly treat validated AI more like a production process than a static release. MLOps (DevOps for ML) practices — automatic data pipelines, continuous integration, version control — will be integrated into pharma IT landscapes. Mature AI systems may run dashboard monitors that automatically alert QA if performance drifts beyond thresholds. This is a departure from the “perform once per release” mentality. EMA recommends monitoring model performance for early detection of drift/degradation and defining acceptable thresholds ([EMA reflection paper](#)).
- **Guideline and Standard Development:** We can anticipate formal updates to regulations. For example, the consultation drafts of revised Annex 11 and new Annex 22 in Europe, FDA’s work on AI credibility assessment, and global ISO standards are developments to track. Proposed provisions should be distinguished from applicable requirements. The guide itself may eventually influence regulatory inspections: inspectors might refer to it as a benchmark for compliant AI use. In the meantime, companies developing AI should track guidance like ISO/IEC 42001 (AI management systems) and adapt to data privacy laws (GDPR) as they affect AI data.
- **Ethics and Trust as Oversight:** AI introduces ethical considerations (bias, explainability), which are receiving more attention. While not historically part of GxP, concepts like “Trustworthy AI” are becoming relevant. Future guidance may explicitly require testing for bias or fairness in clinically sensitive AI applications. The industry is moving toward embedding ethics committees and review boards in tech governance.
- **Cross-Industry Collaboration:** Because AI is domain-agnostic, knowledge from tech, automotive, finance (where AI governance is also emerging) will increasingly inform pharma practice. For instance, tech-sector standards on ML explainability or safety-critical AI may be adapted for life sciences. ISPE’s GAMP AI Guide itself is an example of synthesizing cross-disciplinary input (software engineering, data science, quality). We may see more joint workshops among regulators and industry to calibrate expectations.
- **Technological Innovation in QA:** The QA function itself will increasingly leverage AI tools. Spell-checking 21 CFR documents is trivial with NLP; more importantly, model-based anomaly detection could proactively find issues in plant data or batch records. The concept of “AI validating AI” mentioned above (using ML to spot deviations) will likely grow. In short, regulators will want to see that AI systems used for GxP have at least as much oversight as those used for AI.

Conclusion

Artificial intelligence offers the pharmaceutical industry unprecedented capabilities for analysis, control, and innovation — but it also introduces compliance complexities unlike any seen before. The **ISPE GAMP® Guide: Artificial Intelligence** provides a timely, comprehensive framework to bridge this gap ([ispe.org](#)) ([ispe.org](#)). It reaffirms the immutable goals of GxP (patient safety, data integrity, product quality) while detailing **how** to apply them to AI’s novel challenges. Equally important, it represents a consensus path forward: regulators now have a clear industry voice on what is considered “best practice” for AI validation.

For organizations, this means evolving their validation strategies. Validation and assurance should be proportionate to intended use and risk, while preserving applicable regulatory duties. Data scientists must work with quality teams, and validation narratives must include AI-specific content (data stewardship, model evaluation, drift detection). The upfront investment in this expanded framework is significant, but so is the potential reward: properly validated AI can accelerate processes (in the PwC example, PwC reported a 40% reduction in baseline dashboard-script drafting time (www.pwc.com)), support further evaluation of productivity and quality benefits, without guaranteeing deployment compliance.

Moving forward, industry and regulators will learn in tandem. The GAMP AI Guide, together with emerging regulations (e.g. FDA's AI framework, EU AI Act, ISO standards), will shape the field. Organizations should evaluate readiness for deployment and scaling in their own context, with defined responsibilities and evidence for continued performance. Through diligent adherence to risk-based controls, transparent documentation, and continuous monitoring (as laid out by the guide), AI systems can become reliable, compliant tools that uphold the excellence of GxP programs.

References: (Inline citations preceding. Key sources: ISPE GAMP AI Guide overview (ispe.org) (ispe.org); ISPE Pharmaceutical Engineering article (ispe.org) (ispe.org); FDA and EMA regulatory references (intuitionlabs.ai) (www.ey.com) (www.bioprocessonline.com) ([EMA final reflection paper](#)); industry analyses (www.pwc.com) (www.bioprocessonline.com); peer-reviewed pharmacovigilance validation discussion (pmc.ncbi.nlm.nih.gov).)

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. <https://ispe.org/publications/guidance-documents/gamp-guide-artificial-intelligence>
2. <https://ispe.org/pharmaceutical-engineering/september-october-2025/new-gampr-guide-addresses-challenges-posed-ai>
3. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7892696/>
4. https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle_en.pdf
5. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/part-11-electronic-records-electronic-signatures-scope-and-application>
6. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/computer-software-assurance-production-and-quality-management-system-software>
7. <https://intuitionlabs.ai/articles/deviations-cap-a-change-control>
8. <https://intuitionlabs.ai/articles/fda-ai-drug-development-guidance>
9. <https://www.pwc.com/us/en/industries/health-industries/library/computer-system-validation.html>
10. <https://www.bain.com/about/media-center/press-releases/2024/40-of-pharma-executives-are-baking-expected-savings-from-generative-ai-into-2024-budgets/>
11. <https://www.mckinsey.com/industries/life-sciences/our-insights/scaling-gen-ai-in-the-life-sciences-industry>
12. https://www.ey.com/en_ch/insights/life-sciences/gxp-and-ai-tools-compliance-validation-and-trust-in-pharma
13. <https://www.bioprocessonline.com/doc/trust-but-verify-validating-ai-in-pharma-s-gxp-world-0001>
14. https://health.ec.europa.eu/system/files/2016-11/annex11_01-2011_en_0.pdf

15. <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/artificial-intelligence>
16. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
17. <https://ispe.org/news/ispe-announces-availability-ispe-gampr-guide-artificial-intelligence>
18. <https://www.scribd.com/document/945601934/Gamp-Guide-Ai-Toc>
19. https://health.ec.europa.eu/consultations/stakeholders-consultation-eudralex-volume-4-good-manufacturing-practice-guidelines-chapter-4-annex_en
20. https://health.ec.europa.eu/document/download/5f38a92d-bb8e-4264-8898-ea076e926db6_en?filename=mp_vol4_chap4_annex22_consultation_guideline_en.pdf
21. <https://www.iso.org/ISO-IEC-17025-testing-and-calibration-laboratories.html>
22. <https://www.fda.gov/medical-devices/postmarket-requirements-devices/quality-management-system-regulation-qmsr>
23. <https://www.fda.gov/news-events/press-announcements/fda-releases-artificial-intelligencemachine-learning-action-plan>
24. <https://www.fda.gov/media/166704/download>
25. <https://www.ema.europa.eu/en/news/reflection-paper-use-artificial-intelligence-lifecycle-medicines>
26. <https://www.fda.gov/media/71983/download>
27. <https://www.pharmaceuticalonline.com/doc/how-ai-is-revolutionizing-quality-management-systems-0001>

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore [AI enablement](#), [pharma and biotech software development](#), [data engineering and BI](#), and [Veeva services](#). [Contact IntuitionLabs](#) to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: <https://intuitionlabs.ai>

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.