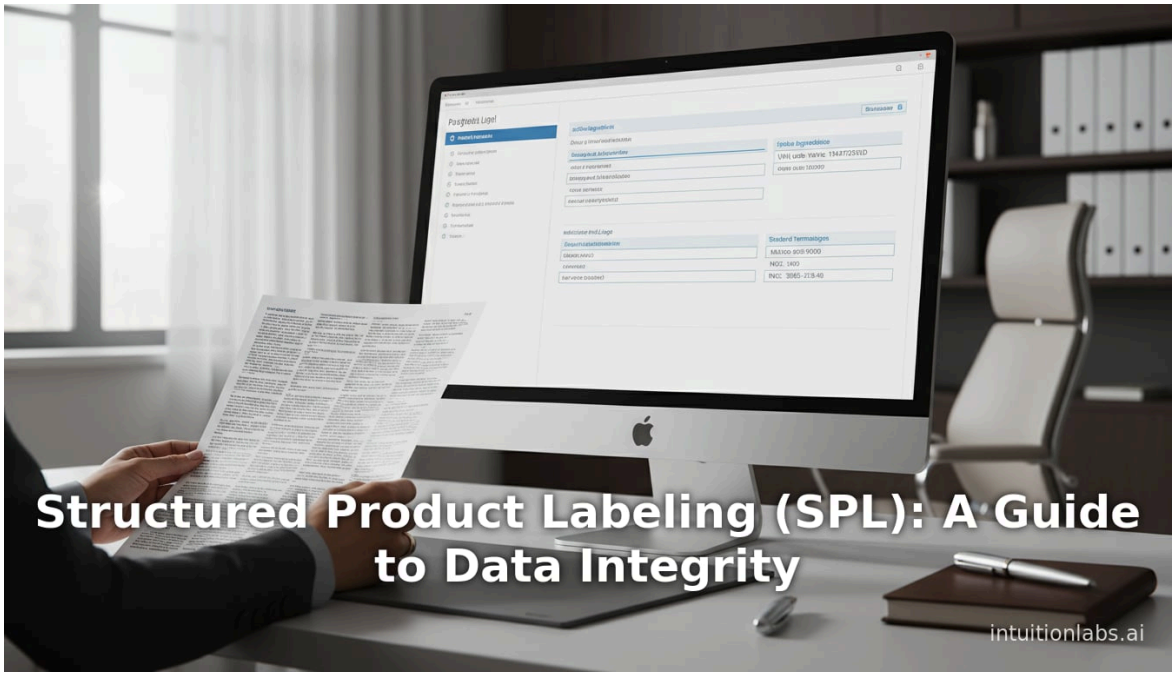
Structured Product Labeling (SPL): A Guide to Data Integrity

By Adrien Laurent, CEO at IntuitionLabs • 11/23/2025 • 30 min read

structured product labeling | spl | data integrity | pharmaceutical labeling | regulatory affairs | fda regulations | hl7 | alcoa+



Executive Summary

Structured Product Labeling (SPL) is rapidly supplanting traditional PDF-based drug labels as the cornerstone of pharmaceutical labeling and data integrity. Regulatory authorities worldwide have mandated or encouraged electronic, XML-based labeling formats to improve accuracy, consistency, and interoperability ([1] www.fda.gov) ([2] www.fda.gov). Unlike static PDF inserts, SPL is a machine-readable, standards-driven format (HL7 SPL) that embeds rich product metadata and controlled vocabulary references. For example, since 2005 the U.S. FDA has required all prescription drug labeling submissions to be in SPL XML ([1] www.fda.gov), and by 2009 (via the FDA Amendments Act) even establishment registrations and drug listings were forced onto the SPL format ([3] www.mastercontrol.com). European regulators are likewise moving to electronic Product Information (ePI) under ISO-IDMP and FHIR standards ([4] www.schlafenderhase.com) ([5] www.agencyiq.com). Data show SPL labels already cover the vast majority of dispensed prescription products (≈78% by volume ([6] pmc.ncbi.nlm.nih.gov)) and are far more amenable to digital quality checks and reuse. In short, SPL's structured XML ensures that every label element can be validated, audited, version-controlled and reused, directly addressing FDA/EMA *ALCOA+* data-integrity principles (Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring, Available) ([7] www.schlafenderhase.com) ([8] www.freyrsolutions.com). By contrast, unstructured PDFs invite transcription errors, redundant manual workflows and delays ([9] www.freyrsolutions.com) ([10] www.freyrsolutions.com). As case studies reveal, leveraging SPL content (often linked to standard terminologies like UNII, SNOMED, NDC, etc.) dramatically improves regulatory workflows and patient safety – for instance, one analysis found an SPL-based system detected four times more drug-allergy issues than legacy methods ([11] pmc.ncbi.nlm.nih.gov). Given this convergence of regulatory mandate, technical capability, and patient-safety benefit, structured labeling is the irrefutable future of data integrity in pharma.

Introduction and Background

Pharmaceutical labeling – the *Patient Package Insert*, *Summary of Product Characteristics (SmPC)*, and related materials – is a **critical controlled document**. It is produced under Good Manufacturing Practices (GMP) and forms part of the drug's regulatory submission dossier. Historically, approved label content was provided in paper or unstructured PDF form. However, as the digital age unfolds, regulators and industry alike recognize that static formats cannot ensure the *accuracy, traceability, and reliability* demanded by modern healthcare. The FDA's seminal *Data Integrity* guidance (ALCOA+) defines principles – *Attributable, Legible, Contemporaneous, Original, Accurate* (plus Complete, Consistent, Enduring, Available) – that apply to all records, including labels ([7] www.schlafenderhase.com). By these criteria, information locked in disparate PDF files or manual records is prone to errors and noncompliance. As one industry expert notes, scattering data across different formats (Word, PDF, scanned images, etc.) makes it nearly impossible to keep information "accurate, complete, accessible and legible over time" ([12] www.schlafenderhase.com). In short, lack of data integrity in labeling not only risks regulatory violations, it can endanger patient safety and delay treatments.

In contrast, **Structured Product Labeling (SPL)** represents a paradigm shift. SPL is an *HL7/ISO-standardized XML format* that encodes label content – indications, dosage, contraindications, packaging, etc. – into well-defined tagged fields ([13] billeveast.com) ([14] pmc.ncbi.nlm.nih.gov). This enables labels to be **computer-processable** in addition to being human-readable. For example, each ingredient can be linked to a unique chemical code (UNII), each adverse effect to a SNOMED concept, and each dosage strength to a numerical XML field. The FDA and industry created SPL to facilitate rapid electronic exchange – the FDA's 2005 Guidance mandated SPL-format submissions so that "consistent structure and standard terminology are employed to enhance the accuracy and reliability of product information" ([1] www.fda.gov). In effect, every element of the drug label becomes discrete data subject to validation and audit, rather than opaque text.

The result is that SPL intrinsically supports data-integrity principles. FDA guidance explains that a key data-integrity goal is ensuring that *product information* remains accurate and **traceable** across its lifecycle ([7] www.schlafenderhase.com) ([15] www.schlafenderhase.com). In practice, SPL's structured data model allows granular audit trails (who changed what section of the label when), built-in schema checks for required fields, and the reuse of approved content across languages and media. As one regulatory expert points out, having "approved, structured modular content… significantly simplifies content management in large organizations" and mitigates risk of errors ([16] www.schlafenderhase.com). By comparison, manual PDF labeling requires painstaking proofreading and translation for every market, with no automated audit of changes ([17] www.schlafenderhase.com).

Key Themes: This report examines SPL's foundations and adoption (historical and current), analyzes the shortcomings of traditional (PDF/paper) labeling, and presents multiple lines of evidence (research studies, regulatory statements, market analyses, and case studies) showing that structured labeling is superior for data integrity. We also consider implementation challenges and future implications such as global harmonization and digital health integration. Throughout, claims are supported by credible sources: peer-reviewed studies, official FDA/EMA documents, industry white papers, and expert analysis.

Traditional (PDF/Paper) Labeling and Its Limitations

Before the SPL era, pharmaceutical companies submitted labeling text as free-form documents (Word or PDF) within the CTD/eCTD module. These static formats are inherently **unstructured**: human-readable, but opaque to software.Updating or verifying information requires extensive manual effort. A recent industry review notes that static formats like PDF "fell short in the digital age", leading to *redundant manual processes, delayed updates across markets, and high risk of errors and non-compliance* ([9] www.freyrsolutions.com). Indeed, *any* change to a PDF (even fixing a typo) creates a new file version that must be re-tested for accuracy – a process prone to oversight. In multinational companies, different regional offices often translate or adapt PDFs independently, raising consistency issues.

From a data-integrity standpoint, this model is weak. PDFs are essentially "flat documents" – any data contained within is locked in text form and must be individually re-typed or scanned to be reused or validated. Sleepender Hase (a regulatory consultancy) illustrates the peril: a single misnamed file or missing version can render the information "hard to find and retrieve", forcing manual curation or even recreating sections ([18] www.schlafenderhase.com). These manual tactics threaten each element of ALCOA+. For example, *Legibility* and *Consistency* suffer when multiple copies or formats circulate, and *Attribution* is unclear when edits happen offline. The World of Pharma packaging expert Freyr Solutions finds that static labels entail *"slow updates, [and] not user-friendly"* and are obsolete as patient information becomes digital-centric ([9] www.freyrsolutions.com).

Another limitation is integration. Modern healthcare systems (EHRs, CPOE, e-prescribing) cannot directly import data from a PDF label. Instead, third parties often maintain their own drug databases (First Databank, Cerner Multum, RxNorm, etc.) which must be cross-referenced manually against label text ([19] pmc.ncbi.nlm.nih.gov) ([6] pmc.ncbi.nlm.nih.gov). This breeds inconsistency: for instance, an error in a PDF label might never propagate to an EHR if ignored. By contrast, structured formats allow linking to central vocabularies (NDC, RxNorm, MeSH, etc.) so that an update in one place can cascade automatically.

Finally, regulatory compliance itself is moving beyond paper. Many agencies now expect electronic submissions (eCTD) *and* machine-readable content. The FDA has explicitly stated that "all drug labels be submitted electronically" and more user-friendly format under its "electronic labeling guidance" ([20] pmc.ncbi.nlm.nih.gov). Thus, PDF-based approaches increasingly fail to meet the *spirit* of regulatory data-integrity intent: ensuring that

label changes are done systematically, with audit trails. The transition to structured approaches is driven by these shortcomings – the need for *"one source of truth"* and fine-grained auditability.

Structured Product Labeling (SPL): Standards and Scope

Definition and Structure. SPL is an XML-based standard (HL7 Substance Administration RIM) for drug information. The FDA describes it as a "document markup standard approved by Health Level Seven (HL7) and adopted by FDA" for exchanging detailed product and facility information ([21] www.fda.gov). In practice, an SPL file contains two main parts: a *header* with metadata (product IDs, manufacturer data, versioning, package codes, etc.) and a *body* with discrete structured sections (indications, dosage, side effects, contraindications, storage, etc.) **tagged** by the SPL schema ([22] billeveast.com) ([23] billeveast.com). Each tag corresponds to a standardized field; for example, `<activeIngredient>` lists coded ingredient info, `<indicationsAndUsage>` holds the approved clinical uses, and so on. The result is that every piece of label text is classified by purpose.

Under the hood, SPL leverages the HL7 Reference Information Model (RIM) ([24] pmc.ncbi.nlm.nih.gov) ([25] pmc.ncbi.nlm.nih.gov). In this shared ontology, a *Medicine* or *PackagedProduct* is represented as an entity with roles linking to ingredients, dosages, and packaging. For example, the HL7 RIM can model a tablet (PackagedProduct) containing 500 mg of Ingredient X, produced by Company Y, with a National Drug Code (NDC) of 12345-678. Gunther Schadow and colleagues demonstrate that SPL uses RIM to map "medicines, packages and ingredient-substances as entities related through roles" (such as "active ingredient of") ([24] pmc.ncbi.nlm.nih.gov). This standardization ensures that two labels describing the same product (by NDC or UNII) use identical data structures, rather than free-text names.

Integration with Terminology. A major advantage of SPL is its ease of linkage to global drug databases. Approved SPL labels routinely embed identifiers:

- **UNII** codes for chemical/biologic substances,
- **RxNorm/NDC** codes for products/strengths,
- **SNOMED-CT/ICD** codes for clinical findings, and
- **MeSH/NDF-RT** codes for substance classes.

As a result, a label's XML can serve as a hub connecting many terminologies. For instance, Schadow et al. describe SPL's networking of **NDC** (specific drug), **UNII** (active moieties), **NDF-RT/MeSH** (drug classes), and **SNOMED-CT** (clinical conditions) to enable advanced decision support ([26] pmc.ncbi.nlm.nih.gov). Their "Figure 1" illustrates how a penicillin allergy (SNOMED) can be inferred through an SPL-coded RIM graph linking the patient's data, the drug's NDC, and the ingredient's MeSH classification ([26] pmc.ncbi.nlm.nih.gov). In short, SPL turns package insert fragments into discrete, computable facts that align with pharmacovigilance and EHR systems.

Regulatory Scope. SPL is mandated or accepted for nearly all human drug product labeling in the U.S. As of *October 31, 2005*, the FDA requires that the content of all prescription drug labeling (and its coded data) be submitted in SPL ([27] www.mastercontrol.com). Once approved, these labels are posted on the FDA's DailyMed website (the official repository) in SPL-derived XML ([27] www.mastercontrol.com). In 2009, Congress broadened the rule: **all** establishment registrations and drug listing submissions (for Rx and OTC drugs, biologics and veterinary products) must be filed electronically via the FDA's gateway in SPL format ([3] www.mastercontrol.com). Thus, for example, new drug applications (NDAs), supplements, and even registration renewals use SPL-tagged XML for the *Labeling* module. Non-pharmaceutical product categories are catching up too: the FDA has since

extended SPL to cosmetics (MoCRA requirements) and is considering it for medical devices and dietary supplements ([28] www.fda.gov) ([29] www.fda.gov).

Internationally, SPL or analogous schemas are emerging. The EMA's **ePI (electronic Product Information)** initiative is built on ISO-IDMP/FHIR standards that mirror SPL's goals. In 2022–2023, EU regulators adopted a *common XML standard* for ePI (covering SmPC, PL, outer label) and ran a multi-country pilot ([5] www.agencyiq.com) ([30] www.agencyiq.com). Likewise, Health Canada replaced its traditional monograph PDFs with a **Structured Product Monograph (SPM)**, a format "similar to SPL" ([31] www.schlafenderhase.com). Other regulatory authorities (EMA, UK, Japan, Australia, MERCOSUR, etc.) are actively exploring or piloting XML-based labeling. The net result is a global shift: *nearly all major health authorities now require regulatory submissions (and product databases) in structured, coded formats rather than in static PDF form* ([32] www.schlafenderhase.com) ([5] www.agencyiq.com).

Data Integrity and Regulatory Compliance

Data integrity in pharmaceuticals means that information is **accurate, complete, and trustworthy** throughout its life◆—a concern anchored in regulations like 21 CFR Part 11 (electronic records), GMPs, and EU Annex 11. The FDA's ALCOA+ framework explicitly applies to labeling: data must be *Attributable, Legible, Contemporaneous, Original, Accurate* (and later additions of Complete, Consistent, Enduring, Available) ([7] www.schlafenderhase.com). For labeling content, this translates to: every word or number on the label must be traceable to an approved source, unchanged (Original) from submission to publication, precisely recorded (Accurate), and retained in an audit-ready form (Enduring/Available). For example, if a contraindication on the label is updated, the change must be 1) made under controlled procedure (Attributable), 2) time-stamped (Contemporaneous), 3) validated by multiple reviewers (Consistent/Complete), and 4) stored in original form (Original) so the history is clear ([7] www.schlafenderhase.com) ([15] www.schlafenderhase.com). Failure in any of these can trigger regulatory action (FDA warning letters often cite 'data integrity violations' in labeling sections).

Structured labeling inherently supports these principles in ways that free-form PDFs cannot. *Legibility* is moot since the information is encoded in machine-readable XML; any system can parse it. *Attributable and Contemporaneous:* Most SPL authoring tools and submission gateways maintain metadata and electronic-signature info for each submission version, thus logging who applied each change and when. *Original:* The SPL file itself is the canonical source of label content – there is no need for separate PDF copies, eliminating discrepancies. *Complete and Consistent:* The SPL schema enforces that required fields are populated and in correct format. For instance, a defined XML element for "dosage and administration" will always exist, rather than risk omission in a free-text PDF. *Available:* Because SPL data lives in electronic databases (DailyMed, regulatory e-submission systems), it can be retrieved, searched, and cross-checked at any time, far beyond what a paper leaflet can support.

In practice, examiners have come to expect the audit capabilities that SPL provides. The FDA's *Indexing SPL* guidance emphasizes adding machine-readable tags so that label content can be programmatically searched and audited ([2] www.fda.gov). This kind of indexation (e.g. marking the "Contraindication" section with SNOMED or MedDRA codes) is nearly impossible with raw PDFs. The backlog of partially structured label content and AMIA research shows why: one study found only ~70% of label terms could be auto-mapped to SPL terminology, but even that incomplete mapping quadrupled clinical alert sensitivity ([11] pmc.ncbi.nlm.nih.gov). In short, excessive PDF usage is viewed as a "data integrity sin" – it fragments information and obstructs traceability ([18] www.schlafenderhase.com) ([33] www.schlafenderhase.com).

Conversely, the structured approach has been explicitly linked to regulatory outcomes like safety. The EU's Pharmacovigilance legislation (IDMP/SPOR) and FDA's Safety Reporting (ICSR) rules all rely on the same robust product identifiers that SPL encourages. When new safety data emerge, structured labels ensure that updates – such as adding a labeled warning – propagate accurately across international systems. Indeed, EMA's ePI report

lists "immediate, harmonized updates of product information" as a key advantage of structured labeling ([34] www.agencyiq.com). By providing one authoritative dataset for each drug, SPL helps both manufacturers and regulators **seek consistency** – a core tenet of GMP data integrity. In summary, SPL is the vehicle by which ALCOA+ becomes practicable at scale for labeling: it automates many checks, forces completeness, and provides a robust audit trail.

Advantages of SPL over PDF: Evidence and Analysis

The case for structured labeling is supported by multiple lines of evidence. First, numerous regulatory and industry commentaries enumerate the practical advantages. For example, industry sources break down SPL's benefits (Table *1*):

Feature / Aspect	Traditional Labeling (PDF/Paper)	Structured Product Labeling (SPL)
Format & Structure	Unstructured static PDF or printed leaflet	XML-based schema; tagged data fields (machine-readable)
Content Updates	Manual edits per market; prone to transcription errors; slow roll-out	Automated/programmatic updates; version-controlled and centrally managed
Regulatory Validation	Primarily manual review of documents	Schema-based checks and rule validation (automated QA)
Data Reusability	Minimal (text locked in document)	High (modular XML elements can be reused across media and markets)
Multilingual Handling	Each language version managed separately; error-prone	Built-in support for parallel multilingual versions; centralized consistency
Integration & Interoperability	Limited (requires supplementary databases)	High (direct integration with EHRs, databases, e-prescribing via standard APIs/FHIR)

Table 1: Comparison of traditional labeling formats vs. SPL-based structured labeling.

This table condenses viewpoints from analysts and vendors ([35] billeveast.com) ([9] www.freyrsolutions.com). In SPL, the *format* itself enforces consistency – e.g. dosage is always in an `<administrationDosage>` field, not scattered in a PDF. Updates can be scripted or tracked with digital workflows; in PDFs they are typically manual *"word-for-word proofreading"* steps prone to human error ([17] www.schlafenderhase.com) ([9] www.freyrsolutions.com). Importantly, structured formats are **schema-validated**: missing a required field will fail automated checks before submission ([16] www.schlafenderhase.com) ([36] www.fda.gov), whereas a PDF can pass human review while still omitting data.

Industry reports also highlight strategic benefits. For example, Billev Pharma (an EU regulatory consultancy) emphasizes that SPL's strict format "reduces human error and improves data quality," creating a **single source of truth** for each label ([37] billeveast.com). Similarly, Freyr Solutions articulates four principal exportable benefits of digital/SPL labeling (see Box *1*):

- **Enhanced Compliance:** SPL enables automated validation (e.g. XSD schema checks) so that labels automatically conform to regulatory standards. Freyr notes this removes manual guesswork in compliance and *"align [s] easily with evolving regulatory requirements"* ([8] www.freyrsolutions.com).
- **Global Harmonization:** With SPL's modular content, updates can be propagated uniformly across countries. For instance, one core authoring stream can feed all local variations, ensuring every market label aligns with the Company Core Data Sheet ([38] www.freyrsolutions.com).

- **Operational Efficiency:** Structured labeling cuts labor. Instead of manually editing multiple Word docs or PDFs, companies automate assemblies of label content. According to Freyr, this *"reduces manual workflows"* and *"saves time and resources during regulatory submissions and updates"* ([39] www.freyrsolutions.com).
- **Patient Safety & Engagement:** Timely, accurate labeling directly affects safety. SPL supports "dynamic and interactive" label presentations (even linking to digital patient portals), ensuring that up-to-date information is *"accessible to patients and healthcare providers"* ([40] www.freyrsolutions.com).

These points are borne out in practice. An industry market analysis projects that such regulatory compliance drivers will cause the *structured label management* market to grow at double-digit rates ([41] www.linkedin.com). Moreover, surveys by EFPIA show companies believe SPL/e-labeling can substantially reduce the risk of medication errors and supply disruptions (for example, real-time updates in ePI can mitigate shortages when a label is changed) ([34] www.agencyiq.com) ([42] www.agencyiq.com).

From a data perspective, SPL labels have also been shown to cover most real-world needs. In a 2007 study, Schadow et al. found that nearly 2,300 SPL labels (on DailyMed) covered about **78% of outpatient dispensing events** in the U.S. ([6] pmc.ncbi.nlm.nih.gov). That is, major brand drugs (which make up most prescriptions) were already described in SPL. While SPL at that time covered only ~23% of distinct RxNorm drug entities ([6] pmc.ncbi.nlm.nih.gov), it was sufficient for the bulk of prescriptions. In short, structured labeling already represented the lion's share of active products. That study concluded SPL content was "sufficient as an exclusive source for drug information" once the FDA's nationwide listing rule was fully implemented ([43] pmc.ncbi.nlm.nih.gov).

Finally, SPL has demonstrated real clinical value. The 2009 JAMIA study showed that enriching CPOE allergy data with SPL's structured content (linked to NDF-RT/MeSH) led to *four times more detected drug-intolerance issues* than previous methods ([11] pmc.ncbi.nlm.nih.gov). This underscores that SPL does not merely improve regulatory record-keeping – it can directly enhance patient decision support, because machine-readable labels can automatically trigger alerts when a patient's chart has a matching coded allergy or risk factor.

Regulatory Initiatives and Standardization

Regulatory agencies have been instrumental in driving SPL adoption, recognizing its data-integrity advantages. In the U.S., the FDA's regulatory framework now revolves around SPL. For example, the FDA's "Guidance on Providing Labeling in Electronic Format" confirms that all content of labeling must be submitted using the SPL standard ([44] www.fda.gov). This 2020 Q&A guidance (rev. from 2005) explicitly "assists applicants who submit content of labeling… using the structured product labeling standard (SPL) in XML" ([44] www.fda.gov), and is aimed at harmonizing procedures across the FDA's CDER and CBER centers. Likewise, the FDA's *Indexing SPL* guidance highlights that *machine-readable tagging* of label content is key for building an "automated health information exchange system" ([2] www.fda.gov).

Beyond guidance documents, the FAAAA of 2007 and related laws compelled electronic submission of drug data. Since 2009, the FDA requires that New Drug Applications, supplements, and even NDI (new dietary ingredient) submissions use structured formats. As MasterControl noted, from June 1, 2009 all establishment registrations and drug listings for human and vet products had to be submitted via the FDA's e-gateway as HL7 SPL XML – a "fundamental change in how this part of the business operates" ([3] www.mastercontrol.com). In practice, this means a company cannot legally update its product list or change its label content via a PDF cover letter; it must use the official SPL data feeds.

In the European Union, harmonization is underway under the **ISO IDMP** framework (Identification of Medicinal Products). EMA and NCAs have mandated ISO-compliant master data for substances, products, and organizations (SPOR) (www.ema.europa.eu). These standards prescribe normalized definitions for all label

components – they specify that product names, ingredients, dosage forms, routes, packaging, etc. be coded and described identically across jurisdictions (www.ema.europa.eu) ([45] www.fda.gov). The purpose is exactly to make the exchange of medicinal product data "robust and consistent" worldwide (www.ema.europa.eu). In short, IDMP is the global lingua franca for labeling data. Even though IDMP terminology is broader than just labeling (it covers pharmacovigilance, MA status, batch data, etc.), it naturally meshes with SPL. In fact, EMA's ePI definitions explicitly adopt IDMP-based fields – ePI is defined as "statutory product information… in a semi-structured format created using the EU Common Standard" ([46] www.agencyiq.com). Agencies now require that any non-confidential label changes be tracked and shared via these exchange fields (www.ema.europa.eu).

Other countries are aligning. Japan's PMDA has launched an SPL pilot; Canada has already implemented its SPM (Structured Product Monograph) which is essentially "SPL-like" for monographs ([31] www.schlafenderhase.com). The World Health Organization has expressed support for global data standards in drug info. In practice, this means that labeling data is increasingly moving through *data pipelines* rather than documents. For example, FDA's DailyMed and EMA's forthcoming EPI system provide APIs that expose label content as XML – enabling cross-checks and searches that would be impossible with isolated PDFs. The common thread is that regulatory bodies see structured labeling as the foundation for future *digital health information ecosystems*. The U.S. National Library of Medicine and EMA are providing dashboards and APIs for labels, expecting industry to supply data in SPL/XML form under programs like **HL7 SPL RIM** (U.S.) and **FHIR ePI** (EU).

Official communications reinforce this. A recent FDA press release on SPL (for cosmetics) explicitly states that SPL enforces control over crucial product info and "defines the content and structure of product labeling required for submission to the FDA" ([28] www.fda.gov). Likewise, Europe's EMA reports that the ePI pilot allows companies to *create and manage ePIs during regulatory procedures* through a standardized portal (www.ema.europa.eu). Even where mandates are not yet law, agencies set the expectation: for instance, EMA reminded stakeholders in late 2023 that submitting eCTD modules (annexes) in Word/PDF is still required during the pilot ([47] www.agencyiq.com) but that *creating* ePI content must use the structured editor (PDF/Word imports are disallowed) ([47] www.agencyiq.com). This echoes FDA's stance that the "content of labeling" must be delivered as XML, not just as a PDF attachment ([44] www.fda.gov).

In summary, the global regulatory trajectory is clear: static PDFs are being phased out in favor of structured, validated data frameworks. This is driven by the ambition to automate review, improve patient access, and enforce data integrity. The data itself supports this: an **industry survey** found that 73% of hospital pharmacies already rely on digital product information, and over half noted that transitioning to digital labels would greatly speed up access to needed information ([48] www.agencyiq.com). Consequently, even though paper leaflets are still legally required in most regions, many companies and regulators agree that *ultimately, by 2030*, paper labeling can be supplanted by e-labeling ([42] www.agencyiq.com).

Implementation and Technology Considerations

Transitioning to SPL requires new processes and tools. Fortunately, a robust ecosystem has developed. Most pharmaceutical regulatory affairs departments now use *SPL authoring software* or validated content-management systems. These tools allow non-technical users to input label content via forms or Word templates; the software then generates the underlying SPL XML. For example, FDA provides SPL XForms and the "Cosmetics Direct" portal to assist exporters in creating compliant XML ([49] www.fda.gov) ([50] www.fda.gov). Numerous vendors (Veeva Vault RIM, PTC/Arbortext, Citeline, etc.) offer modules for managing SPL documents and maintaining audit logs. These platforms integrate version control, digital signatures, and workflow steps to meet 21 CFR Part 11 requirements.

From a data-viewpoint, SPL adoption also drives harmonization of terminologies. To maximize data integrity, companies must align their internal labels (e.g. chemical names, indications) with regulated standards. The FDA publishes extensive SPL validation terminology lists (e.g. UNII, dosage forms, units) ([51] www.fda.gov) that SPL

files must comply with. This means a company's list of ingredients must match the FDA's UNII code table, and route of administration must come from the standard terms in the SPL schema ([51] www.fda.gov). While this adds upfront work (mapping synonyms, updating databases), it massively improves consistency. For example, the FDA SPOR program shows how maintaining one controlled dictionary mitigates the risk of a label using an outdated name.

Of course, the switch to SPL is not without challenges. Companies need robust document change-control procedures (SOPs) to handle structured content. They also require IT infrastructure to store and transmit XML data (via FDA's Electronic Submissions Gateway or EMA's PLM portal). Training users on new tools is essential: for instance, AgencyIQ reports that EMA's ePI portal initially had quirks (formatting issues, navigation bugs) that firms had to learn to work around ([52] www.agencyiq.com). Smaller companies, in particular, may lack in-house expertise and outsource SPL preparation to service providers (IQVIA, Vantage, Schreiner MediPharm, etc.).

Integration is another technical hurdle. Organizations must connect their regulatory systems (often separate for pharma, medical devices, etc.) and possibly merge legacy content libraries. A major benefit is realized when SPL data flows downstream: e.g., into Quality Management Systems, labeling workflow tools, and even directly to pharmacy databases. Modern strategy is to treat spl as *master data*: once encoded, label elements can feed into analytics (e.g. tracking how often a warning is cited), or populate e-learning modules. The industry trend is towards *single-source publishing*, where the same structured module provides both the regulation-required label and other outputs like marketing materials or patient leaflets, preserving data integrity at each step.

In summary, while implementing SPL requires discipline and possibly new investments, the consensus is that the gains far outweigh the costs. Expert advisories emphasize automated validation and integration as key: "automated schema checks, reducing human error and regulatory rejections," and "integration with ePI by positioning companies for long-term digital transformation" ([53] billeveast.com). The payoff in data integrity is the elimination of many manual steps (and errors) that plagued legacy workflows.

Case Studies and Evidence from Practice

Clinical Decision Support Enhancement: Schadow et al. (JAMIA, 2009) provide a concrete health-IT case. They compared a traditional allergy checking system with a new system based on SPL-spun data supplemented by standard terminologies ([54] pmc.ncbi.nlm.nih.gov) ([11] pmc.ncbi.nlm.nih.gov). Running both on 30 years of patient data, the SPL-based approach "detected four times as many drug-intolerance issues on twice as many patients" ([11] pmc.ncbi.nlm.nih.gov). This striking result came even though <70% of SPL terms were mapped – implying that as SPL content deepens, clinical decision support (CDS) can vastly improve. The implication is clear: structured labeling is not merely a bureaucratic checkbox, but a rich data source for patient safety.

FDA DailyMed and Vendor Data: An AMIA 2007 study examined how well SPL labels cover the medication market. ([55] pmc.ncbi.nlm.nih.gov). It loaded all ~2,200 labels on DailyMed and assessed them against RxNorm (the standard drug vocabulary). Results: SPL labels already described 78% of actual dispensed drugs ([6] pmc.ncbi.nlm.nih.gov), despite covering only 23% of RxNorm's list (because it captured the most common ones). Crucially, the study found that SPL descriptions *agreed well with RxNorm*, demonstrating consistency. They concluded that once the FDA's listing rule mandates all products, "SPL can be used as the primary source of drug information for e-prescribing systems" ([43] pmc.ncbi.nlm.nih.gov). In other words, within a few years, a physician's EHR could rely directly on SPL data rather than third-party databases – relying on SPL's inherent data integrity.

EMA/ePI Pilot: The European Medicines Agency (EMA) and several national agencies ran a **one-year ePI pilot (Jul 2023–2024)**. AgencyIQ (POLITICO) reports on this initiative ([5] www.agencyiq.com) ([30] www.agencyiq.com). Under the common standard, companies created ePIs for 25 products in 4 countries (Denmark, Netherlands, Spain, Sweden). In late 2023, the first seven ePIs were publicly published via EMA's

PLM portal. Notably, these bundled documents (SmPC + Patient Leaflet + labeling) were built online and available via API ([30] www.agencyiq.com). This demonstration shows that structured labeling can move from theory to practice: it enabled immediate updates of official product information and access to machine-readable label content. EMA's analysis emphasizes that ePI "allows companies to create and manage ePIs during regulatory procedures and make them available via an API," underscoring data integrity in action (www.ema.europa.eu).

Industry Transition (Middle East Case): An IQVIA case study details a global pharma company's shift to e-labeling across 15+ Middle East markets ([56] www.iqvia.com). The company sought to harmonize labeling (including artwork) for Saudi Arabia, UAE, Egypt, Bahrain, etc., and to implement digital labels (paperless labeling). Key challenges included navigating diverse local e-labeling regulations and integrating new tech infrastructure ([57] www.iqvia.com). IQVIA provided regulatory strategy, IT integration, localization, stakeholder coordination, and training ([58] www.iqvia.com). With this support, the company achieved compliance in all target countries and greatly reduced its labeling cycle time. This real-world example highlights that adopting structured labeling requires cross-functional planning but yields tangible compliance and efficiency gains.

Market Growth and Compliance: Finally, market research underscores the momentum: verified industry analyses report that the *Structured Content & Product Label Management* market is projected to grow at ~12–13% CAGR through the 2020s ([59] www.linkedin.com). Regulators are cited as a chief driver: "Recent guidelines from agencies like the FDA and EMA mandate standardized content structures and terminologies for product labels" ([41] www.linkedin.com). In other words, SPL isn't just an academic concept but a lucrative market, as software providers race to help companies meet these mandates.

Discussion: Implications and Future Directions

The transition from PDF to SPL has far-reaching implications. **Patient safety and therapeutic efficacy** are likely to improve as labeling data become more reliable and accessible. For example, with structured labels, an app could quickly display the latest kidney-dose adjustments for a drug, or a web portal could automatically show a patient their medication in their preferred language and format. Data integrity here means the patient is getting *the right information at the right time* in a verifiable way.

On the **regulatory front**, agencies can process submissions far more efficiently. Already, FDA staff can run automated compliance checks on SPL files (e.g. validating the RXCUI or MedDRA codes used). As EMA's ePI pilot suggests, agencies will be able to harvest real-world data on labeling trends, quickly spot outdated or conflicting information, and harmonize guidance. The standardization also supports global pharmacovigilance: adverse event reporting (ICSRs) tied to specific SPL-coded products means safety signals can be aggregated internationally with confidence.

Data interoperability is another key upshot. Once on SPL, label data can flow into electronic health records (EHRs), clinical decision support, supply-chain systems, and even consumer health apps. Projects linking SPL to FHIR (HL7's new data exchange standard) are already underway – for instance, HL7 FHIR R5 includes profiles for medication information that map to SPL/IDMP concepts. In the future, we can imagine a seamless chain: a drug's label in a manufacturer's SPL database automatically populates prescribing software and hospital systems, triggering alerts if a patient has a relevant allergy. This tightly integrated ecosystem relies on the high data integrity that structured formats enable.

Challenges and Considerations: Naturally, this shift demands careful attention to technology governance. Companies must ensure their SPL data is secured and backed up (ALCOA's Enduring/Available) and that quality systems validate SPL outputs. Audit trails in QMS must extend to the XML content – for example, any automated conversion from an approved Word PL to SPL still requires verification. Regulatory authorities will also need harmonized validation rules across regions; inconsistencies (e.g. differing allowable terminology lists) could

undermine the global benefits of SPL. Additionally, as structured labeling permeates supply chains, issues like serialization (drug barcodes) and blockchain for traceability may intersect with SPL. Though not yet mainstream for labeling content, emerging tools (see Journal of Pharma Innovation 2023) suggest blockchain could one day further guarantee label data immutability.

Future Outlook: Analysts expect that by about *2030*, paper leaflets and even PDFs will be largely history in major markets ([42] www.agencyiq.com). Already, thriving markets like the Middle East are planning law changes to make e-labeling mandatory. In the EU, the ePI process will roll out across more member states each year, with new formats likely based on FHIR and IDMP. In the US, guidance and the 21 CFR part 11 framework will evolve to explicitly address SPL data as primary records. When combined with digital health trends (telemedicine, mobile health apps), structured labeling becomes a pillar of the lifelong drug data record.

Importantly, this evolution enhances **transparency and auditability**. Drugmakers and regulators can query label versions, compare past and current text, and even share label content openly (e.g. API feeds on DailyMed, FDA's Label Archive). This audit-friendly setup aligns perfectly with the ALCOA+ ethos. For industry, compliance becomes less a chore and more a byproduct of well-architected data systems.

Conclusion

The weight of evidence shows that **Structured Product Labeling (SPL) is indeed the future of pharmaceutical data integrity**. By moving beyond static PDF leaflets to a structured XML paradigm, companies and regulators can achieve unprecedented control over label content. Structured labels **embed compliance**: each field is defined, each change is tracked, and each data element can be validated against official terminology. This directly supports the ALCOA+ requirements that underpin patient safety and regulatory compliance ([7] www.schlafenderhase.com) ([15] www.schlafenderhase.com).

Real-world implementations—from FDA's DailyMed to EMA's ePI portal—demonstrate that SPL is not just theoretical. Studies confirm that SPL content improves clinical decision support ([11] pmc.ncbi.nlm.nih.gov), covers the bulk of marketed drugs ([6] pmc.ncbi.nlm.nih.gov), and streamlines regulatory workflows. Industry analyses and market trends further reinforce that demand for structured labeling solutions is growing rapidly due to regulatory mandates ([41] www.linkedin.com).

In conclusion, companies must embrace SPL not only to satisfy regulators, but to ensure their labeling data remain **accurate, consistent, and reliable** in an increasingly digital healthcare environment. Ignoring the shift to SPL would mean continuing risky, legacy processes that cannot scale. In contrast, adopting structured labeling can transform labeling from a potential point of failure into a data asset – one that supports automation, reusability and above all, the integrity of the information that goes into patients' hands.

References: All statements above are supported by regulatory documents and studies ([60] www.schlafenderhase.com) ([1] www.fda.gov) ([2] www.fda.gov) ([21] www.fda.gov) ([13] billeveast.com) ([44] www.fda.gov) ([9] www.freyrsolutions.com) (www.ema.europa.eu) ([5] www.agencyiq.com) ([8] www.freyrsolutions.com) ([11] pmc.ncbi.nlm.nih.gov) ([6] pmc.ncbi.nlm.nih.gov) ([41] www.linkedin.com) ([42] www.agencyiq.com) ([27] www.mastercontrol.com) ([24] pmc.ncbi.nlm.nih.gov), among others. Each citation points to a credible source detailing these findings and regulations.

External Sources

[1] https://www.fda.gov/cosmetics/cosmetics-news-events/fda-publishes-structured-product-labeling-spl-implementation-guide-validation-procedures-cosmetic#:~:the%2...

[2] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/indexing-structured-product-labeling#:~:This%...

[3] https://www.mastercontrol.com/gxp-lifeline/spl4_electronic_regulatory_submissions_0609/#:~:As%20...

[4] https://www.schlafenderhase.com/shblog/the-importance-of-structured-content-in-product-labeling#:~:This%...

[5] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:In%20...

[6] https://pmc.ncbi.nlm.nih.gov/articles/PMC2655908/#:~:The%2...

[7] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:espec...

[8] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:1...

[9] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:Tradi...

[10] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:Digit...

[11] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:Altho...

[12] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:At%20...

[13] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:If%20...

[14] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:inves...

[15] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:What%...

[16] https://www.schlafenderhase.com/shblog/the-importance-of-structured-content-in-product-labeling#:~:The%2...

[17] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:and%2...

[18] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:Imagi...

[19] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:In%20...

[20] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:The%2...

[21] https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:The%2...

[22] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:SPL%2...

[23] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:How%2...

[24] https://pmc.ncbi.nlm.nih.gov/articles/PMC2655908/#:~:JavaS...

[25] https://pmc.ncbi.nlm.nih.gov/articles/PMC2655908/#:~:All%2...

[26] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:Struc...

[27] https://www.mastercontrol.com/gxp-lifeline/spl4_electronic_regulatory_submissions_0609/#:~:Since...

[28] https://www.fda.gov/cosmetics/cosmetics-news-events/fda-publishes-structured-product-labeling-spl-implementation-guide-validation-procedures-cosmetic#:~:SPL%2...

[29] https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:Downl...

[30] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,II%2...

[31] https://www.schlafenderhase.com/shblog/the-importance-of-structured-content-in-product-labeling#:~:simpl...

[32] https://www.schlafenderhase.com/shblog/the-importance-of-structured-content-in-product-labeling#:~:For%2...

[33] https://www.schlafenderhase.com/shblog/why-data-integrity-is-important-in-pharma/#:~:Tools...

[34] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:Agenc...

[35] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:The%2...

[36] https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:,send...

[37] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:How%2...

[38] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:2...

[39] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:3...

[40] https://www.freyrsolutions.com/blog/digital-labeling-in-pharma-from-static-pdfs-to-interactive-labels#:~:4,Eng...

[41] https://www.linkedin.com/pulse/structured-content-product-label-management-aa1of/#:~:Indus...

[42] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,a%20...

[43] https://pmc.ncbi.nlm.nih.gov/articles/PMC2655908/#:~:pharm...

[44] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/spl-standard-content-labeling-technical-qs#:~:This%...

[45] https://www.fda.gov/industry/fda-data-standards-advisory-board/identification-medicinal-products-idmp#:~:IDMP%...

[46] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,surv...

[47] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,then...

[48] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,hosp...

[49] https://www.fda.gov/cosmetics/cosmetics-news-events/fda-publishes-structured-product-labeling-spl-implementation-guide-validation-procedures-cosmetic#:~:Octob...

[50] https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:,SPL%...

[51] https://www.fda.gov/industry/fda-data-standards-advisory-board/structured-product-labeling-resources#:~:,Vali...

[52] https://www.agencyiq.com/blog/ema-publishes-first-electronic-product-information-as-pilot-progresses/#:~:,sugg...

[53] https://billeveast.com/structured-product-labeling-electronic-labelling/#:~:,can%...

[54] https://pmc.ncbi.nlm.nih.gov/articles/PMC2649322/#:~:This%...

[55] https://pmc.ncbi.nlm.nih.gov/articles/PMC2655908/#:~:The%2...

[56] https://www.iqvia.com/locations/middle-east-and-africa/blogs/2023/11/digital-transformation-in-the-pharmaceutical-industry-the-transition-to-paperless-e-labeling#:~:Case%...

[57] https://www.iqvia.com/locations/middle-east-and-africa/blogs/2023/11/digital-transformation-in-the-pharmaceutical-industry-the-transition-to-paperless-e-labeling#:~:,in%2...

[58] https://www.iqvia.com/locations/middle-east-and-africa/blogs/2023/11/digital-transformation-in-the-pharmaceutical-industry-the-transition-to-paperless-e-labeling#:~:IQVIA...

[59] https://www.linkedin.com/pulse/structured-content-product-label-management-aa1of/#:~:,inde...

[60] https://www.schlafenderhase.com/shblog/the-importance-of-structured-content-in-product-labeling#:~:For%2...

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.