IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

Agentic AI for Pharma Regulatory Document Automation

> 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 · 2026-03-05 · agentic ai · clinical study reports · new drug applications · ectd workflows · pharmaceutical compliance · ai governance

Original article: https://intuitionlabs.ai/articles/agentic-ai-pharma-regulatory-document-automation



In this article

1. Executive Summary
2. Introduction and Background
3. Agentic AI: Concepts and Capabilities
4. Regulatory Documents: NDA, CSR, and eCTD Overview
5. New Drug Applications (NDAs)
6. Clinical Study Reports (CSRs)
7. eCTD Workflows
8. AI-Powered Automation: Case Studies and Evidence
9. Data Analysis and Quantitative Insights
10. Challenges, Risks, and Regulatory Considerations
11. Future Directions and Implications
12. Conclusion
13. References

Executive Summary

Agentic AI – broadly defined as autonomous, goal-directed AI systems – is poised to transform regulatory document processing in the pharmaceutical industry. By orchestrating multi-step workflows with minimal human intervention, agentic AI can dramatically accelerate the preparation of key regulatory submissions such as **Clinical Study Reports (CSRs)**, **New Drug Applications (NDAs)**, and the electronic Common Technical Document (**eCTD**) dossiers, while also improving consistency and traceability. For example, a recent case study using an AI-powered writing assistant ("AutoIND") showed an approximate **97% reduction in drafting time** for IND submission summaries, with expert regulatory writing still required before submission (arxiv.org). Similarly, advanced retrieval-and-reasoning agents (e.g. "RegGuard") are being developed to parse complex regulations and align them with company policies to reduce compliance burdens (arxiv.org).

Nonetheless, agentic AI introduces new challenges in assurance and governance. Hallucinations and omissions remain a risk – as one study of AI-generated trial summaries noted, current systems tend to **"introduce unsupported statements"** that make them unsuitable for direct use without expert review (arxiv.org). Compliance concerns (such as 21 CFR 11 requirements for validated electronic records (en.wikipedia.org)), data privacy, and the need for rigorous validation of AI outputs are significant hurdles. Addressing these requires integrating robust retrieval (e.g. RAG) and policy frameworks (arxiv.org), as well as establishing audit trails and oversight mechanisms. Current "hybrid" approaches – where AI generates first drafts or parses large documents and human experts revise and validate – show great promise. In sum, agentic AI offers **considerable efficiency gains and new capabilities** for regulatory writing, but industry adoption hinges on ensuring quality, accountability, and compliance at every step.

Introduction and Background

The preparation of regulatory submissions in pharmaceuticals is traditionally labor-intensive. A **New Drug Application (NDA)** – the formal request to the U.S. Food and Drug Administration (FDA) to approve a new drug for marketing – must *"tell the drug's whole story"* (www.fda.gov). NDAs compile data from all human and animal studies, detail the drug's chemistry and manufacturing, and present clinical safety and efficacy data, among other elements (www.fda.gov). Likewise, **Clinical Study Reports (CSRs)** are extremely detailed documents that summarize the design, conduct, and results of clinical trials. A CSR is generally "a very long" scientific report providing *"much detail about the methods and results of a trial"*, serving as an authoritative account of efficacy and safety (en.wikipedia.org). All of this content must then be assembled into the **electronic Common Technical Document (eCTD)** format for submission. The eCTD backbone organizes submission content into five modules (FDA). For submissions to FDA's CDER and CBER, eCTD requirements cover NDAs, ANDAs, BLAs and commercial INDs, subject to applicable exemptions and waivers. For noncommercial INDs, including investigator-sponsored and expanded-access INDs, eCTD is optional but encouraged (FDA).

Historically, the volume of submission documents has only grown. Since 1938, an NDA has been required for any new drug in the U.S. (www.fda.gov). ICH's **Common Technical Document (CTD)** guidance recommends a five-module organization for submission dossiers (FDA). FDA has accepted submissions in eCTD format since 2003 (FDA). eCTD offers critical advantages (structured navigation, incremental updates, validation checks) that can significantly shorten review cycles, but it also imposes strict technical requirements (file naming, format, sequencing) on sponsors. In the United States, FDA's eCTD requirements apply to specified CDER/CBER submissions, including commercial INDs; noncommercial INDs are optional but encouraged (FDA). Requirements must be checked for each jurisdiction.

In parallel with these regulatory evolutions, **artificial intelligence (AI)** has advanced rapidly. Traditional AI models (rule-based systems, classifiers, or even basic generative models) generally perform defined tasks under human direction. The latest generation of **Large Language Models (LLMs)** – the so-called *generative AI* – can generate human-like text from prompts, assisting in tasks from drafting emails to summarizing articles. However, generative AI is **fundamentally reactive**: a language model like ChatGPT produces outputs in response to each user prompt and lacks autonomous goal-setting (www.techtarget.com). In contrast, **agentic AI** refers to systems that can interpret high-level objectives, plan multi-step workflows, retrieve information, and take actions with minimal step-by-step instruction (www.techtarget.com). An agentic AI might, for example, be given the goal "Prepare the clinical pharmacology summary in CTD section 2.7.2 for expert review", then autonomously gather relevant data, draft the text, check it against regulations, and refine the output, rather than waiting for a human to specify each subtask.

Agentic AI is still emerging, but it promises new capabilities for automation. Such systems typically combine multiple components – e.g. a planner, one or more LLMs, retrieval modules, tools or APIs – working in concert (www.techtarget.com). These agents can perceive their environment (via APIs, documents, or user inputs), retrieve domain knowledge (often through **Retrieval-Augmented Generation**, or RAG), reason about the task, collaborate among sub-agents, and even learn over time (www.techtarget.com). For highly structured processes like preparing regulatory documents, an agentic approach allows automation of not just drafting

text, but also checks, cross-references, data extraction, and file assembly, in a traceable way. In the sections below, we examine how agentic AI can be applied to CSRs, NDAs, and eCTD workflows, evaluating current capabilities, use-cases, and data from recent studies.

Agentic AI: Concepts and Capabilities

Agentic AI systems are built around the idea of *autonomy and goal orientation*. As TechTarget explains, agentic AI is the class of AI that can execute **"autonomous action and decision-making"** (www.techtarget.com). Unlike a simple LLM that waits for a user prompt, an agent can **"set its own goals"** (within human-defined objectives) and choose strategies to achieve them (www.techtarget.com). In practice, this means the agent can break a complex task into subtasks, invoke appropriate skills (sometimes other AI agents or tools), and adjust its approach based on new information. For example, an agent tasked with assembling an eCTD docket might separately handle data retrieval, formatting, and consistency checking, coordinating these steps without a human micromanaging each one.

Key **capabilities** distinguish agentic AI from traditional AI or LLM chatbots. These include (see Table 1 below):

- **Autonomy and Planning**: Agentic AI can pursue objectives proactively. It is not purely reactive to prompts; it can plan multi-step workflows and execute them. Traditional AI, by contrast, typically requires a user or developer to define each step. Agentic systems may use techniques like reinforcement learning or reasoning chains to simulate planning (www.techtarget.com).
- **Knowledge Retrieval and Context**: Agents actively fetch and integrate external information. They often use RAG or knowledge bases to ground their actions in up-to-date data (www.techtarget.com). A generative model alone has static knowledge in its parameters, but an agent can query regulations, prior documents, or databases on demand (essential for accurate regulatory content).
- **Multi-Agent Collaboration**: Complex agents are often implemented as *teams* of specialized sub-agents. For example, one sub-agent might extract data from a clinical study dataset, while another formulates narrative text and another validates facts. These sub-agents can orchestrate and delegate tasks dynamically (www.techtarget.com), whereas traditional AI rarely negotiates or hands off subtasks.
- **Continuous Learning and Memory**: Agentic systems can incorporate memory and adapt over time. Though still nascent, some agents bookmark context or user feedback to improve. Traditional models, once trained, do not usually learn from each session in real-time.
- **Governance and Compliance**: Crucially for regulated domains, agentic AI architectures are being designed with oversight in mind (arxiv.org). For instance, emerging frameworks propose embedding policy-checking "open policy agents" into the system to enforce rules (arxiv.org), and user-controlled token schemes to regulate agent interactions (arxiv.org). Such features are largely absent in simpler AI tools.

Table 1. *Comparison of Traditional (Generative) AI vs. Agentic AI*

Capability	Traditional AI / Generative Models	Agentic AI
Autonomy	Reactive: generates content in response to user input; no self-initiated planning (www.techtarget.com)	Proactive: sets and pursues goals, plans multi-step tasks (www.techtarget.com)
Task Scope	A standalone model responds to prompts; fixed LLM workflows can chain multiple steps through predefined code paths (Anthropic)	Handles complex workflows, multi-hop tasks; can divide tasks into subtasks (www.techtarget.com)
Interaction	"One-shot" or conversational; relies on human for direction	Capable of multi-agent collaboration and delegation; operates continuously
Knowledge Integration	Uses internal model knowledge; may retrieve some data if explicitly prompted	Actively uses external knowledge (RAG, databases, APIs) to inform decisions (www.techtarget.com)
Reasoning & Adaptation	Limited reasoning (as per training); no on-the-fly learning	Employs reasoning (e.g. chain-of-thought, ReAct) and can adapt plans based on feedback
Governance & Control	Fixed LLM workflows can include guardrails, including a separate model that screens inputs (Anthropic)	Agent deployments also need appropriate guardrails and extensive testing; controls must be designed for the implementation (Anthropic)

Agents thus extend the capabilities of LLMs toward more autonomous, end-to-end process automation. Notably, although agentic AI allows greater independence, it is still governed by human-set objectives. In practice, an agent for regulatory documents could receive a defined goal such as "Draft a report for a placebo-controlled efficacy study for expert review and placement in CTD section 5.3.5.1." The agent chooses how to fulfill this: gathering source data, drafting, correcting format, etc., all without intermediate prompts. Research on agent frameworks (e.g. ArGen) is actively exploring how to embed **policy and regulatory rules** directly into such systems via "policy-as-code" and reward mechanisms (arxiv.org), providing a path toward AI that is not only capable, but demonstrably compliant.

Regulatory Documents: NDA, CSR, and eCTD Overview

New Drug Applications (NDAs)

An **NDA** is the primary regulatory submission for drug approval in the U.S. It aggregates everything known about a drug: results of all clinical trials, pharmacokinetics and pharmacodynamics, toxicology, manufacturing processes, and proposed labeling. As the FDA notes, the NDA must demonstrate that *"the drug is safe and effective… and the manufacturing methods… are adequate"* (www.fda.gov). In effect, it demands the entire *"drug's whole story"*, including clinical data, formulation details, and quality controls (www.fda.gov). Because NDAs are so expansive, they typically run to thousands of pages. For perspective, one FDA guidance notes that an NDA bundling clinical and statistical sections alone spans many hundreds of pages.

NDAs are organized using the **Common Technical Document (CTD)** framework – five modules comprising, among others, Module 3 (Quality), Module 4 (Nonclinical), and Module 5 (Clinical). Module 2 contains high-level summaries (Quality Overall Summary, Clinical Overview, etc.), while Module 1 holds administrative forms.

FDA's eCTD requirement for NDAs, BLAs and ANDAs took effect in May 2017; commercial IND requirements followed in May 2018, subject to applicable exemptions and waivers (FDA guidance). Noncommercial INDs, including investigator-sponsored and expanded-access INDs, are optional but encouraged for eCTD (FDA). Submission workflows should apply the requirements for the relevant application category.

Clinical Study Reports (CSRs)

Clinical Study Reports are a major component of NDA Module 5 (and IND submissions). A CSR is a **scientific report** that documents the design, conduct, and results of a single clinical trial in full detail. By definition, a CSR is *"a document, typically very long, providing much detail about the methods and results of a trial"* (en.wikipedia.org). It is *"scientific"* in character – essentially a comprehensive record of efficacy and safety data rather than a marketing summary (en.wikipedia.org). In ICH terminology, the content of a CSR is defined by **ICH Guideline E3**, which standardizes its format and required sections (title page, synopsis, tables, narratives on safety, etc.) (www.gmp-navigator.com). In practice, CSRs can be hundreds or thousands of pages each (covering each phase 1–3 study), and sponsors often produce dozens of CSRs for an NDA.

Automation of CSR writing is appealing, as these reports require repetitive summarization of trial data (for endpoints, adverse events, etc.) and strict adherence to style and templates. Any AI-generated CSR must faithfully reflect the **raw trial data**; hallucinating a result or omitting an adverse event would be a critical error. Thus, agentic AI solutions for CSRs would focus on **data ingestion and templating** – pulling in statistical outputs and generating narrative buffers that experts refine. (In the CliniDigest study, summaries used an average of 54% of sources, a source-use metric that does not establish completeness of the underlying facts (study).) Ensuring an audit trail (so every statement in a CSR can be traced back to source data) is also crucial in a regulated environment.

eCTD Workflows

The **eCTD** is the *standardized electronic dossier* used for submissions in ICH regions. It enforces a consistent folder structure: Module 1 (regional forms), Module 2 (summaries), Module 3 (quality), Module 4 (nonclinical), and Module 5 (clinical, i.e. CSRs) (en.wikipedia.org). A correctly constructed eCTD includes an XML "backbone" index that lists all documents and their metadata, and every PDF leaf must follow naming and size limits. Technical validation (XML schema checks, checksum verification) is automated by FDA/EMA tools, and failed checks can cause submission rejection. In short, the eCTD workflow involves not just authoring content, but also generating precise metadata, links, and navigational indexes.

Table 2 below summarizes these document types, their contents, and potential AI automation opportunities and challenges.

Table 2. *Regulatory Document Types, Content, and AI Automation Considerations*

Document Type	Contents and Role	Agentic AI Automation Tasks	Challenges and Quality Controls
NDA (New Drug App.)	Entire drug dossier (Modules 1–5): - Module 5: Clinical (CSRs, trial reports) - Module 4: Nonclinical studies - Module 3: CMC/Quality - Module 2: Summaries - Module 1: Forms/labeling (FDA)	• Generate first drafts of summary sections (Nonclinical Overview, Clinical Overview) using data sources. • Organize documents into eCTD folder + XML index. • Cross-check labeling and CMC info for consistency. • Answer reviewer FAQs via internal knowledge base. (arxiv.org)	• **Accuracy:** Must exactly reflect study data and chemistry. No hallucination of results or processes. • **Format compliance:** Strict eCTD formatting (PDF specs, metadata); errors in index cause submission rejection. • **Traceability:** Audit-worthy output (logs of source data used). • **Content integrity:** Protect sensitive IP, ensure data privacy.
CSR (Clinical Study Report)	Detailed narrative of a single trial: design, methods, outcomes, tables of results (en.wikipedia.org) (www.gmp-navigator.com). Scientific (not promotional) report of efficacy/safety.	• Auto-summarize trial protocol and statistical outputs into narrative paragraphs. • Compile tables and listings automatically from trial datasets. • Consistency check subpopulations, endpoints, dosage info. • Retrieve relevant regulatory guidelines (e.g. ICH E3) to ensure format. (en.wikipedia.org)	• **Factual correctness:** Any error in results/interpretation can jeopardize approval. Expert review mandatory. • **Completeness:** Check completeness against the protocol, statistical analysis plan, study results and applicable guidance. Use an appropriately adapted E3 structure; E3 is a guideline rather than a rigid template (ICH E3 Q&A). (en.wikipedia.org) • **Style & Clarity:** Results should be concise and free of misinterpretation; e.g. a study found no effect must be reported as such, not glossed over (en.wikipedia.org).
eCTD Workflow	The container/format for submissions (en.wikipedia.org): XML index + PDF documents. Ensures consistency across regions.	• Automate folder creation and PDF linking via RAG protocols. • Validate naming conventions, file sizes, and generate audit logs. • Ingest regulatory objectives (e.g. M4/M2 guidelines) to prioritize content. • Use AI to inspect completed eCTD for missing modules or incorrect cross-references.	• **Technical Rigidity:** Even small deviations in file naming or sequence break compliance. Agent must strictly enforce rules. • **Validation checks:** Automated eCTD validation (via FDA/EMA validator) is a gate; AI must pre-validate before submission. • **Synchronization:** If content is updated (e.g. add amendment), agent must correctly update version flags in XML (New vs Replace).

This breakdown highlights that agentic AI can potentially tackle both the *content generation* (writing sections of CSRs or summaries) and the *assembly tasks* (building the eCTD structure, enforcing format). For example, an LLM agent might draft the Nonclinical Overview by reading animal toxicology reports and auto-generating text, while another sub-agent compiles the Module 1 forms and navigational index. In Table 3 below, we further illustrate via recent case studies how AI techniques have been applied to such regulatory text tasks.

AI-Powered Automation: Case Studies and Evidence

Recent research and industrial projects have begun to quantify the benefits of AI assistance in regulatory writing and compliance tasks. Several notable examples include:

- **AutoIND Regulatory Writing (LLM in IND preparation).** A preprint evaluated nonclinical summary drafting. AutoIND generated initial summaries from 18,870 source pages in 3.7 hours; these were not 18,870 generated pages. The paper reports a 96.3% time reduction against an estimated manual benchmark. The authors conclude that regulatory writing experts must refine the drafts and templates, particularly their emphasis and conciseness, before submission (study).
- **Pharmaceutical Regulatory QA Agent ("RegGuard").** University researchers developed *RegGuard*, an AI assistant for navigating complex regulatory texts (arxiv.org). RegGuard ingests heterogeneous sources (FDA guidances, HTA policies, etc.) and uses novel RAG techniques to answer compliance questions. Its key innovations include HiSACC (contextual chunking) and ReLACE (domain-tuned cross-encoder) to improve document retrieval. In industry evaluations, RegGuard significantly increased answer relevance and groundedness while **mitigating hallucinations** (arxiv.org). For instance, when a user posed a query about an FDA regulation, RegGuard returned precise, evidence-backed answers linked to source text, unlike a plain LLM which might simply guess statutory language. The system's architecture also emphasizes auditability – every answer is traceable to specific regulation passages (arxiv.org). RegGuard exemplifies how agentic methods can add compliance-specific layering to AI.
- **GMPilot (AI Compliance Agent for cGMP).** Another recent project, *GMPilot*, targeted FDA good manufacturing practice (cGMP) questions using an AI agent (arxiv.org). Built on a curated regulatory knowledge base, GMPilot employs Retrieval-Augmented Generation (RAG) and a ReAct reasoning framework to support quality professionals. In simulated inspections, GMPilot provided rapid, structured answers by retrieving relevant regulations and past inspection findings (arxiv.org). For example, when asked about deviation investigations, GMPilot could quickly list FDA criteria and past observations, improving "responsiveness and professionalism" of QA review (arxiv.org). GMPilot's design prioritizes traceability (providing citations for each answer) and shows how AI agents can be specialized for narrow regulatory domains.
- **Clinical Trial Summarization (CliniDigest & TrialsSummarizer).** In related domains of evidence synthesis, AI has tackled summarizing large sets of trial descriptions. *CliniDigest* used GPT-3.5 to condense batches of trial records into concise summaries (arxiv.org). In one experiment, 85 individual trial descriptions (around 10,500 words total) were summarized into a single 200-word paragraph with

citations (arxiv.org). On average, CliniDigest summaries used 54% of sources; this measures source use, not coverage of all source facts. Another system, *TrialsSummarizer*, built BART-based multi-document summaries of RCT outcomes given a query (arxiv.org). Both systems found that **fluency and relevance were high**, but crucially they exhibited a "tendency to introduce unsupported statements" (hallucinations) (arxiv.org), underscoring the importance of evidence-checking in regulated outputs.

- **AI Regulatory QA (RIRAG).** The RIRAG framework addressed regulatory text comprehension for question answering (arxiv.org). RIRAG generated thousands of question–answer pairs from financial regulations to train systems. In that setting, the authors stressed that regulatory documents are *"lengthy, complex, and frequently updated,"* requiring significant effort to interpret (arxiv.org). They proposed evaluation metrics focused on capturing *all relevant obligations* without contradiction. Though not specific to pharma, such research highlights techniques (like automated QA pair generation) that could help agents learn regulatory content.

These cases illustrate both the potential and the limits of current AI in regulatory workflows. Quantitatively, AutoIND reports a **96.3% reduction in initial nonclinical IND-1 summary drafting time** against an estimated manual benchmark, with expert revision still required for submission-ready quality (study). Qualitatively, they underscore the need for anchoring AI outputs: *RegGuard* and *GMPilot* explicitly link answers to regulations, and CliniDigest provides citations in its summaries. They also reveal continuing challenges of hallucination and clarity, echoing the expert recommendation that humans *"remain essential to mature outputs to submission-ready quality"* (arxiv.org).

Table 3 summarizes selected case examples, the tasks they tackled, and their outcomes. These examples form the empirical basis for optimistic projections about AI impact.

Table 3. *Case Studies of AI Applications in Regulatory/Clinical Documentation*

Study/Tool	Task/Document Type	AI Technique	Outcome / Key Findings
AutoIND (arxiv.org)	IND Application writing (nonclinical summaries)	LLM-assisted drafting	Drafting time cut from ~100 h to ~3–4 h per batch (~97% reduction). Quality scores ~70–78%. Expert review found omissions making drafts unsuitable for submission (arxiv.org).
RegGuard (arxiv.org)	Regulatory compliance QA	Retrieval/RAG + LLM (domain-tuned)	Improved answer quality (relevance, groundedness) for complex queries. Significantly *mitigated hallucination* risk while providing audit trails (arxiv.org).
GMPilot (arxiv.org)	FDA cGMP compliance support	RAG + ReAct reasoning agent	Provided real-time, traceable answers in simulated inspections. Structurally retrieved regulations and past observations, aiding QA responsiveness (arxiv.org).
CliniDigest (arxiv.org)	Clinical trial summarization	GPT-3.5 batch summarizer	Condensed 85 trial descriptions (≈10,500 words) into a 200-word summary with citations. Across 457 trials in 27 medical subdomains, field-level summaries averaged 153 words and used 54% of sources (arxiv.org).

Study/Tool	Task/Document Type	AI Technique	Outcome / Key Findings
TrialsSummarizer (arxiv.org)	Evidence summarization	Multi-doc summarization (BART + architecture)	Produced fluent summaries of trials retrieved for specified queries. However, tended to introduce unsupported statements, deeming outputs unsafe for unchecked use (arxiv.org).
RIRAG (baseline) (arxiv.org)	Regulatory document QA	RegNLP question-answer generation	Created large QA datasets from regulations to facilitate compliance QA. Highlighted regulatory text complexity and need for full-obligations metrics (arxiv.org).

Each row indicates both the *promise* seen (e.g. rapid summarization) and *caveat* (e.g. residual hallucinations). The trend is that **combining retrieval with generation** yields the best balance: systems like RegGuard and GMPilot explicitly augment LLMs with concept retrieval and rule-based reasoning to boost accuracy. The implication is that a truly reliable agentic solution for NDAs/CSRs will likely be hybrid – using AI to draft and organize, but with embedded factual checks and human oversight.

Data Analysis and Quantitative Insights

Though comprehensive real-world data on AI-based regulatory workflows is still emerging, available studies and industry reports provide quantitative insights:

- **Time Savings.** The AutoIND study (arxiv.org) reports that initial nonclinical IND-1 summary drafting took 3.7 hours versus an estimated 100 hours for manual preparation, a 96.3% time reduction. Expert regulatory writers still need to revise the drafts to submission-ready quality. Similarly, in evidence summarization tasks, AI condenses large document sets by factors of 50–70 (85 trials → 200-word summary (arxiv.org)).
- **Adoption Trends.** Market.us forecasts that the global agentic AI market will expand from $5.2 billion in 2024 to approximately $196.6 billion by 2034, with a 43.8% CAGR between 2025 and 2034 (Market.us). This is a forecast about agentic AI, not artificial general intelligence (AGI).
- **Quality Metrics.** In the AutoIND results, the AI drafts scored ~70–78% on correctness, completeness, conciseness, etc. (arxiv.org). Expert review found omissions making drafts unsuitable for submission. In question-answer settings, RegGuard demonstrated higher relevance/groundedness than unaugmented models (arxiv.org), implying that RAG techniques materially reduce hallucination rates.
- **Regulatory Compliance.** Assess Part 11 applicability for the specific electronic records and relevant predicate rules. FDA recommends documenting those decisions (FDA guidance).
- **Case Study ROI.** AutoIND evaluated first-draft composition of nonclinical IND written summaries (study). A whole-NDA ROI estimate would require separate evidence about drafting, review, revision and submission assembly effort.

Challenges, Risks, and Regulatory Considerations

While the promise is great, deploying agentic AI in regulated document workflows faces critical hurdles:

- **Accuracy and Hallucination Risk.** As noted, AI systems can **hallucinate** plausible-sounding but incorrect information. TrialsSummarizer explicitly showed that state-of-the-art LLM summaries *"introduce unsupported statements"*, making them unsafe without verification (arxiv.org). In regulatory writing, any misrepresentation (even minor) can have severe consequences. Therefore, a core safeguard is **evidence grounding**: linking every AI statement back to source data or a rule. Tools like RAG, knowledge graphs, or explicit citation tracing are essential. For example, RegGuard and GMPilot mitigate hallucinations by coupling LLM outputs with retrieved regulatory text (arxiv.org).
- **Regulatory Compliance of the AI itself.** Determine which electronic records fall within Part 11 using the applicable predicate rules and intended use. FDA describes enforcement discretion for specified validation and audit-trail provisions while retaining predicate-rule obligations (FDA guidance). Document a risk assessment to guide validation and controls; identical output on every run should not be presented as a universal legal requirement.
- **Intellectual Property and Data Use.** Utilizing LLMs raises questions about IP: if the agent accesses published literature or proprietary data, proper licensing must be ensured. Agentic systems may need to handle PHI/PII carefully – advanced AI architectures often propose datastores or on-premise models with access controls (arxiv.org). Additionally, training or fine-tuning models on private regulatory docs would require careful segregation as per industry standards.
- **Skill and Organizational Change.** Regulatory professionals will need new skills (AI oversight, prompt engineering, data strategy). Agents are not plug-and-play; they require well-structured inputs (e.g. organized databases of past CSRs) and expert-curated rules. Change management is non-trivial. Early adopters recommend extensive cross-functional teams: regulatory experts to validate content, IT/AI specialists to manage models, and compliance officers to set governance policies.
- **Ethical and Governance Issues.** Agentic AI introduces questions of "who is responsible" if an autonomous agent makes an error (www.techtarget.com). Industry is starting to develop frameworks (e.g. IBM's AI governance guidelines) to manage such issues. Within pharma, this overlaps with existing computer system validation (CSV) practices: any agentic component might itself be treated as a validated computerized system. Frameworks like ArGen illustrate how to embed multifaceted rules (ethical, safety, regulatory) into AI systems (arxiv.org). Companies may implement "silver bullet" governance layers (blockchain for auditable actions (arxiv.org), user-managed tokens (arxiv.org)) to ensure agents remain under control.

Future Directions and Implications

Looking ahead, the integration of agentic AI into regulatory affairs could unfold along several lines:

- **Increasing Autonomy with Oversight ("Human-in-the-Loop 2.0").** Initially, humans will remain in the loop to review and correct AI outputs. Over time, trust in the agent can grow if its performance is rigorously tracked. Just as clinical trial automation is evolving toward "closed-loop" systems, regulated workflows may see **"closed-loop regulatory writing"** where agentic tools iteratively refine documents based on reviewer feedback and validation checks.
- **Regulatory Acceptance of AI-assisted Documents.** EMA's September 2024 reflection paper addresses AI-assisted drafting and review of medicinal product information in section 2.3.5 and recommends close human supervision (EMA). This is a scoped recommendation for product information, rather than a comprehensive framework for autonomous dossier preparation.
- **Global Harmonization and Standards.** Agentic AI could eventually aid in harmonizing submissions. For example, an agent might automatically adjust content for different regional requirements (e.g. Module 1 variations between FDA and EMA). If agents are trained across multiple guidelines (FDA, EMA, PMDA), they can flag regional differences in real time. In turn, industry groups (like ICH) may develop guidance on the use of AI in writing, akin to how 21 CFR 11 and ICH Q10 provide frameworks for software and quality systems.
- **Expanding Use Cases: Beyond NDAs.** While initial focus is on NDAs and INDs, agentic AI can assist any regulated document workflow: post-marketing safety reports, biannual reports, even training records. For instance, at FDA it's conceivable to have agents monitor regulatory updates and summarize revisions as they impact ongoing studies. On the CRO side, contracts, clinical protocols, and monitoring reports may also be partially automated.
- **Long-term: AI-Driven Submissions?** In the far future, one can imagine fully AI-driven submissions: a model that, given a new drug's data, autonomously generates a complete dossier (drafts, data tables, summaries, and an eCTD archive) and interacts with regulators for clarifications. While still science fiction, ongoing research in multi-agent collaboration and autonomous AI (e.g. chaining LLMs to web search, data tools, each other (www.techtarget.com)) hints at that potential. Ensuring such a system is safe, however, will require robust frameworks like SAGA (arxiv.org) and traceable blockchains (arxiv.org), as well as perhaps legal reforms clarifying who bears liability.

Conclusion

Agentic AI represents a promising evolution of AI for the regulatory domain. By combining generation capabilities of LLMs with planning, retrieval, and tool use, agentic systems can tackle the *scope and complexity* of CSRs, NDAs, and eCTD workflows far beyond what traditional AI has done. The evidence, though early, shows dramatic efficiency gains: e.g. a reported 96.3% reduction in initial nonclinical IND-1 summary drafting time versus an estimated manual benchmark, with expert revision still required for submission-ready quality (arxiv.org). It also demonstrates, however, that rigorous oversight is vital: output quality must match regulatory standards, and every AI action must be auditable. Careful integration (via RAG, knowledge graphs, and governance layers (arxiv.org)) is key to reaping benefits while ensuring compliance.

As pharmaceutical regulation becomes ever more data-rich and complex, the need for intelligent automation will only grow. Agentic AI offers a path to handle this scale – potentially reducing time-to-market and freeing experts to focus on higher-level decisions. The journey will require multidisciplinary collaboration between AI researchers, regulatory scientists, and quality professionals. Continued research (as cited above) and pilot projects will refine these tools. Ultimately, a future in which an AI copilot helps craft precise, evidence-backed regulatory submissions – perhaps even conversing with the FDA's reviewers – may lie ahead. For now, the cautious but proactive approach is warranted: adopt agentic AI for document automation *today*, but under tight controls and human supervision, to accelerate drug development **safely** and reliably.

References

1. European Medicines Agency, *Electronic common technical document (eCTD)* (glossary), EMA Glossary: eCTD (accessed Apr. 2026) (en.wikipedia.org).
2. U.S. FDA, *New Drug Application (NDA)*, FDA resource (Jan. 2026), FDA NDA Guidance (accessed Apr. 2026) (www.fda.gov).
3. International Conference on Harmonisation (ICH), *Clinical Study Report Structure and Content (E3)*, ICH guideline (1995) (www.gmp-navigator.com).
4. Wikipedia, "Clinical study report," *Medicine* (Apr. 3, 2023) (en.wikipedia.org).
5. Eser *et al.*, "Human-AI Collaboration Increases Efficiency in Regulatory Writing," *arXiv* (Sept. 2025) (arxiv.org).
6. Wang *et al.*, "GMPilot: An Expert AI Agent For FDA cGMP Compliance," *arXiv* (Mar. 2026) (arxiv.org).
7. Yang *et al.*, "RegGuard: AI-Powered Retrieval-Enhanced Assistant for Pharmaceutical Regulatory Compliance," *arXiv* (Jan. 2026) (arxiv.org).
8. Kumar & Roussinov, "NLP-based Regulatory Compliance — Using GPT-4.0 to Decode Regulatory Documents," *arXiv* (Dec. 2024) (arxiv.org).
9. Ershov, "Compliance as Code with Graphs and Language Models: Regulatory Knowledge Graph," *arXiv* (Feb. 2023) (arxiv.org).
10. Gokhan *et al.*, "RIRAG: Regulatory Information Retrieval and Answer Generation," *arXiv* (Sept. 2024) (arxiv.org).
11. Hu *et al.*, "Enabling Regulatory Multi-Agent Collaboration: Architecture, Challenges, and Solutions," *arXiv* (Sept. 2025) (arxiv.org).
12. Syros *et al.*, "SAGA: A Security Architecture for Governing AI Agentic Systems," *arXiv* (Apr. 2025) (arxiv.org).
13. White *et al.*, "CliniDigest: LLM-Based Large-Scale Summarization of Clinical Trial Descriptions," *arXiv* (July 2023) (arxiv.org).
14. Ramprasad *et al.*, "Automatically Summarizing Evidence from Clinical Trials: A Prototype," *arXiv* (Mar. 2023) (arxiv.org).
15. Wikipedia, "Title 21 CFR Part 11 (Electronic Records; Electronic Signatures)," *Code of Federal Regulations* (May 2025) (en.wikipedia.org).
16. Wisbey & Craig, *"Agentic AI Explained: Key Concepts and Enterprise Use Cases," TechTarget* (Oct. 28, 2025) (www.techtarget.com).

Source links

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

1. https://arxiv.org/abs/2509.09738
2. https://arxiv.org/abs/2601.17826
3. https://arxiv.org/abs/2303.05392
4. https://en.wikipedia.org/wiki/Title_21_CFR_Part_11
5. https://arxiv.org/abs/2509.07006
6. https://www.fda.gov/drugs/types-applications/new-drug-application-nda
7. https://en.wikipedia.org/wiki/Clinical_study_report
8. https://www.fda.gov/media/89851/download
9. https://www.fda.gov/drugs/electronic-regulatory-submission-and-review/electronic-common-technical-document-ectd
10. https://www.fda.gov/media/71610/download
11. https://www.techtarget.com/searchenterpriseai/definition/agentic-AI
12. https://arxiv.org/abs/2504.21034
13. https://www.anthropic.com/engineering/building-effective-agents
14. https://www.fda.gov/media/135373/download
15. https://www.gmp-navigator.com/guidelines/gmp-guideline/ich-e3-structure-and-content-of-clinical-study-reports
16. https://arxiv.org/abs/2307.14522
17. https://en.wikipedia.org/wiki/Electronic_common_technical_document
18. https://database.ich.org/sites/default/files/E3_Q&As_R1_Q&As.pdf
19. https://arxiv.org/pdf/2509.09738
20. https://arxiv.org/abs/2603.20815
21. https://arxiv.org/abs/2409.05677
22. https://market.us/press-release/agentic-ai-market
23. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/part-11-electronic-records-electronic-signatures-scope-and-application
24. https://arxiv.org/abs/2509.09215
25. https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle_en.pdf
26. https://www.ema.europa.eu/en/glossary-terms/electronic-common-technical-document
27. https://arxiv.org/abs/2412.20602
28. https://arxiv.org/abs/2302.01842

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.