

AI-Assisted Global Regulatory Intelligence for Software as a Medical Device (SaMD): A Traceable, Risk-Based Framework for Multijurisdictional Compliance

Dr. Pushpinder Paul

Director, QCE & Independent Researcher- Quality and Regulatory Compliance, Quality Compliance & Excellence (QCE), H.P., India.
Email: [pushparyan27\[at\]gmail.com](mailto:pushparyan27[at]gmail.com)

Abstract: ***Background:** Software as a Medical Device (SaMD), including software functions that use artificial intelligence (AI) or machine learning (ML), is governed by evolving and partly divergent regulatory requirements. Manufacturers operating across markets must identify applicable requirements, interpret their legal status, map them to quality-system processes, and maintain evidence that requirements remain current. Manual regulatory surveillance is difficult to scale and can lead to missed changes, inconsistent interpretations, and weak traceability. **Objective:** This paper proposes a human-supervised, AI-assisted global regulatory intelligence (GRI) framework for SaMD. The framework links authoritative-source monitoring, structured requirement extraction, jurisdictional comparison, quality-management-system mapping, risk-based gap assessment, and controlled change implementation. **Methods:** A conceptual framework was developed through a targeted review of publicly available regulatory and consensus documents from the International Medical Device Regulators Forum (IMDRF), the US Food and Drug Administration (FDA), the World Health Organization (WHO), the Central Drugs Standard Control Organization (CDSCO), and the International Organization for Standardization (ISO). The approach is a narrative, document-based synthesis rather than a systematic review or empirical validation study. **Results:** The proposed architecture contains eight connected modules: source governance; regulatory change detection; document classification and version control; requirement extraction and semantic normalization; jurisdictional applicability; evidence and QMS mapping; risk prioritization; and human approval with audit-trail retention. A requirement-level data model and a staged validation plan are provided. The framework explicitly separates binding law, regulator guidance, consensus documents, and voluntary standards. **Conclusion:** AI can improve the speed, consistency, and traceability of SaMD regulatory intelligence, but it cannot replace accountable regulatory interpretation or legal review. A defensible system should use authoritative sources, preserve source-level evidence, disclose uncertainty, prevent unsupported inference, and validate performance against expert-labeled regulatory requirements. The proposed framework is a research and implementation blueprint; its effectiveness remains to be demonstrated empirically.*

Keywords: Software as a Medical Device; SaMD; artificial intelligence; regulatory intelligence; regulatory change management; medical-device software

1. Introduction

Software increasingly performs medical functions independently of dedicated hardware or as software functions integrated into broader medical devices. SaMD may support screening, diagnosis, prognosis, monitoring, treatment selection, or clinical decision-making. AI/ML can expand the analytical capability of these products, but can also introduce lifecycle risks associated with data representativeness, model drift, changes in performance, explainability, cybersecurity, and the relationship between model outputs and clinical workflows.

The regulatory challenge is not limited to determining whether software is a medical device. Manufacturers must identify the applicable legal framework, classification and conformity-assessment route, premarket evidence, quality-system expectations, cybersecurity and software-lifecycle controls, post-market surveillance obligations, and requirements for modifications. These expectations differ across jurisdictions and may change at different times. For example, IMDRF has published foundational SaMD documents and, in January 2025, final Good Machine Learning Practice (GMLP) guiding principles [1–3]. The FDA has issued guidance on AI-enabled device software functions and predetermined change control plans (PCCPs)

[4]. WHO has summarized regulatory considerations for AI for health while explicitly describing its publication as a resource rather than a binding regulatory framework [5]. CDSCO has also published a Medical Device Software guidance document in 2026, while its published material cautions readers to consult applicable legislation and current official clarifications [6].

Regulatory intelligence (RI) is the systematic collection, verification, interpretation, dissemination, and maintenance of information about laws, regulations, guidance, standards, and regulatory decisions relevant to a product or organization. Traditional RI often depends on expert monitoring, spreadsheets, email alerts, and periodic review. Such approaches remain important but may struggle with the volume of documents, terminology variation, cross-references, multiple languages, and differences between legally binding obligations and nonbinding recommendations.

AI-assisted GRI offers potential benefits through retrieval, classification, semantic comparison, change summarization, and mapping of requirements to controlled QMS processes. However, generative systems may hallucinate, misread legal modality, overlook exceptions, or cite an outdated version. In a regulated environment, an unverified AI summary

cannot be treated as the authoritative requirement. This paper therefore proposes a traceable and risk-based framework in which AI supports, but does not own, regulatory decisions.

2. Aim and Research Questions

The aim is to propose a practical framework for AI-assisted global regulatory intelligence for SaMD that supports consistent identification, interpretation, traceability, and implementation of evolving regulatory requirements.

The framework addresses five questions: (1) How can authoritative regulatory information be monitored and converted into structured requirements? (2) How can requirements be compared across jurisdictions without falsely assuming equivalence? (3) How can requirements be mapped to QMS processes, technical documentation, and lifecycle evidence? (4) How should AI-generated outputs be risk-ranked, reviewed, and controlled? (5) What validation evidence is needed before the system is relied upon in regulated operations?

3. Methods

3.1 Design

This article uses a targeted narrative review and conceptual design method. Publicly accessible primary sources were prioritized, including IMDRF technical documents, FDA guidance, WHO publications, CDSCO materials, and ISO standard descriptions. The purpose was to identify recurring regulatory-intelligence tasks and derive a workflow architecture. This was not a systematic review: no exhaustive database search, formal screening protocol, inter-reviewer agreement analysis, or quantitative synthesis was performed.

3.2 Source hierarchy

The framework classifies source material into four levels: (A) binding legislation and legally enforceable regulations; (B) official regulator rules, notices, and guidance, with the legal status of each document explicitly recorded; (C) consensus documents and international harmonization materials; and (D) voluntary standards, technical reports, and other supporting references. This hierarchy is a governance control, not a universal ranking of legal authority. Applicability must be determined in the relevant jurisdiction and context.

3.3 Scope and limitations of the review

The conceptual comparison focuses on India, the United States, the European Union, and international reference materials. It does not claim to provide a complete country-by-country legal inventory. Regulatory requirements are product-, intended-use, risk, and date-dependent. Any operational deployment would require a controlled jurisdiction list, a defined product profile, and current primary-source verification.

4. Regulatory Context for AI-Enabled SaMD

4.1 International harmonization

IMDRF's SaMD work provides common concepts for risk categorization, quality management, and clinical evaluation [1,2]. Its GMLP guiding principles emphasize multidisciplinary lifecycle management, sound software engineering and security, representative datasets, separation of training and test data, appropriate reference datasets, human-AI team performance, clinically relevant testing, transparent user information, and post-deployment monitoring [3]. These principles provide useful normalized topics for an intelligence system, but they do not replace national legislation.

4.2 United States

The FDA regulates medical-device software functions through applicable statutory and regulatory pathways, which can include 510(k) clearance, De Novo classification, or premarket approval, depending on the device and circumstances. FDA materials on AI-enabled medical devices identify relevant guidance and lifecycle resources [4]. The FDA's August 2025 final guidance on marketing-submission recommendations for PCCPs for AI-enabled device software functions describes recommendations for planned modifications, protocols, and assessment of impacts [7].

4.3 European Union

In the EU, SaMD qualification, classification, conformity assessment, clinical evaluation, post-market obligations, and quality-system requirements may be governed by Regulation (EU) 2017/745 (MDR), including the software classification principles in Annex VIII. AI-enabled medical-device software may also intersect with the EU Artificial Intelligence Act where the relevant legal conditions are met. The interaction between the MDR and the AI Act is time-sensitive and must be assessed against the current consolidated legal text, applicable transition provisions, and official Commission guidance. The proposed framework therefore stores each instrument separately and links them through explicit applicability rules rather than treating them as one merged checklist.

4.4 India

In India, the Medical Devices Rules, 2017 and applicable amendments, notifications, and CDSCO guidance are central sources for medical-device regulation. CDSCO's Medical Device Software guidance (document identifier CDSCO/MD/GD/MDSW/01/2026) provides public-awareness guidance and directs users to the statutory provisions and subsequent official clarifications for professional and legal reliance [6]. For this reason, the system should retain the source document, issue date, version, legal status, affected product category, and a reviewer-approved interpretation.

4.5 Standards and quality systems

ISO 13485:2016 establishes quality-management-system requirements for medical devices and was confirmed as current in ISO's published catalogue in 2025 [8]. ISO 14971:2019 provides a risk-management framework for medical devices, including software [9]. Standards may be voluntary, incorporated by reference, recognized by a regulator, or otherwise relevant under specific legal arrangements. The intelligence system must therefore record the exact edition, amendments, recognition status, and jurisdictional use rather than assuming that every ISO standard is universally mandatory.

5. Proposed AI-Assisted Gri Framework

5.1 Architecture and workflow

The proposed framework contains eight modules. **Module 1-** source governance- maintains an approved registry of official regulator, legislature, standards-body, and international-harmonization sources. **Module 2-** change detection- monitors new documents, revisions, withdrawals, consultations, and effective dates. **Module 3-** document control- captures title, issuing body, jurisdiction, publication date, version, status, URL, retrieval timestamp, and cryptographic hash where technically feasible. **Module 4-** requirement extraction- uses natural-language processing and retrieval-augmented generation to propose atomic requirements with verbatim source excerpts. **Module 5-** applicability assessment- maps requirements to product characteristics, intended purpose, risk class, market, lifecycle phase, and effective date. **Module 6-** QMS/evidence mapping-links approved requirements to procedures, design inputs, risk controls, software lifecycle records, clinical evidence, cybersecurity files, labeling, PMS, and change-control records. **Module 7-** risk prioritization- scores potential impact, urgency, confidence, and implementation lead time. **Module 8-** human approval and implementation- routes proposed interpretations to qualified reviewers, records decisions, updates controlled matrices, and initiates change controls where necessary.

The system should use retrieval from an approved document corpus rather than rely solely on a model's parametric knowledge. Every extracted requirement should link to an exact source location, quoted text, document version, and retrieval date. Where the system cannot identify a reliable passage or determine applicability, it should return "unresolved- expert review required," not fabricate a conclusion.

5.2 Requirement-level data model

Each requirement record should contain at minimum: unique requirement ID; jurisdiction; issuing authority; source title and identifier; source URL; document version and status; publication, effective, and transition dates; legal-status category; affected device/product profile; lifecycle phase; normalized requirement statement; verbatim excerpt; section/page/paragraph locator; modality (must/shall, should, may, or descriptive); applicability rationale; related requirements; associated QMS process; required evidence;

risk rating; AI confidence indicator; reviewer identity and approval date; decision history; and next review date.

A normalized requirement must preserve the source's legal modality. For example, 'shall' in a binding provision should not be softened to 'may'; conversely, a recommendation in guidance should not be upgraded to a binding obligation. Machine-generated summaries should always remain secondary to the original source text.

5.3 Jurisdictional comparison without false equivalence

A comparison matrix should distinguish (a) common regulatory objective, (b) jurisdiction-specific legal mechanism, (c) evidence required, (d) responsible legal entity, (e) timing, and (f) unresolved interpretation. Similar wording does not prove legal equivalence. For example, software lifecycle, risk management, clinical evaluation, cybersecurity, and post-market monitoring may appear across multiple regimes but differ in scope, evidence, conformity-assessment route, and enforceability. AI may identify candidate similarities; a qualified reviewer must decide whether requirements are equivalent, complementary, conflicting, or non-comparable.

Recommended comparison labels are: aligned objective; partially aligned; jurisdiction-specific; potentially conflicting; and not yet assessed. Each label should have a rationale and source citations. The model should not generate a single global 'compliant/noncompliant' conclusion without an approved product profile and a jurisdiction-specific applicability determination.

5.4 Risk-based prioritization

A transparent prioritization score can support triage but should not replace formal risk management. One illustrative score is $\text{Priority} = (\text{Patient/clinical impact} \times \text{Applicability} \times \text{Urgency} \times \text{Implementation effort factor})$, adjusted by confidence and uncertainty. Each input should use a documented scale, and high-severity requirements should be escalated even when the AI confidence score is low. Low confidence is a trigger for review, not evidence of low risk. The score is a proposed operational aid and has not been validated as a clinical or regulatory risk model.

5.5 Human oversight and accountability

Human review is required before an extracted statement becomes an approved requirement, before an applicability decision affects market access, and before an identified change alters a submission, technical file, QMS procedure, labeling, or post-market process. Reviewers should have suitable regulatory, software, quality, clinical, cybersecurity, and risk-management competence according to the issue. The system should record the source consulted, AI output, reviewer changes, rationale, approval, and resulting controlled-document or change-control reference.

5.6 Security, privacy, and data governance

The intelligence platform should apply role-based access, least privilege, secure integration, backup and recovery,

retention controls, supplier assessment, and protection against prompt injection or malicious content embedded in retrieved documents. Public regulatory text can still contain untrusted instructions or links; the retrieval layer must treat source content as data, not as instructions to the model. Confidential product submissions, personal data, or commercially sensitive information should not be sent to external AI services without an approved legal, privacy, and security basis.

6. Proposed Implementation Workflow

Step 1: Define the product profile, intended purpose, target population, software function, AI/ML characteristics, markets, and lifecycle stage.

Step 2: Approve the source register and jurisdiction list.

Step 3: Ingest and version-control authoritative documents.

Step 4: Detect and classify changes, including whether a document is a law, regulation, guidance, consultation, standard, or informational publication.

Step 5: Extract atomic requirements with source-level evidence and preserved legal modality.

Step 6: Apply product- and jurisdiction-specific rules to propose applicability.

Step 7: Map requirements to existing QMS processes and objective evidence.

Step 8: Identify gaps and prioritize them using documented risk criteria.

Step 9: Obtain qualified review and approve or reject the proposed interpretation.

Step 10: Initiate controlled changes, assign owners and due dates, verify implementation, and retain evidence.

Step 11: Reassess when source documents, intended use, software architecture, model, data, markets, or product risk changes.

The process should be integrated with the organization's document control, regulatory change management, design and development, risk management, software lifecycle, cybersecurity, clinical evaluation, post-market surveillance, complaint handling, and CAPA processes. Regulatory intelligence is not complete when a change is detected; it is complete when the change's applicability is decided, implementation is controlled, and effectiveness is verified.

7. Illustrative Use Case

Consider an AI-enabled SaMD product that analyzes medical images to support clinical decision-making and is intended for the United States, European Union, and India. A new official document concerning AI model modifications is detected. The system captures the source and status, compares its version with the previous approved version, extracts potentially affected passages, and identifies candidate links to software change control, performance evaluation, risk management, clinical evidence, labeling, and post-market monitoring.

The system then generates a jurisdiction-specific impact record. For the United States, it may flag whether the FDA's PCCP guidance is relevant to the proposed modification and submission strategy [7]. For the EU, it may flag MDR software obligations and, where applicable, AI Act

interactions for current legal review. For India, it may flag the applicable Medical Devices Rules and current CDSCO software guidance [6]. These are candidate routes for review, not automatic legal conclusions. The regulatory lead confirms applicability, records the rationale, and initiates controlled actions. If the new document is a draft or consultation, it is logged for monitoring but is not automatically treated as a final requirement. This use case illustrates how AI can reduce search and comparison effort while preserving regulatory accountability. It does not demonstrate a measured reduction in review time or an improvement in compliance outcomes; those claims require empirical evaluation.

8. Validation and Performance Evaluation

Before operational reliance, the system should be validated against a curated, version-controlled benchmark of regulatory documents and expert-labeled requirements. The benchmark should cover different jurisdictions, document types, legal modalities, update types, and ambiguous applicability cases. At least two qualified reviewers should independently annotate a representative subset, with disagreements resolved through documented adjudication. The benchmark should include both positive cases (requirements that should be detected) and negative cases (similar text that is not an applicable requirement).

Recommended performance measures include: (1) source retrieval precision and recall; (2) requirement extraction precision, recall, and F1; (3) citation entailment- the proportion of generated claims fully supported by cited source passages; (4) legal-modality classification accuracy; (5) version and effective-date accuracy; (6) applicability classification agreement with expert reviewers; (7) change-detection latency; (8) the rate of unsupported or fabricated claims; (9) reviewer override rate; and (10) time required for an expert to complete a defined regulatory-intelligence task. Results should be reported by jurisdiction and document type, not only as an aggregate score.

Safety-critical error analysis is essential. False negatives may omit a new obligation; false positives may cause unnecessary work or misdirect resources; incorrect modality can distort legal meaning; and stale-source errors can lead to use of superseded requirements. A human-factors evaluation should examine whether reviewers over-trust AI-generated outputs, whether source links are usable, and whether uncertainty indicators lead to appropriate escalation. A prospective pilot may compare the AI-assisted workflow with the existing manual process using the same task set. Predefined endpoints could include time per reviewed change, completeness of detected requirements, citation accuracy, and expert-rated usability. The protocol should define inclusion criteria, sample size rationale, adjudication methods, acceptance thresholds, and handling of model updates. No performance results are reported in this conceptual paper.

9. Discussion

The central contribution of the proposed framework is to treat regulatory intelligence as a controlled lifecycle process

rather than a document-search task. Its unit of control is the traceable requirement record, linked to the exact source and version, applicability rationale, QMS process, evidence, reviewer decision, and implementation status. This creates a bridge between regulatory monitoring and operational quality management.

Several design principles are particularly important. First, authoritative sourcing is more important than fluent generation: a concise answer without a verifiable source is unsuitable for compliance decisions. Second, jurisdictional differences should be represented explicitly rather than hidden by a harmonized checklist. Third, legal modality and document status must be preserved. Fourth, uncertainty should trigger review, not be concealed by a confident-sounding summary. Fifth, regulatory change management should be integrated with formal QMS controls so that a detected change results in accountable action.

The approach is consistent with the lifecycle orientation of SaMD and AI/ML guidance, including IMDRF's SaMD clinical evaluation framework [1], GMLP principles [3], and WHO's broad regulatory considerations for AI [5]. However, the framework is not itself a regulator-endorsed methodology and does not create a presumption of compliance. The usefulness of AI depends on source coverage, data quality, retrieval accuracy, model choice, reviewer competence, and the organization's control environment.

The most significant risks are hallucinated requirements, omission of exceptions, incorrect legal-status classification, source-version errors, over-generalization across jurisdictions, and automation bias. These risks can be mitigated- but not eliminated- through closed-corpus retrieval, source quotations, version control, structured schemas, confidence and uncertainty handling, human review, change logs, periodic challenge testing, and monitoring of model performance. For highly consequential determinations, the system should facilitate expert analysis rather than make the final decision.

A further challenge is the dynamic nature of regulatory texts. Effective dates, transition periods, corrigenda, amendments, national implementation, standards recognition, and regulator interpretations may differ. The platform therefore needs a maintained regulatory source register and a clear process for resolving conflicting or incomplete information. Regulatory intelligence should also include a record of when a source was last checked; the absence of a detected update should not be interpreted as proof that no update exists.

10. Limitations

This paper proposes a conceptual framework and does not report a systematic review, deployed software, expert-panel study, or empirical validation. The regulatory landscape is dynamic, and the examples are not a complete legal analysis for any product or jurisdiction. No claim is made that the framework ensures compliance, replaces regulatory counsel, or is endorsed by a regulator or standards body. The prioritization score is illustrative and requires validation before use.

11. Future Research

Future work should build a versioned benchmark of SaMD regulatory requirements; compare rule-based, retrieval-augmented, and hybrid AI approaches; measure requirement extraction and citation entailment across jurisdictions; evaluate multilingual source monitoring; test reviewer-AI interaction; and examine whether structured requirement records improve the completeness and timeliness of regulatory change implementation. Additional work should investigate ontology design for requirement relationships, conflict detection, standards recognition tracking, and secure integration with electronic QMS platforms. A multicentre study involving manufacturers, regulatory professionals, and notified-body or conformity-assessment experts could test generalizability and identify implementation barriers.

12. Conclusion

AI-assisted global regulatory intelligence can support SaMD manufacturers by organizing authoritative sources, identifying document changes, extracting candidate requirements, comparing jurisdictions, mapping requirements to QMS evidence, and prioritizing expert review. To be credible in a regulated environment, the system must preserve source-level traceability, distinguish law from guidance and standards, retain version and effective-date metadata, disclose uncertainty, and require qualified human approval before regulatory decisions are implemented. The framework presented here offers a structured blueprint for development and empirical testing. Its practical value should be established through transparent validation against expert-reviewed regulatory benchmarks rather than assumed from AI capability alone.

References

- [1] International Medical Device Regulators Forum (IMDRF). Software as a Medical Device (SaMD): Clinical Evaluation. IMDRF/SaMD WG/N41 FINAL:2017. 21 September 2017. Available at: <https://www.imdrf.org/documents/software-medical-device-samd-clinical-evaluation> (accessed 20 October 2026).
- [2] International Medical Device Regulators Forum (IMDRF). Software as a Medical Device: Application of Quality Management System. IMDRF/SaMD WG/N23 FINAL:2015. Available via the IMDRF SaMD working-group document library: <https://www.imdrf.org/working-groups/closed-working-groups/software-medical-device-samd> (accessed 20 October 2026).
- [3] International Medical Device Regulators Forum (IMDRF). Good Machine Learning Practice for Medical Device Development: Guiding Principles. IMDRF/AIML WG/N88 FINAL:2025. 29 January 2025. Available at: <https://www.imdrf.org/documents/good-machine-learning-practice-medical-device-development-guiding-principles> (accessed 20 October 2026).
- [4] U.S. Food and Drug Administration (FDA). Artificial Intelligence-Enabled Medical Devices. FDA Digital Health Center of Excellence. Available at:

- <https://www.fda.gov/medical-devices/digital-health-center-excellence/artificial-intelligence-enabled-medical-devices> (accessed 20 October 2026).
- [5] World Health Organization. Regulatory Considerations on Artificial Intelligence for Health. Geneva: WHO; 2023. ISBN 978-92-4-007887-1. Available at: <https://www.who.int/publications/b/64825> (accessed 20 October 2026).
- [6] Central Drugs Standard Control Organization (CDSCO), Medical Devices & IVD Division. Guidance Document on Medical Device Software. Document No. CDSCO/MD/GD/MDSW/01/2026. Available at: https://www.cdsc.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/Guidance-document-on-Medical-Device-Software-under-MDR-2017.pdf (accessed 20 October 2026).
- [7] U.S. Food and Drug Administration (FDA). Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and Food and Drug Administration Staff. August 2025. Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence> (accessed 20 October 2026).
- [8] International Organization for Standardization (ISO). ISO 13485:2016 Medical devices—Quality management systems—Requirements for regulatory purposes. Geneva: ISO; 2016. Available at: <https://www.iso.org/standard/59752.html> (accessed 20 October 2026).
- [9] International Organization for Standardization (ISO). ISO 14971:2019 Medical devices—Application of risk management to medical devices. Geneva: ISO; 2019. Available at: <https://www.iso.org/standard/72704.html> (accessed 20 October 2026).
- [10] European Parliament and Council of the European Union. Regulation (EU) 2017/745 of 5 April 2017 on medical devices. Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/eli/reg/2017/745/oj> (accessed 20 October 2026).
- [11] European Parliament and Council of the European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> (accessed 20 October 2026).
- [12] Government of India. Medical Devices Rules, 2017, as amended. Ministry of Health and Family Welfare. Official current consolidated text and notifications should be verified through CDSCO before use: <https://cdsc.gov.in/> (accessed 20 October 2026).

Appendix A: Framework Tables

Table 1: Core modules of the proposed framework

Module	Primary function	Key control/output
1. Source governance	Maintain approved source register	Authority, jurisdiction, source type, owner, review frequency
2. Change detection	Identify new, amended, withdrawn, or draft documents	Change event with date and source link
3. Document control	Version and archive source records	Version, status, retrieval date, hash/record ID
4. Requirement extraction	Create atomic candidate requirements	Verbatim excerpt, locator, normalized text, modality
5. Applicability	Relate requirements to product and market	Applicability rationale and reviewer decision
6. QMS/evidence mapping	Link requirement to controlled processes	Procedure, design file, risk file, clinical/PMS evidence
7. Risk prioritization	Triage impact and urgency	Transparent score, uncertainty flag, escalation
8. Human approval	Approve interpretation and implementation	Reviewer, rationale, change-control link, audit trail

Table 2: Minimum requirement record fields

Field group	Minimum fields
Identity	Requirement ID; jurisdiction; issuing authority; source identifier
Source control	Title; version; status; URL; publication/effective date; retrieval date; exact locator
Interpretation	Verbatim excerpt; normalized statement; legal modality; source type; applicability rationale
Product context	Intended purpose; device/software function; risk class; market; lifecycle phase
Implementation	QMS process; required evidence; owner; due date; linked controlled records
Governance	AI model/version; confidence/uncertainty; reviewer; approval; decision history; next review

Table 3: Proposed validation metrics

Metric	What it assesses	Suggested evaluation method
Requirement extraction precision/recall/F1	Completeness and correctness of extracted requirements	Compare against expert-labelled benchmark
Citation entailment	Whether cited passages support generated claims	Blind expert review of claim–evidence pairs
Modality accuracy	Correct preservation of must/shall/should/may	Compare against labelled source clauses
Version/effective-date accuracy	Use of the current document and correct timing	Test amended, superseded, draft, and transition cases
Applicability agreement	Alignment with expert product/jurisdiction decisions	Inter-reviewer agreement and adjudication
Change-detection latency	Time from official publication to system alert	Timestamped test set
Unsupported-claim rate	Frequency of hallucinated or ungrounded statements	Adversarial and routine challenge set
Task completion time	Efficiency relative to the manual workflow	Controlled comparative pilot