

Structured Product Labeling (SPL): Automation & AI Trends

2/23/2026 • 40 min read

- structured product labeling
- spl automation
- fda compliance
- pharmaceutical labeling
- regulatory affairs
- ai in pharma
- xml data standards
- hl7 standards
- generative ai
- electronic labeling



Executive Summary

Structured Product Labeling (SPL) is a regulatory data standard that has transformed how pharmaceutical and related products' labeling information is created, exchanged, and maintained. Developed under the HL7 Version 3 framework and formally adopted by the FDA in 2005 (^[1] www.pharmtech.com) (^[2] open.fda.gov), SPL defines an XML-based format for drug and device labeling. By encoding every element of prescribing and safety information in a machine-readable, controlled-vocabulary format, SPL enables automated validation, reuse of content (e.g. in package inserts or public databases), and systematic regulatory review. The transition from free-form text (e.g. PDF or paper labels) to structured, tagged content has yielded measurable benefits in data consistency and processing efficiency – for example, one industry review noted that thousands of labels can now be automatically indexed and cross-checked by health authorities, reducing manual errors (^[3] www.globalvision.co) (^[4] www.globalvision.co).

Despite these advances, most SPL creation workflows still involve significant manual effort. Regulatory teams often hand-author XML or use specialized tools to map content into the SPL schema (^[5] www.globalvision.co) (^[4] www.globalvision.co). The current state of automation ranges from basic schema-validation tools to sophisticated content management systems, but gaps remain. Errors in SPL are common: a study noted “even a single schema error triggers an FDA rejection” (^[4] www.globalvision.co), underscoring the need for robust tooling. As the volume of required labeling content grows (today DailyMed distributes on the order of 7–8 thousand active labels (^[6] dailymed.nlm.nih.gov)) and global regulations push for electronic labeling (e-labels/QR codes, IDMP), pharmaceutical companies face a scalability challenge.

Artificial Intelligence (AI) offers new opportunities to advance SPL automation. Generative language models can accelerate drafting of label text, while natural language processing (NLP) can extract and classify information from unstructured documents into SPL fields. Early surveys indicate strong industry interest: one report found ~70% of pharma leaders view AI as an immediate priority, especially in low-risk areas (^[7] www.fiercepharma.com). AI-enabled use cases in labeling include automated content generation (subject to expert review) (^[8] appian.com) (^[9] kpmg.com), machine-assisted translation for multilingual markets (^[10] appian.com) (^[9] kpmg.com), intelligent comparison of label versions (^[11] appian.com), and continuous compliance checks against regulatory changes. Notably, AI can help populate SPL documents from other sources (patient information leaflets, clinical study reports, etc.), and flag inconsistencies (for example, mismatched National Drug Codes or missing warnings). Leading vendors (e.g. Appian, KPMG) have outlined pilot solutions: Appian's platform uses embedded AI “skills” for label text generation, data extraction, and content summarization (^[8] appian.com) (^[11] appian.com), while KPMG's “AI LabelWise” touts GenAI-powered drafting, verification, and translation of labeling content (^[9] kpmg.com).

However, industry professionals remain cautious. In a 2025 survey of 40 pharma compliance experts, 65% said they do *not* trust AI to create regulatory submissions (^[12] www.fiercepharma.com), citing risks of “hallucinations,” lack of audit trails, and regulatory uncertainty. Nonetheless, many organizations are piloting AI for lower-risk tasks (e.g. internal review of promotional materials) (^[13] www.fiercepharma.com). The consensus view is that AI should be integrated into workflows with strong human oversight and explainability, ensuring that any machine-generated label content is verified by experts.

This report provides a comprehensive examination of SPL and its automation to date, and identifies where AI can augment this domain. We begin with the historical and regulatory background of SPL, followed by the current state of SPL automation (tools, workflows, obstacles). We then survey specific AI opportunities—drawing on expert discussions, case examples, and market strategies—and present data on technology trends and adoption. The final sections explore cross-functional impacts (e.g. digital health, e-labeling) and outline future directions. Throughout, we provide extensive citations to regulatory guidance, technical literature, industry analyses, and real-world examples.

Introduction and Background

Product labeling—the prescription information, warnings, and usage instructions accompanying drugs and devices—is a critical component of healthcare. Accurate labeling ensures that clinicians, patients, and supply chain managers have up-to-date, authoritative information for safe use. In the past, labels were created as free-text documents (paper leaflets or PDFs), which were difficult to update, reprint, or analyze. The push for electronic submissions and global harmonization, however, led regulators to require structured, machine-readable labeling formats.

In the early 2000s, the US FDA spearheaded efforts to digitalize labeling. In February 2004, FDA published Draft Guidance (21 CFR Parts 314 and 601) on “*Providing Regulatory Submissions in Electronic Format — Content of Labeling*”, laying the groundwork for SPL (^[1] www.pharmtech.com). On September 24, 2004, FDA announced it would begin accepting SPL submissions for drug labeling. By late 2005, it became mandatory: Pharmaceutical companies were required to submit all prescription drug labeling in SPL XML format, discontinuing PDF or Word submissions (^[1] www.pharmtech.com) (^[14] www.fda.gov).

Definition of SPL. The FDA characterizes Structured Product Labeling (SPL) as “**a document markup standard approved by Health Level Seven (HL7) and adopted by FDA as a mechanism for exchanging product and facility information.**” (^[2] open.fda.gov) HL7 is a standards-setting body (ANSI-accredited) for healthcare data; SPL is based on HL7 Version 3 (V3). The SPL standard defines schemas and controlled vocabularies for all parts of a label (e.g. ingredients, dosage, indications, contraindications, packaging). Each label is delivered as an XML document with precisely tagged sections. An Australian government health portal similarly notes that “Structured Product Labeling (SPL) is an HL7 standard that sets out the content on prescription drug labels in an XML format,” designed for machine-readability of ingredients, dosages, etc. (developer.digitalhealth.gov.au). In short, SPL transforms a label from an unstructured document into a structured data file.

Goals and benefits. The adoption of SPL and related structured content in labels was driven by safety, efficiency, and interoperability goals. By enforcing standardized formats and terminologies, SPL dramatically reduces human error in labeling. For example, an approved SPL can automatically republish the exact FDA-reviewed boxed warning into updated printed inserts without retyping (^[15] www.pharmtech.com). Regulatory agencies can then programmatically parse and aggregate labeling data across products, identifying inconsistencies or important safety signals across thousands of products (^[3] www.globalvision.co). Standardization also supports automation in review workflows: HL7 R3 schemas and validation rules allow FDA’s systems to instantly reject malformed label submissions (^[4] www.globalvision.co), thereby focusing human regulators’ efforts on content quality rather than formatting.

Ultimately, SPL aims to ensure that the “single source of truth” about a drug’s safe use is maintained accurately. Structured labels can be fed into prescribing software, clinical decision support, and public databases (e.g. DailyMed) to improve patient safety. As technologists note, “timely, accurate information means better treatment decisions with fewer errors” (^[16] www.powershow.com). Especially in an age of global regulatory submissions, SPL also facilitates sharing of approved content between agencies and countries (when combined with international standards like ISO IDMP).

Evolution beyond Rx drugs. Initially focused on prescription drugs and biologics, SPL has expanded. OTC drugs and device labels also use SPL as part of drug listings. More recently, the US has extended SPL mandates to other categories: for instance, under the Modernization of Cosmetics Regulation Act (MoCRA, 2022), the FDA now requires cosmetics product listings and facility registrations in SPL format (^[17] www.fda.gov). The FDA updated its SPL Implementation Guide in December 2023 to include cosmetics, signaling that SPL is becoming the de facto standard for any FDA-regulated product listing (^[18] www.fda.gov) (^[14] www.fda.gov).

Global regulatory bodies have adopted or are planning related structured formats. In Canada, Health Canada transitioned to a *Structured Product Monograph (SPM)* format analogous to SPL (^[19] www.schlafenderhase.com). European regulators historically used the XEVPRM (EudraVigilance Product Report Message) format, but are moving toward the ISO IDMP standards (particularly the SPOR framework) for standardized product identification and labeling (^[19] www.schlafenderhase.com). Ultimately, agencies across regions are cohering on similar principles: in each market, there is increasing emphasis on XML/IDMP-based electronic labels over paper. Table 1 (below) summarizes the major standards and timelines.

Region/Authority	Structured Label Standard	Mandate/Status (as of 2025)
USA (FDA)	HL7 V3 SPL (XML)	SPL mandated for drug labeling (21 CFR 314.601) since 2005 (^[1] www.pharmtech.com) (^[14] www.fda.gov); now extended to OTC, cosmetics (MoCRA 2022) (^[17] www.fda.gov).
Canada	Structured Product Monograph (SPM) (similar to HL7 SPL)	Adopted SPM for drug submissions, aligning with SPL concepts (^[19] www.schlafenderhase.com).
European Union (EMA)	XEVPRM (legacy); evolving to ISO IDMP/SPOR-based format	Traditional ePI via XEVPRM; EU Regulation requires mandatory ePI for orphan drugs (2023) and encourages broader e-labeling. EMA moving to IDMP standards (through SPOR) for product data (^[20] www.schlafenderhase.com) (^[19] www.schlafenderhase.com).
Other (e.g. Japan)	eCTD labeling modules (technical guidelines)	Most regulators accept eCTD submissions; some use standardized electronic label templates, but no global SPL mandate.

Table 1. Examples of structured labeling standards by jurisdiction and current status (sources cited).

The SPL standard itself consists of multiple document types and sections. At its core is the **Labeling Content** file, containing the patient package insert and prescribing information in richly tagged XML. Other document types include **Drug Listing** (listing of product codes and codes for NDC), **Establishment Registration**, **Health Care Provider Information**, and specialized annexes (e.g. Medication Guides). Within the labeling content, common HL7 terminologies are used (Argonaut coexists with LOINC codes for sections like “Contraindications”, and code systems like RxNorm can appear for ingredients (^[21] www.globalvision.co)). For example, each SPL label uses a *SetId* and *VersionId* to track document revisions, and embeds structured elements such as *ingredient sub-parts* with their quantities (^[21] www.globalvision.co).

Current State of SPL Automation

This section examines how SPL creation, management, and submission are currently handled in industry. We review the typical manual processes, the tools and platforms designed for SPL, and metrics on adoption and efficiency to the extent available.

Traditional (Manual) SPL Workflows

Most pharmaceutical companies still rely on semi-manual processes for SPL authoring. A common approach is for medical writers or regulatory specialists to draft the label in Word/PDF, then have HL7/XML specialists or FDA liaison staff manually transcribe that content into an SPL-authoring tool or hand-scripting the XML. This is labor-intensive and error-prone. Regulatory guidance emphasizes that even a minor XML mistake causes rejection, so teams must perform painstaking QA. As one industry note bluntly states, “one misplaced XML tag can send an entire drug submission back for correction.” (^[22] www.globalvision.co) Indeed, FDA’s automated gatekeepers strictly enforce SPL schemas: a submission with any schema or terminology violation will be kicked back without review.

The manual route has become increasingly untenable as workloads grow. Accelerated approvals and continuous postmarketing updates mean that labels change frequently (new safety bulletins, indications, or global harmonization). In the past, companies could afford to hand-code occasional label updates; today, regulatory staffs can face dozens of label changes per product per year. Without automation, this volume causes delays and risks: one mistaken NDC code, outdated study result, or omitted warning could trigger a serious compliance issue.

Table 2 compares key aspects of manual vs. automated SPL processes (including emerging AI extensions). In a manual workflow, medical writers create narrative label text and spreadsheets of data, then an XML expert encodes them into HL7 XML. Each new label or change must be re-validated against the current SPL schema, often using FDA-provided validation software. Even then, any subtle content-language mismatch might go unnoticed for days. By contrast, an automated SPL process leverages specialized software to enforce structure from the start, and may incorporate AI-based checks.

Process Phase	Manual Process	Automation & AI Opportunities
Content Authoring	Medical writers draft label text in Word/PDF. Required content often held in disparate documents (clinical reports, CSLR, etc.).	Generative AI (LLMs) can assist drafting. For example, specialized language models or AI-enabled authoring tools can produce first-pass label text based on standard templates, which human writers then revise (^[8] appian.com) (^[9] kpmg.com). This speeds creation of boilerplate sections (indications, dosage) while preserving expert review.
Data Entry / Structuring	Regulatory info (NDC, drug class, dose, etc.) is keyed into forms or XML tables by hand.	NLP / Knowledge Extraction: AI can parse existing documents (e.g. previous labels, study reports, spreadsheets) to auto-fill structured fields. Tools using named-entity recognition identify ingredients, strengths, frequencies, etc., populating SPL fields with high accuracy (^[23] appian.com).
Schema Validation	Before submission, XML is validated against the SPL schema using FDA tools or vendor platforms. Errors (tag mismatches, missing fields) cause rejection.	AI-Assisted QA: Machine learning can predict and flag likely schema violations before running formal validation. For example, anomaly detection models could learn from past rejections to catch subtle format issues preemptively. (Linked content checks, like cross-matching NDC and establishment numbers, can be automated as noted by industry sources (^[3] www.globalvision.co.))
Version Control & Audit Trails	Changes are tracked manually via filenames and documentation. Maintaining history can be cumbersome.	Integrated Document Management Systems: SPL platforms (often regulated RIM systems) automatically record every SPL version, author, and change reason in audit logs. While not strictly "AI," this is key automation. AI can further assist by analyzing change logs and summarizing revision histories.
Content Comparison	Teams manually compare new vs. old label versions (line-by-line proofreading).	Automated Diff Tools / NLP Comparison: Advanced algorithms (including LLM-based analysis) can compare two label texts and highlight meaningful differences. Appian notes that generative AI can automate "content comparisons" by flagging added or changed clauses between label versions (^[11] appian.com).
Translation Management	Labels translated by linguistic vendor; manual insert into each market's version.	Machine Translation with Post-Edit: Neural translation models can auto-translate labels into multiple languages. Reportedly, AI can route labels by market and translate, then forward to human reviewers (^[10] appian.com). This slashes turnaround for multi-country launches while still requiring human verification.
Regulatory Compliance Checking	Reviewers cross-check label against regulations/guidelines by hand.	AI Compliance Monitoring: Emerging systems continuously scan global regulatory updates and apply NLP rules to determine impact on labels. For instance, an AI agent could alert a R&D team if a new contraindication guidance is issued that may affect their product, enabling proactive label updates (^[24] appian.com).

Table 2. Manual vs. automated approaches in the SPL workflow. AI entries illustrate opportunities often cited by platforms such as Appian and KPMG (^[8] appian.com) (^[11] appian.com) (^[10] appian.com) (^[9] kpmg.com). All AI outputs require expert validation.

Software Platforms and Tools

To manage the complexity of SPL, many companies adopt specialized software. These solutions typically integrate with broader **Regulatory Information Management (RIM)** systems or content management platforms. Vendors like Lofware, ArisGlobal, Sparta Systems, and Veeva offer modules for SPL authoring, storage, and submission.

A common feature of modern SPL software is *automatic XML conversion and validation*. Instead of hand-coding XML, users input label content via a user interface or by importing source documents. The tool then generates compliant XML and runs consistent schema checks in real-time. As one industry guide notes, "teams see efficiency gains almost immediately after adopting SPL software" (^[25] www.globalvision.co): the system "converts documents into compliant XML, validates schema and structure in real time, and produces audit-ready reports before submission" (^[25] www.globalvision.co). By automating these routine tasks, companies can focus their human resources on content accuracy rather than format fixes.

For example, an SPL platform might automatically verify that a given National Drug Code has been properly registered under the associated manufacturer, as controlled vocabularies are cross-checked (^[3] www.globalvision.co). If a drug's listed NDC does not match an establishment registration, the software will flag the discrepancy. This type of automation aligns with the FDA's own systems: by design, "the FDA's systems can automatically verify that your NDC aligns with your establishment registration during drug listing processes," according to a case study (^[3] www.globalvision.co).

SPL tools also manage **version history**. Each time labeling is updated, the software records old and new content side by side with metadata (who made the change, rationale, date). This built-in version control is crucial for audits: instead of digging through file folders, teams can retrieve any prior approved version in minutes. Over time, this transforms auditing from a "stressful, last-minute task" into routine work (^[26] www.globalvision.co).

Integration is another key capability. SPL systems often plug into other enterprise platforms (e.g. content authoring, artwork management, or clinical trial databases) ⁽²⁷⁾ www.globalvision.co). This interconnectivity reduces duplicate data entry. For instance, if a structural change in a clinical study result necessitates a label change, data flows through the system linking those records rather than manual copying. In short, modern SPL solutions create a **continuous flow of approved data** from creation to submission ⁽²⁷⁾ www.globalvision.co.

Adoption Metrics and Trends

Industry reports suggest that SPL adoption is now widespread among major pharmaceutical companies, though precise statistics vary. Data from the FDA's public repositories underscores the scale: as of early 2026, the U.S. National Library of Medicine's DailyMed platform (which disseminates SPL documents) contained on the order of 7,000–8,000 active labels ⁽⁶⁾ dailymed.nlm.nih.gov). (By comparison, back in 2007 only ~2,273 labels were available, covering 78% of brand-name marketed drugs ⁽²⁸⁾ pmc.ncbi.nlm.nih.gov.) The jump from ~2.3K to several thousand highlights a near-complete market conversion over two decades.

Market analyses project continued growth. A 2023 report estimated the global *structured product label management* market at over USD \$53 billion, growing ~14% annually ⁽²⁹⁾ www.researchcorridor.com). Drivers include stringent regulations in healthcare and consumer packaged goods, plus industry digital transformation trends. Since labeling compliance is non-negotiable by regulators like the FDA, any efficiency gain (automation, centralized data control, AI) in labeling systems yields significant ROI.

Anecdotally, surveys of industry insiders indicate high SPL engagement. One account claimed ~73% of major pharma firms are piloting SPL programs, with 45% fully compliant ⁽³⁰⁾ docuvera.com). (Large pharmaceutical corporations purportedly approach 90% SPL adoption. ⁽³¹⁾ docuvera.com) Although such figures come from vendor literature, they align with the fact that FDA stopped accepting paper labeling back in 2005 ⁽¹⁾ www.pharmtech.com). Importantly, any approved new drug or new labeling supplement in the U.S. now must be in SPL – there is no optionality for traditional formats.

Internationally, adoption lags the U.S. timeline but is accelerating. For example, Health Canada's shift to SPM formalizes structured labeling for drugs, mirroring U.S. SPL requirements ⁽¹⁹⁾ www.schlafenderhase.com). The EU, under EMA oversight, has launched initiatives (like the Coalition for Clinical Trial Information) to promote e-labeling and IDMP. Although some markets still allow PDF/HTML patient leaflets (especially in low-demand regions), the global trend is clear: structured content is the future of labeling.

Despite this progress, challenges in implementation remain. **Technical expertise** is a major barrier: companies must hire or train staff fluent in XML, XSLT, LOINC, and other technical aspects of SPL. **System integration** is non-trivial: many legacy document-management platforms cannot natively handle SPL's complexity, requiring custom workarounds. And **change management** is needed to shift corporate culture from "writing a word doc and calling it compliant" to a disciplined XML authoring process. These factors explain why complete automation (zero-manual touch) is still aspirational for many organizations.

AI Opportunities in SPL

As the volume and complexity of labeling increase, life science companies are exploring how Artificial Intelligence (AI) can further automate and augment SPL processes. AI promises to tackle tasks that are currently manual and knowledge-intensive. This section details key areas where AI and machine learning can add value, along with real-world examples and reported benefits.

Natural Language Processing and Generation in Labeling

One major opportunity lies in leveraging Natural Language Processing (NLP) and Large Language Models (LLMs) for content creation and handling. Labeling involves large amounts of text (indications, usage instructions, warnings, etc.) which must be precise and consistent. AI tools can assist in multiple ways:

- **Automated Text Generation.** Generative AI can draft first-pass label sections based on underlying data. For instance, developers can train or prompt LLMs (GPT-like models) on a company's approved wording style and regulatory templates. The AI then produces suggested wording for doses, contraindications, or summary sections. Human experts review and edit, significantly speeding up the process. Appian reports that its platform's generative AI skill is used to *generate label text*, expediting label creation (with human oversight for verification) ⁽⁸⁾ [appian.com](#)). This can reduce rote writing: mundane details like reformatting an indication or dosage table can be done by AI, freeing writers to focus on new or complex clinical data.
- **Content Summarization.** Complex clinical documents (study reports, safety analyses, CCDS dossiers) often contain nuggets to be reflected in labeling. Traditionally, regulatory writers manually identify and distill the relevant points. AI-powered summarization can scan large text bodies and highlight key findings. For example, Appian describes using generative models to extract and summarize key data points from clinical trial documents and Company Core Data Sheets ⁽³²⁾ [appian.com](#)). This means instead of reading dozens of pages to find a new pharmacokinetic study result, a writer can ask the AI to pull out the "most relevant new safety data from Study XYZ," saving hours of work. These summaries then inform label updates (e.g. adjusting adverse reaction sections).
- **Automated Translation.** With global markets, labels must be prepared in many languages. Machine Translation (MT), especially neural MT, is now highly accurate for established languages. LLMs can automate translations of entire labels or individual sections. KPMG notes this benefit: its AI LabelWise offers "*rapid translation services... efficient and precise translations*" for labels, dramatically reducing the time and cost of preparing multi-language packs ⁽⁹⁾ [kpmg.com](#)). In practice, an AI system can detect the label's market and automatically invoke the appropriate language model; human linguists then perform targeted post-editing. This hybrid approach often outpaces traditional translation chains, slashing turnaround from weeks to days.
- **Content Comparison and Gap Analysis.** AI can compare versions of labels or compare proposed changes against current text. For example, when a new safety finding emerges, one must ensure all relevant products' labels reflect it. Manually checking each label's text is tedious. NLP algorithms or LLMs can align two text versions and highlight changes or omissions. Appian's content-comparison AI "analyzes old and new labels" to ensure critical updates are captured across related labels ⁽¹¹⁾ [appian.com](#)). Similarly, AI can compare a proposed label draft to regulatory guidelines (a knowledge base of rules), spotting clauses that may violate style or content rules.
- **Information Extraction.** Beyond summary, AI can *structure unstructured data*. For example, existing approved labeling (in plain text or PDF) can be fed into an AI to extract discrete fields (active ingredients, adverse reactions, etc.) into a database. This accelerates migrating legacy labels into SPL format. Preliminary research (e.g. using LLM-based retrieval QA or LLM+RAG frameworks) shows promise: one study used foundation LLMs to parse SmPC (Summary of Product Characteristics) documents and build an IDMP-compliant data model ⁽³³⁾ [pmc.ncbi.nlm.nih.gov](#)). While that study is experimental, it illustrates the concept: an LLM can map free-text paragraphs into standardized data fields. Such tools could be trained to specifically output valid SPL XML snippets given raw content.

AI for Quality, Compliance, and Review

AI also has roles in quality assurance and compliance tracking within SPL workflows:

- **Semantic Validation.** Traditional SPL validation ensures the file structure matches the schema, but does not interpret semantics. AI can perform a deeper contextual check. For instance, an LLM could be used to verify that the label text makes sense in context (e.g. no contradictory statements) or that numerical dosage instructions are consistent with drug strength and route. While still emerging, a form of AI-driven sanity check could flag potential logic errors that lone schema validators cannot detect.
- **Regulatory Intelligence.** Global labeling rules evolve continuously. AI agents can monitor regulatory databases and publications for changes (e.g. new FDA guidances, or updated WHO guidelines) and then map these changes to potential label impacts. For example, if the FDA issues a new safety alert for a drug class, an AI assistant could automatically alert product teams whose products fall in that class, prompting label review. Appian mentions "real-time compliance monitoring" where AI scans global regulatory updates and flags possible label impacts ⁽²⁴⁾ [appian.com](#)). In future, such agents could even suggest specific label text edits to meet new guidelines.

- **Multi-market strategies.** Different countries have nuanced labeling requirements. AI can assist in aligning label content across regions. For example, a label created for the U.S. can be automatically checked against EU-specific rules (and vice versa). A centralized AI-driven system can automatically route a label to country-specific workflows, as Appian notes for translation tasks (^[10] appian.com). This kind of global orchestration is complex to do manually, but AI (especially with knowledge of local regulations) can improve consistency and speed.
- **Audit Trail and Explainability.** One widely-cited concern is that AI decisions must be traceable. In labeling, it is legally critical to know *why* a certain text was included. Thus, any AI tool in SPL must produce an audit log or rationalization. Hybrid AI+human workflows (where AI suggests and humans approve) are commonly proposed to ensure accountability. For example, if an LLM drafts a section, the system logs the model version and prompt used so reviewers can trace the output back.

Industry Perspectives on AI in Labeling

Several industry players and thought leaders have begun articulating how they see AI fitting into labeling operations:

- **Appian (2025).** An Appian life-sciences consultant points out that many pharma companies currently deploy AI only in isolated ways (e.g. separate NLP tools, stand-alone translation services). They argue that to truly benefit, "AI needs to be in a process" – meaning integrated into end-to-end workflows (^[34] appian.com) (^[35] appian.com). Appian's model is to embed AI "skills" in its process automation platform. Their blog lists concrete AI features (mentioned above) and emphasizes security (Appian's Private AI). Key quote: "Labeling process involves lots of data...Data fabric gives AI access to exactly the data it needs" (^[36] appian.com). By leveraging cloud/GPU-powered intelligence, they promise to unlock latent efficiency in traditional RIM and labeling processes.
- **KPMG (2023/2024).** KPMG has developed an "AI LabelWise" solution for labeling. According to a KPMG article, its core benefits include *"Intelligent Content Creation: Utilizes GenAI for generating, verifying, and proofreading label content"* and *"Rapid Translation Services: efficient and precise translations"* (^[9] kpmg.com). Another highlight is *"Automated Data Consolidation"*: using AI to gather disparate product documents into a central database for labeling (e.g. aggregating chemistry/marketing files). KPMG frames these as ways to reduce the risk of human error and to shorten time-to-market. Though marketing, this narrative underscores confidence among consultants that AI can drive meaningful improvements.
- **Surveys and Research.** Independent observers have begun gauging industry sentiment. A FiercePharma summary of a 2025 Klick Health survey reported that a majority of pharma compliance professionals *do not* trust AI to author compliance submissions: 65% said "they don't trust the technology" for that purpose (^[12] www.fiercepharma.com). The top concerns were hallucinations (40% concerned), lack of audit trail (20%), and lack of explainability (12.5%) (^[37] www.fiercepharma.com). Notably, respondents were more receptive to AI for "review" tasks – i.e. using AI to check or review content rather than create it (^[38] www.fiercepharma.com). This suggests that a prudent near-term use case is AI-assisted reviewing/proofing (match to guidelines, consistency checks) while leaving final authorship to humans.

Another survey (Define Ventures) indicated ~70% of pharma leaders say AI is an immediate priority and ~80% are raising AI budgets (^[7] www.fiercepharma.com). This seemingly contradictory stance (high priority vs. low trust) reflects an industry at an inflection point: companies want to innovate with AI but need robust frameworks to ensure regulatory safety.

AI Case Studies and Pilot Projects

While full implementations are rare (SPL is highly regulated and risk-averse), several technology proof-of-concepts illustrate AI gains in related regulatory tasks:

- **Pharmaceutical Promotional Review (TCS).** Though not SPL-specific, it signals the power of AI in compliance. Tata Consultancy Services (TCS) built an AI system for a global pharma to review marketing materials. According to TCS, the AI solution cut review time by **88%** and caught over **90%** of potential compliance issues in medical content (^[39] www.tcs.com). It used state-of-the-art LLMs for text/image extraction, blended with a regulatory knowledge base to flag violations (^[40] www.tcs.com). This project demonstrates that even in very sensitive content (promotional claims), AI can significantly amplify human review. The same principles could be applied to label review: checking a label's claims and safety language against external regulations.

- **Pharma Compliance Verification (ThyncAI, 2023).** A case study reported on the use of an AI-driven platform (ThyncAI) for a pharma client to verify marketing materials against FDA guidelines (^[41] thyncai.com). The platform reportedly automates compliance checklists and ensures accurate terminology. Beyond promotional content, similar tools can check static label sections for compliance violations (e.g. forbidding certain promotional wording in medical context).
- **IDMP Data Extraction (Academic Research).** Researchers have used large language models to extract structured data for IDMP (medicinal product identification) from existing documents. For example, a preliminary experiment used foundation LLMs and retrieval-augmented generation (RAG) to convert narrative SmPC text into deeply nested data models for IDMP (^[33] pmc.ncbi.nlm.nih.gov). While still 'preliminary', this approach hints at a future where an LLM fine-tuned on drug labeling could auto-fill IDMP core elements (e.g. substance structure, dosage forms) from an unstructured label text.

Taken together, these examples suggest AI can automate routine, high-volume tasks in labeling and catch issues more reliably than humans alone. However, they also underscore the need for domain-specific tuning and governance; generic LLMs must be carefully controlled in life sciences contexts.

Data Analysis and Evidence

Quantitative data on SPL automation is scarce in the public domain, but several lines of evidence illustrate trends and impacts:

- **Label Count Growth:** DailyMed download statistics show a sharp uptick in SPL labels over time. The number of active labels grew from roughly 2,300 in 2007 (^[28] pmc.ncbi.nlm.nih.gov) to nearly **8,200** by December 2025 (then ~6,900 in January 2026) (^[6] dailymed.nlm.nih.gov). This reflects both natural company pipelines of NCEs and legacy conversion of older products. The increasing throughput implies that any inefficiency (e.g. manual rework) scales accordingly.
- **Submission Efficiency Gains:** Although formal published metrics are rare, industry claims hint at meaningful gains. One vendor article, for instance, asserts that an FDA analysis showed **45% faster processing** for SPL-compliant submissions and up to **78% fewer labeling errors** (^[42] docuvera.com). (Note: these figures are from a sponsored blog and should be interpreted cautiously.) Still, the basic premise holds: a submission formatted correctly the first time avoids weeks of back-and-forth for corrections. Given that labeling deficiencies historically have been a common cause of FDA complete response letters, the efficiency benefit is significant.
- **Survey Data on Workloads:** The FiercePharma piece noted that in some companies, over 50 new marketing or regulatory assets require review each quarter (^[43] www.fiercepharma.com). By analogy, companies that frequently update labels would see a large annual number of review cycles as well. Time savings in each cycle (say, by automating 20-30% of tasks) compound across the entire portfolio.
- **AI Adoption Statistics:** One Statista report (2025) indicated that, in pharma, about 50% of respondents had partially employed AI in at least one area (e.g., drug discovery, operations, safety) (^[44] www.statista.com). Specific to regulatory, surveys (as above) show ~80% boosting AI budgets (^[7] www.fiercepharma.com). Gartner and consulting firms predict that by 2025, a majority of life science companies will have at least pilot projects in regulatory processes. While a minority have fully automated SPL workflows today, that share is growing.
- **Regulatory Updates Monitoring:** The speed of regulatory change is accelerating. For example, the FDA's adoption of the PADDs database (due Jan 2023) and the EU's implementation of ISO IDMP timing (2022–2025) mean companies must churn out SPL-coded IDMP reports on fast schedules. AI can help meet such increasing demands, but quantitative forecasts on "how much faster" are still hypothetical.

Overall, the data points to a clear trend: **Structured labeling is now mainstream, and automation yields tangible improvements in compliance and efficiency.** AI promises to further enhance these gains, but evidence of tangible AI ROI in SPL specifically is just beginning to emerge.

Case Studies and Real-World Examples

While comprehensive academic studies on SPL automation are limited, several real-world implementations illustrate how companies and regulators approach the problem.

- **Electronic Product Information (ePI) in Europe:** In the EU and UK, regulators are progressively enabling e-labeling (electronic leaflets). A 2025 multinational study of 182 countries found that 74.7% allow use of QR codes on packaging to link to label info, and 67.3% accept PDF leaflets as formats (^[45] www.frontiersin.org). This “hybrid” approach (printed leaflet plus optional e-content) is becoming standard. For industry, this means label content typically exists both as SPL (for official submission) and as PDF/HTML for distribution. AI could help here by ensuring consistency: e.g., an updated SPL document could auto-generate the HTML for a QR code link. Case reports (like IntuitionLabs’ QR-code e-label map (^[46] intuitionlabs.ai)) show patchwork regulations; companies tackling global launches often rely on SPL frameworks internally, then convert to region-specific ePI formats.
- **Regulatory Training Initiatives:** Recognizing the SPL learning curve, agencies have begun systematic training. For example, the FDA offers formal SPL standard training for industry professionals. This underscores the novelty of the field. The FDA’s “SPL Standard Training” (ongoing as of 2025) indicates both the complexity of the standard and the expectation of full electronic compliance (^[47] www.fda.gov) (^[14] www.fda.gov). Companies investing in training implicitly acknowledge that tool support is necessary but not sufficient; intellectual investment is needed to use SPL effectively.
- **Vendor Case: ArisGlobal or Veeva:** Although detailed customer stories are seldom public, some vendors highlight successes. For instance, Veeva (a leading regulatory software firm) has case studies where biopharma clients using Veeva Vault RIM for labeling saw faster cycle times. One whitepaper noted that structured content reuse in Vault Labeling reduced content creation time by ~30% (source: Veeva, internal data). Similarly, ArisGlobal’s “LifeSphere” RIM clients have automated XEVPRM/SPL to meet EU regs; Aris commercials claim error rates dropped “by over 70%” with their consolidated content platform. While we lack independent verification, these industry anecdotes illustrate common logic: centralized structured repositories allow parallel global updates and reduce re-keying.
- **Regulatory Agency Data Initiatives:** Beyond compliance, structured labels are being exploited as public data. NLM’s DailyMed (based on FDA SPL submissions) now serves as a rich dataset for research and apps. For example, openFDA provides an API for drug labeling (CrashReports, adverse event mining, etc.) that relies on SPL-structured content (^[48] open.fda.gov). Data scientists can query this large corpus for pharmacovigilance signals or to build machine-learning models of drug–drug interactions. The conversion of labels into an open database is a societal benefit of SPL that is already in use.
- **Controlled Vocabulary Alignment (WHO/ISO):** In the IDMP space, there is an ongoing effort to align SPL with ISO’s definitions. The FDA plans to submit its SPL vocabularies (substance codes, dosage forms) into the global ISO IDMP Substance and Product dictionaries. This means SPL serves as a bridge: for example, “the IDMP framework seeks to establish a common language for identifying medicinal products,” and SPL labelling provides the content that populates those IDMP descriptors (^[49] www.pharmaceuticalcommerce.com). Pilot projects are under way where company product information (from SPL) is automatically mapped to ISO codes, a process ripe for AI assistance (entity recognition).

Implications and Future Directions

The automation of SPL via AI has broad implications for the pharmaceutical and healthcare ecosystems. This section discusses potential benefits, risks, and strategic considerations, as well as longer-term trends.

Efficiency and Cost Impacts

Automation promises to reduce manual labor and error costs. Even a conservative 20-30% time saving on label drafting and submissions could translate to millions in savings for large companies. By catching errors early, AI-assisted validation reduces the expensive resubmission cycles. More efficient labeling processes mean faster regulatory approvals, which directly impacts time-to-market and ROI on drug development. Given that “accelerated product approval timelines” were noted as a benefit of structured labeling adoption (^[42] docuvera.com), adding AI into the workflow could further compound these time savings.

Moreover, centralizing and automating translations and global reviews reduces the risk of a single missed local requirement stopping an entire shipment. Regulatory fines and supply disruptions (due to noncompliant labels) can be costly; improved automation mitigates that risk. We expect that as AI matures in this space, companies will reallocate some resources from rote tasks to higher-level regulatory strategy, improving overall productivity.

Quality, Safety, and Compliance

The ultimate goal of SPL is to protect patient safety through accurate information. AI can enhance that goal by making consistency checks more robust. For instance, an AI tool that flags when a black-box warning in one country's label is not present in another's can help eliminate safety oversights. Continuous AI monitoring of labeling means that when a safety alert or new study result emerges, companies can rapidly identify all affected products and update labels in hours instead of weeks.

However, new risks arise. Chief among them is "hallucination" – AI fabricating plausible but false content. In a compliance document, a fabricated data point could be dangerous. Therefore, any AI-written label content must be validated back against trusted sources (the clinical data, regulatory submissions). Compliance-minded companies may require human "sanity checks" for all AI output. This presumably limits immediate full automation; rather, AI serve as extremely fast drafting or checking aides under human control.

Legal and regulatory frameworks for AI use in this context are still emerging. The FDA has not issued guidance on AI usage in labeling specifically, but the general principles of evidence and provenance apply. Companies should maintain auditable records of how a label was generated. There is ongoing industry discussion about creating *explainable AI* logs for LLM outputs in pharma. For example, any text proposed by an LLM might be tagged with the source references it used. If formalized, such measures could satisfy regulators' demand for accountability.

Technological Evolution

Looking forward, several technology trends will influence SPL automation:

- **Integration with FHIR and IDMP.** HL7 has already released a FHIR R5 *SPL Mapping* Implementation Guide (^[50] packages2.fhir.org), illustrating a future vision where SPL data can interoperate with the broader HL7 FHIR ecosystem. In this model, labeling information could instantly populate FHIR Medication resources in healthcare systems. Further, ISO IDMP (the new global standard for product identification) is slated to require SPL generation as part of MA/Dossier submissions. Companies planning for 2027-2028 IDMP deadlines will likely leverage AI to map existing SARS to IDMP-compliant data (especially as manual coding of thousands of records would be prohibitive).
- **Electronic Labeling (e-Label, QR codes).** Technologies like QR codes and NFC tags are enabling "digital label" strategies. The global e-labeling studies (^[45] www.frontiersin.org) indicate wide adoption of QR codes to retrieve label info. Companies will need workflows where updating the central SPL document automatically refreshes the content linked by these digital codes. AI might streamline this process by generating the web-friendly HTML or JSON payload behind an e-label.
- **Pretrained Domain Models.** We anticipate the development of pharma-specific LLMs trained on regulatory text. An LLM fine-tuned on FDA-authored labels, or on the corpus of DailyMed, would likely hallucinate far less (its vocabulary would be constricted to regulatory language). Such models could be offered as regulatory writing assistants: imagine a "RegAI" that an RA specialist can query ("What is the approved USP for 5 mg Tablet Z in the US?") and get a direct, cited answer from the latest SPLs or guidances. Research in this direction is already underway (^[33] [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)).
- **Quality and Explainability Tools.** As advanced models become integrated, new kinds of QA tools will become important. For example, AI could compare multiple information sources for consistency. If an LLM writes a sentence about a drug interaction, a downstream AI might check it against a drug interaction database for plausibility. If mismatches occur, the system alerts the author. These meta-AI quality gates will likely be critical for industry adoption.

Challenges and Considerations

Several key considerations could limit or shape how AI is adopted:

- **Trust and Transparency.** As surveys show, regulatory professionals are uneasy about “black box” AI content. To build trust, solutions often emphasize a “human-in-loop” model. All AI-generated text in a labeling submission is treated as a draft needing sign-off, and the system records version-by-version who approved each change. Ideally, LLM outputs would also provide rationale fields (e.g. “These statements were generated based on the input summary with 85% confidence”). Technology to “explain” LLM outputs is an active research area; regulators may in future demand such explainability for any AI tools used in regulated content.
- **Validation of Training Data.** If companies train internal LLMs on proprietary documents (e.g. past labels, clinical data), they must ensure data privacy and correctness. In regulated industries, using real data to train general-purpose models is often restricted. Some vendors offer *private AI* where the company’s data is not exposed outside the system (^[51] [appian.com](#)). That approach will likely be favored until policies evolve.
- **Regulatory Clarity.** Currently, no major health authority has explicitly banned or fully approved AI use in labeling. This gray area means companies must tread carefully. Some firms may voluntarily disclose AI use in submission cover letters (similar to disclosing use of e-signatures or CDISC Standard). Insurance and liability are also questions: if an AI error leads to a mislabeling incident, who is responsible? Industry consortia and legal teams are likely to develop guidelines (for example, best practices for auditing AI outputs).
- **Global Harmonization.** Different regulators may have different views on structured content. While FDA has long embraced SPL, others are newer to the concept. Companies validating AI approaches must ensure compliance not only with FDA CFR but also with EMA/EU directives, PMDA (Japan), etc. However, the global convergence on XML/IDMP suggests alignment is increasing.

Future Directions

The future of SPL automation and AI integration can be envisioned along a few paths:

- **Full Content Lifecycle Integration:** Ideally, data flows from discovery to labeling seamlessly. For example, core data (e.g. chemical structure, acceptable trade names) is captured in early development systems; AI pipelines ensure this information automatically populates initial SPL drafts. In this vision, the SPL document is dynamically linked to the master product registry and pharmacovigilance database, with AI monitoring signals that move across systems. This is the “digital thread” approach that some advanced RIM vendors and agencies (FDA’s Center for Biologics Evaluation and Research) have articulated.
- **Routine E-label & E-brochure, Paperless Transmission:** As mobile device usage continues to rise, regulators may increasingly permit replacing paper inserts with fully digital ePI accessed by QR codes. Pharmaceutical companies would then use SPL primarily as backend data; consumer-facing labels might be HTML/CDA generated via AI from SPL. Already, some countries allow removal of paper leaflets in exchange for e-versions (^[45] [www.frontiersin.org](#)). AI could generate user-friendly patient leaflets in various languages directly from the structured SPL source, creating a single point of truth that serves regulators, clinicians, and patients in different formats.
- **AI Governance Frameworks:** We expect industry-led consortia to issue best practices for AI in regulatory content. Similar to how Good Clinical Practice (GCP) guides human trials, an “AI-GxP” framework may emerge covering how AI/ML tools should be validated, tested, and controlled in labeling processes. Standards bodies (GH-VE, ISO) might develop specific guidelines for AI use in regulated submissions. The FDA itself is exploring how to regulate AI in other areas (e.g., medical devices), which may spill over into labeling. Communities of practice (HL7, RAPS) will likely convene working groups on this topic.
- **Research and Real-World Evidence:** As data accumulates, we may see peer-reviewed studies measuring the impact of SPL automation. Academic-industry collaborations could analyze FDA review times or error rates pre- and post-AI adoption. Such evidence would bolster (or refine) hypotheses about efficiency gains. For example, a study could track a company’s submission error rate before/after implementing an AI check. To date, most published work has focused on SPL’s foundational benefits (^[42] [docuvera.com](#)) (^[28] [pmc.ncbi.nlm.nih.gov](#)); similar studies on AI’s incremental effect would close the loop.

Conclusion

Structured Product Labeling has fundamentally reshaped pharmaceutical regulatory documentation over the past two decades. What was once a laborious PDF-based process has become a data-driven workflow, with all key labeling content encoded in machine-readable XML. This transformation has delivered tangible benefits in compliance, safety, and efficiency (^[1] [www.pharmtech.com](#)) (^[25] [www.globalvision.co](#)). However, the journey is not complete: even with modern SPL software, many steps still demand expert human attention.

- [43] <https://www.fiercepharma.com/marketing/pharma-pros-skeptical-letting-ai-loose-regulatory-compliance-submissions-survey#:~:Relat...>
- [44] <https://www.statista.com/statistics/1538057/ai-adoption-in-pharmaceutical-industry-by-area#:~:AI%20...>
- [45] <https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1662782/full#:~:the%20...>
- [46] <https://intuitionlabs.ai/articles/e-labeling-regulations-by-country#:~:~e,Eur...>
- [47] <https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:Struc...>
- [48] <https://open.fda.gov/data/spl#:~:~The%20...>
- [49] <https://www.pharmaceuticalcommerce.com/view/challenges-and-benefits-of-implementing-idmp-in-the-pharmaceutical-industry#:~:Chall...>
- [50] <https://packages2.fhir.org/xig/resource/hl7.fhir.us.spl%7Ccurrent/ImplementationGuide/hl7.fhir.us.spl#:~:~Title...>
- [51] <https://appian.com/blog/acp/life-sciences/ai-pharmaceutical-labeling-process#:~:~With%...>
-

IntuitionLabs - Industry Leadership & Services

North America's #1 AI Software Development Firm for Pharmaceutical & Biotech: IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

Elite Client Portfolio: Trusted by NASDAQ-listed pharmaceutical companies.

Regulatory Excellence: Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

Founder Excellence: Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

Custom AI Software Development: Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

Private AI Infrastructure: Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

Document Processing Systems: Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

Custom CRM Development: Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

AI Chatbot Development: Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

Custom ERP Development: Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

Big Data & Analytics: Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

Dashboard & Visualization: Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

AI Consulting & Training: Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at <https://intuitionlabs.ai/contact> for a consultation.

DISCLAIMER

The information contained in this document is provided for educational and informational purposes only. We make no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

Any reliance you place on such information is strictly at your own risk. In no event will IntuitionLabs.ai or its representatives be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from the use of information presented in this document.

This document may contain content generated with the assistance of artificial intelligence technologies. AI-generated content may contain errors, omissions, or inaccuracies. Readers are advised to independently verify any critical information before acting upon it.

All product names, logos, brands, trademarks, and registered trademarks mentioned in this document are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, trademarks, and brands does not imply endorsement by the respective trademark holders.

IntuitionLabs.ai is North America's leading AI software development firm specializing exclusively in pharmaceutical and biotech companies. As the premier US-based AI software development company for drug development and commercialization, we deliver cutting-edge custom AI applications, private LLM infrastructure, document processing systems, custom CRM/ERP development, and regulatory compliance software. Founded in 2023 by [Adrien Laurent](#), a top AI expert and multiple-exit founder with 20 years of software development experience and patent holder, based in the San Francisco Bay Area.

This document does not constitute professional or legal advice. For specific guidance related to your business needs, please consult with appropriate qualified professionals.

© 2025 IntuitionLabs.ai. All rights reserved.