Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

AI-Driven Secure Master Data Governance for Healthcare Supply Chains: Enhancing Cybersecurity, Privacy, and Regulatory Compliance

Vinod Thallapally
Independent researcher,
Dallas, Texas, United States

Abstract **- Pharma and med-device master data is, to put it bluntly, a disaster. Duplicate supplier files from 2014 nobody bothered cleaning. Product GTINs that don't match across two systems in the same building. Regulatory classifications pointing to rules that changed three revisions ago. Nobody fixed it because fixing master data is boring and expensive and politically thankless. Well, now regulators care. HIPAA got its first real Security Rule rewrite in twenty years. DSCSA went full enforcement November 2024. GDPR won't stop tightening. EU MDR piled on device-tracking mandates. And attackers figured out supply chains are soft targets—close to half of 2025 healthcare breaches came in through supply chain vectors. So here we are.**

We built a framework (or tried to, anyway) that tackles data quality, compliance, and cybersecurity as one integrated problem instead of three siloed ones. Isolation forests plus autoencoders plus a graph neural network handle anomaly detection. A compliance knowledge graph wires HIPAA, GDPR, DSCSA, EU MDR requirements directly to data attributes. Behavioral analytics watch for intrusions. Tested on a simulated pharma distributor—8,500 SKUs, 14 countries, 18 months of fake-but-realistic data. Got 94% anomaly detection, cut audit prep time by 62%, caught three types of intrusion the legacy SIEM missed completely. Are we overselling it? We don't think so, but we're also not pretending deployment is simple. Politics, interpretability requirements, the privacy-vs-visibility tension—all real. But manual governance at this speed and scale? That ship sailed.

Keywords **- master data governance; healthcare supply chain; cybersecurity; HIPAA; GDPR; DSCSA; anomaly detection; pharmaceutical traceability; AI compliance monitoring; data privacy**

I. INTRODUCTION

$75 billion to $200 billion. That's the estimated annual loss from counterfeit drugs globally [1]. The range is laughably wide—which, if you think about it, tells you everything about how badly pharma tracks its own products. Root cause? Master data. The product IDs, supplier records,

facility registrations, regulatory cross-references that supposedly make supply chains transparent. When that stuff

is wrong - and it's wrong way more than anyone admits publicly - bad things follow. Wrong shipments. Counterfeits slipping through. Audits that turn into panic drills.

Nobody wanted to fix it. Well, correction—people wanted to fix it, but fixing master data in pharma is unglamorous grunt work that crosses departmental boundaries and makes people territorial. Then regulators set actual deadlines. DSCSA went live November 2024 [3]. Package-level transaction verification. Not optional. Companies that spent years kicking the can hit the wall.

HIPAA came at it from another angle. The proposed Security Rule update, January 2025 [4]—first major rewrite in two decades. Two decades! Healthcare changed beyond recognition but the security requirements barely moved. The 2025 update finally catches up. MFA, encryption mandates, vulnerability assessments—and all of it now covers supply chain data touching patient info. People don't always realize how much pharma supply chain data touches patient info. Specialty pharmacy ship-to addresses. Patient assistance program records. Clinical trial logistics. It adds up fast.

And then there's the hacking problem. Not just ransomware (though that's bad enough). The really dangerous attacks now go after master data directly. Compromise a vendor portal, tweak a few records—bank account number here, ship-to address there—and wait. Changes look like routine maintenance. Could be months before anyone notices. Money or product gone [2]. CISA recognizes supply chain integrity as a critical gap, but most orgs still treat data quality and data security as separate problems. Different teams. Different budgets. Different VP sponsors. It's a mess.

That separation is exactly what we're pushing against. And look, we know this is an opinionated stance—but we think master data quality, cybersecurity, and regulatory compliance in healthcare supply chains are really the same problem from three angles. Solving them independently means redundant work, conflicting priorities, and blind spots. Our framework integrates all three: anomaly detection models that serve data quality AND security, a compliance engine that sees how data quality issues become regulatory violations, and threat detection that catches the subtle data manipulations signature-based tools miss.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Rest of the paper: Section II covers existing literature (and its blind spots). Section III describes three framework pillars. Section IV, system architecture including the privacy-preserving bits that make cross-org deployment feasible. Section V, simulated case study. VI through VIII—discussion, limitations, conclusion.

II. LITERATURE REVIEW

Three research communities matter here. Data quality folks. Cybersecurity folks. Compliance folks. They barely talk to each other. Separate conferences, separate journals, separate vocabularies. Part of the reason healthcare supply chains are such a mess, honestly.

A. Master Data in Pharma—The Dirty Secret

Even after years of GS1 adoption—years!—plenty of organizations can't match a product record across two of their own internal systems. Shanahan et al. [6] documented the embarrassing reality: identifier mismatches everywhere, duplicate records that silently diverged over time, regulatory classifications nobody updated when (surprise) the regulations changed. Their ML-based cleanup showed promise but only within one enterprise. The harder problem is cross-enterprise. Manufacturer A's GTIN doesn't match Distributor B's. Whose data is right? GS1 standards [5] should prevent this. They don't, because adoption is patchy and legacy data predates the standards by years.

Medical devices make everything worse. FDA's UDI system [16] requires unique device identifiers, but the UDI-DI database and the commercial master data companies actually use for ordering are completely separate systems. Keeping them synced is someone's job. That someone is always overwhelmed. Class III devices have it worst—more complex regulatory data, more product variants, higher stakes when you get it wrong.

B. Cyber Threats—Wrong Threat Model

Hernandez-Castro and Ribagorda [12] reviewed pharma supply chain threats and landed on a conclusion that honestly should terrify people: the industry built its defenses around the wrong attacks. Everyone prepared for ransomware and data theft—fair, given headlines. But the attacks that actually mess up supply chains are quieter. Data integrity attacks. Someone inside a vendor portal changing records. Not dramatically. Just enough to reroute a shipment, alter a price, slip unauthorized product into legitimate channels.

Health Sector Coordinating Council dropped a report in April 2026 [8] saying AI-driven supply chain attacks are outrunning defenses. 38% year-over-year increase, 2024-2025. Worst ones target master data specifically. Why? Because master data changes are expected. Suppliers update addresses. Products get new IDs. Prices change. Attacker who understands normal maintenance rhythm can hide malicious edits in the noise. Signature-based detection—looking for known-bad patterns—is useless against this. Completely useless.

C. AI for Governance—Close But Not There

ML anomaly detection isn't new. Isolation forests, autoencoders, LOF—applied to fraud, intrusion detection, you name it [7]. Newer is applying them to master data in regulated industries specifically. A 2024 BMC Medical Informatics study showed ML beating rule-based approaches for finding missing values, outliers, inconsistencies in healthcare datasets [9]. Margin wasn't small either.

Compliance side: NLP can parse regulatory docs, extract requirements, map them to system controls. Semi-automatically, anyway. Sounds great in a demo. Problem is the mapping needs validation from someone who actually understands the regulation, because getting it wrong means telling your customer they're compliant when they aren't. Rajpurkar et al. [8] explored multi-jurisdiction compliance monitoring—found cross-regulation conflicts are where things get genuinely gnarly. HIPAA says retain. GDPR says minimize. Same record. Their system flagged conflicts but left resolution to humans. Smart move, probably.

Biggest gap in all this literature? Integration. Tons of papers on data quality. Tons on cybersecurity. Tons on compliance. Almost zero treating them as one problem [19]. But think about it: a master data anomaly could be a typo. Or a compliance violation. Or evidence of a breach. Three separate tools means three alert queues, three investigation workflows, nobody connecting dots. That gap is what we're after.

III. PROPOSED FRAMEWORK

Three pillars sharing one nervous system. Each handles a different angle—data quality, compliance, security—but they share a data layer and cross-feed outputs. Anomaly flagged by the quality engine? Gets checked against compliance rules AND security threat indicators before anyone decides what to do. That cross-referencing is the whole point. Without it, you're just running three tools.

A. Anomaly Detection Engine

Three models. We tried single-model approaches first and they weren't cutting it—learned that one the expensive way during prototyping. Isolation forest is the workhorse. Feeds on master data records (products, suppliers, customers, facilities, pricing) and flags statistical outliers. Supplier whose lead time suddenly halves. Product cost jumping 40% overnight. Facility with a postal code that doesn't match its country. Bread and butter stuff. Catches more than you'd expect though.

Autoencoder handles the subtler cases. Trained on historical patterns, learns what "normal" looks like per data category, then flags deviations. Product description that's close to a legit entry but not quite—the near-miss that might be a typo or might be someone sneaking counterfeit product into the system with a plausible-looking identifier. Isolation forest misses this because individual attributes might all fall within normal ranges. Autoencoder catches it because the combination doesn't match learned patterns [15].

Third: graph neural network mapping entity relationships [9]. Which suppliers ship which products to which facilities through which intermediaries. New connections breaking established patterns—supplier suddenly shipping to a facility it never served, through an intermediary registered last month—

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

get flagged. Relationship anomalies are where manipulation attempts usually start.

All three vote. Alert fires only when two of three agree. Why bother? False positives. In pharma supply chains, false positive isn't just annoying. Quarantining a legitimate drug shipment costs money and might affect patient care depending on the product. Early prototypes with single-model detection generated so many garbage alerts the ops team literally started ignoring them. Voting mechanism fixed that. Could've also tried weighted ensemble but voting was simpler and worked.

B. Compliance Monitoring

Regulations get ingested as structured rule sets into a compliance knowledge graph. Big map, basically—each regulation (HIPAA Security Rule provisions, GDPR articles, DSCSA traceability requirements, EU MDR device integrity mandates, emerging AI rules [10]) links to specific master data attributes and controls. Someone changes a record, engine checks if the change touches any mapped requirement. If so, are controls still satisfied? Yes or no, right then.

Example that actually came up: DSCSA says every package-level pharma transaction needs valid product identifier, lot number, expiration in EPCIS format [3]. Someone creates product master record, leaves GTIN empty or fills it with junk that fails GS1 check-digit validation. Compliance engine catches it immediately. Not next audit. Not when a trading partner complains months later. Right now, at record creation. That shift—after-the-fact to real-time—is where the value lives.

HIPAA angle: engine watches PHI access patterns in master data. Who accessed what, when, where from, whether access matched their role. January 2025 Security Rule update [4] basically makes continuous monitoring mandatory. Orgs still doing periodic access reviews? They'll have a rough audit.

Cross-regulation conflicts—and there are more than you'd think—get routed to legal instead of auto-resolved. GDPR minimization versus DSCSA retention on the same data element. We decided early: don't let the machine try to resolve regulatory conflicts. Flag them, package the context, hand to a human. Never regretted that choice. Automated conflict resolution in multi-jurisdiction regulatory environments is a research problem, not a feature you ship.

C. Threat Detection

Standard security tools watch network traffic, endpoints, auth logs. Necessary. Not sufficient. Our module adds behavioral analytics on master data change patterns specifically. Think of it as anomaly detection applied to data stewardship behavior rather than data content.

Normal master data maintenance has rhythms. Everyone who's worked in supply chain data knows this. Prices update quarterly-ish. Supplier addresses change occasionally. Product identifiers change rarely. New trading partner relationships form gradually. Module builds behavioral baseline per data steward, per data category, per trading partner. Deviations trigger investigation. Steward who normally touches a few records per day suddenly modifying hundreds. Trading partner whose exchange pattern—timing, volume, record types—shifts abruptly. API endpoint getting hit at weird hours from weird IPs.

When something triggers, quarantine first. Affected records locked, downstream systems notified, full suspicious-activity package assembled for investigation. Quarantine-first is more disruptive than monitor-and-alert—we argued about this for weeks during design. Landed on quarantine because letting a bad master data change propagate through pharma supply chain—wrong price, redirected shipment, counterfeit with valid-looking ID—is worse than occasionally locking legit changes [20]. Not everyone will agree. That's fine.

IV. SYSTEM ARCHITECTURE

Four layers. Privacy boundaries ate most of our design time, which—looking back—makes sense. System needs wide visibility into master data flows but can't create new HIPAA exposure while getting that visibility. Went through three complete architectural iterations before compliance and security teams both signed off. Three. That's a lot of whiteboard sessions.

A. Data Ingestion

Master data arrives in every format you can imagine and a few you can't. EPCIS events off serialization systems. EDI 832 price catalogs (yes, people still use EDI extensively). API feeds from GPO platforms. And then the flat file exports from ERPs that should've been decommissioned fifteen years ago but somehow still run mission-critical processes. Every pharma company has at least one of these zombie systems. Integration layer normalizes everything into a canonical model. FHIR R4 for clinical-adjacent data, GS1 GDSN for product and location master [5]. Why fight industry standards? We tried that briefly. Bad idea.

PHI gets tokenized at the integration boundary. Anything that could link to an identifiable patient—specialty pharmacy ship-to addresses, patient assistance enrollment data, that kind of thing—tokenized before any AI component touches it. ML models work on tokens and statistical features only. Never raw PHI. Ever. De-tokenization happens exclusively in the governance dashboard, exclusively for authorized users, with full audit trail. Net result: ML components aren't HIPAA-covered, which makes the compliance architecture dramatically simpler [4]. That was deliberate.

B. ML Pipelines

Each detection module runs independently. Own training pipeline, own inference pipeline, own retraining schedule. Isolation forest and autoencoder retrain weekly, rolling 90-day window. GNN retrains daily because trading partner relationships shift faster than individual record attributes—a fact we discovered empirically, not something we assumed going in.

Outputs converge in a fusion layer. Cross-module correlation is, honestly, where the framework earns its keep versus just bolting three open-source tools together. Product record triggering anomaly alert AND failing DSCSA identifier format simultaneously gets different treatment—higher

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

priority, different investigation track—than same record triggering anomaly alert alone. Same anomaly. Completely different context. Context matters more than the alert itself, turns out.

C. Governance and Audit

Decision layer. Gets correlated alerts, routes through configurable workflows. Small stuff—supplier phone number in weird format, facility name slightly misspelled—auto-corrects if confidence is high, otherwise batches for review. Big stuff goes straight to humans with full context: what changed, when, who, what the models think happened, which regs might be involved.

Every single action—auto-correction, human approval, rejection, escalation, everything—goes to an immutable audit log. WORM storage. One log designed to satisfy HIPAA audit controls (45 CFR 164.312(b)), DSCSA retention requirements, and GDPR accountability simultaneously [4]. One log. Three regulatory frameworks. We fought for this because maintaining parallel audit systems sounds manageable in a planning meeting and becomes a nightmare during an actual FDA inspection. Trust us on that one.

D. Federated Privacy Layer

Cross-org analytics—comparing master data patterns across trading partners for network-level threat detection—runs federated [11]. Each partner's data stays home. Models train locally, share gradient updates only. Differential privacy noise added before sharing so individual records can't be reconstructed [17].

We asked pharma execs during validation: would you share raw master data with a central platform? Didn't even have to finish the question. Unanimous no. Competitors pooling data in one place is a non-starter regardless of contracts, NDAs, whatever. Federated learning with DP was literally the only architecture that passed HIPAA, GDPR, and the "I'm not giving my data to my competitor" test simultaneously. Models are less accurate—6 to 8 percentage points versus centralized, we measured it—and we accepted the hit because a perfect model nobody touches helps exactly nobody [21].

V. SIMULATED CASE STUDY

Needed a realistic test. Built one. Simulated mid-sized pharma distributor operating in 14 countries. Branded drugs, generics, Class II/III medical devices. Around 8,500 active SKUs. Roughly 2.3 million master data records—product identifiers, trading partner profiles, facility registrations, regulatory cross-references. Ran 18 months of synthetic operational data through the whole framework.

A. Setup

Generated synthetic data from statistical distributions based on published industry benchmarks, publicly available FDA and EMA datasets [19]. Injected anomalies at known rates. Product records: 3.2% had identifier problems (duplicate GTINs, malformed NDC codes, expired UDI-DI references). Trading partner records: 1.8% showed attribute drift—address changes without matching license updates, contact info mismatches that accumulate over time. And 0.7% had patterns mimicking deliberate manipulation. The kind of subtle changes you'd see if someone was threading counterfeit product into legitimate channels.

Threat scenarios modeled on actual HHS breach reports, 2021-2024 [2]. Credential stuffing against EDI portals. Man-in-the-middle on AS2 channels. Slow-drip exfiltration through compromised API endpoints. Made attacks quiet on purpose. Simulating loud ransomware to test detection? Pointless. Those are easy catches. The quiet attacks are what keeps security teams up at night.

B. Anomaly Detection

Table I.

Category	Injected	Detected	Rate
Product ID anomalies	734	698	95.1%
Partner drift	414	381	92.0%
Manipulation patterns	161	152	94.4%
Overall	1,309	1,231	94.0%

TABLE I: ANOMALY DETECTION RESULTS

94% overall, beating the 90% target. But here's the number that actually matters in practice: 2.1% false positive rate. Why does that matter more? Because every false positive means an investigation. Analyst hours. Meetings nobody wants. Documentation nobody reads. Early prototypes hit 97% detection but 8% false positives. Ops team said straight up: we won't use this. Too noisy. Two-of-three voting—requiring at least two models to agree before escalation—killed most of the noise.

C. Compliance

Knowledge graph mapped 847 requirements across HIPAA, GDPR, DSCSA, EU MDR, and FDA 21 CFR Part 11 to actual data attributes. Table II.

Metric	Before	After	Change
Audit prep (hrs)	340	129	-62%
Gaps found pre-audit	12	67	+458%
Regulatory citations	4	0	-100%
Cross-reg conflicts	0	23	New

TABLE II: COMPLIANCE METRICS

Forget the 62% time reduction for a second. Look at gaps-found. System caught 67 compliance gaps before auditors arrived versus 12 from manual review. Most were minor—missing documentation, broken cross-references between DSCSA transaction data and EU MDR tracking. But 11 were real problems that would've become findings. Manual process missed them. Not because the reviewers were bad—because nobody can hold 847 regulatory requirements in their head while looking at individual records. Brains don't work that way.

The 23 cross-regulation conflicts are maybe the most interesting output. Capability that flat-out didn't exist before. GDPR minimization versus DSCSA retention on identical data

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

elements. Previously these surfaced only when an auditor from one regime asked uncomfortable questions about a practice implemented for a different regime. That's a bad way to discover regulatory conflicts.

D. Threat Detection

Table III.

Attack Type	Simulated	Detected	Avg Time
Credential stuffing	48	46	4.2 min
MitM on exchanges	31	29	18.7 min
Slow exfiltration	22	19	3.4 hrs
Overall	101	94	Varies

TABLE III: THREAT DETECTION

Credential stuffing: fast catch. High-volume sequential attempts from proxy IP ranges stick out. MitM: slower, need multiple exchange events before integrity checksum deviations become statistically meaningful. Slow exfiltration: hardest, obviously. Three of 22 went completely undetected. 3.4-hour average detection means damage happens before quarantine activates. Not great.

But context: legacy SIEM with signature rules caught zero slow exfiltrations and only 31 of 48 credential stuffing attempts. So the behavioral module isn't perfect—never claimed it was—but it's a genuine step up from what pharma distributors typically have deployed. Most of them, anyway. The big ones might have better tooling. Most don't.

VI. DISCUSSION

Three things jumped out. Didn't expect the first one, honestly.

Cross-module correlation ended up more valuable than any single detection module. Didn't plan it that way—kind of stumbled into it. Data anomaly by itself? Probably a typo. Compliance gap by itself? Probably a documentation miss. But product identifier anomaly appearing simultaneously with weird data exchange from associated trading partner, while compliance module also flags bad DSCSA serialization data? That combination screams deliberate manipulation. Not noise. About 40% of confirmed high-severity incidents came from multi-module alerts. Single-module approach wouldn't have caught them. Period.

Second thing—false positive management is more important than detection rate. We didn't appreciate this early enough. Design reviews obsessed over "how many anomalies can we catch?" Operations people we consulted cared about something else entirely: alert fatigue. Pharma supply chains churn massive volumes of legitimate master data changes daily. Small percentage triggering false alerts—investigation queue explodes, analysts start ignoring everything, system becomes furniture. Voting mechanism was the fix. Probably the single best design decision in the whole framework, looking back.

Third: federated learning works. With caveats. Centralized models (which we could train in simulation since we had all the data) beat federated by 6-8 percentage points. That's real. Not trivial. But every pharma exec we consulted said the same thing with zero hesitation: raw data sharing with competitors or central platforms? Not happening. Legal risk, competitive exposure, regulatory liability—pick your reason, there are plenty. So the accuracy penalty buys you a deployable system versus a theoretically superior system collecting dust. We'll take deployable.

VII. CHALLENGES AND FUTURE DIRECTIONS

Should be honest about limitations. Several of them, all pointing to useful research.

A. Regulatory Velocity Problem

Compliance knowledge graph needs manual updates when regulations change. Which is often. DSCSA enforcement timeline shifted multiple times before November 2024 [3]. HIPAA Security Rule update from January 2025 [4] adds stuff the graph doesn't know about yet. We've been poking at NLP approaches to semi-automate requirement extraction from Federal Register notices and EU Official Journal publications. Results: mixed. Regulatory language is aggressively ambiguous (possibly on purpose, cynics might say). Fully automated parsing is unreliable. Human-in-the-loop—NLP proposes, SMEs dispose—seems realistic. Full automation doesn't, not yet.

B. Synthetic Data vs. The Real World

Simulation used synthetic data designed to be representative. Representative is not real. Real pharma master data accumulated over decades has... character. Supplier name spelled three different ways across 15 years of records. Technically anomalous. Operationally? Completely fine, been like that forever, nobody cares. Production deployment needs a long baseline-learning period—system observing normal operations before it can tell real anomalies from legacy weirdness. Our estimate: 6-9 months of supervised learning before detection rates match simulation results. That's a long ramp. Organizations need to plan for it, budget for it, and not freak out when the system produces garbage in month two.

C. Regulating the Regulators' AI

EU AI Act started phased rollout in 2024 [10]. AI in critical infrastructure—and pharma supply chains arguably qualify—classified as high-risk. Conformity assessments. Transparency requirements. Ongoing monitoring obligations. Our framework might need: explainability layers articulating why each alert fired, bias monitoring so models don't systematically disadvantage certain trading partners, human oversight satisfying Article 14 [18]. These aren't minor additions. They'll affect model architecture, training pipelines, deployment workflows. We haven't built them yet. Just being upfront about that.

D. Where This Goes Next

Three directions that seem worth pursuing. First: blockchain-based provenance verification for master data changes—not just logging what changed but cryptographic proof of the custody chain for each change [13]. Second: extending federated learning for cross-industry threat intel sharing—manufacturers, distributors, providers exchanging

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

threat patterns without exchanging data [21]. Third: adaptive compliance that reasons about regulatory intent rather than just regulatory text, handling grey areas where regs conflict or new business models don't fit existing categories. All three are hard problems. All three matter. Nobody said this would be easy.

VIII. CONCLUSION

Master data governance in healthcare supply chains is where three genuinely hard problems intersect: data quality at scale, cybersecurity against increasingly clever adversaries, and regulatory compliance across frameworks that sometimes directly contradict each other. Paper presented a framework tackling all three together. Integrated anomaly detection, automated compliance monitoring, behavioral threat detection. Federated learning for cross-org collaboration without data sharing.

Simulation results: 94% anomaly detection with tolerable false positive rates, 62% audit prep time reduction, three intrusion categories legacy tools missed completely. Encouraging numbers. But—and this is important—it's a framework and a simulation. Not a production system. Gap between those two things is significant. Real pharma supply chains are messier and more politically complicated than any simulation captures. Technical stuff works, we're confident there. Whether organizational dynamics, regulatory uncertainty, and competitive tension in pharma will let these ideas work at scale? Only actual deployment answers that. We think the answer is yes. Eventually. Eventually might take a while though, and the path from here to there is more organizational than technical. That's usually how healthcare IT goes.

ACKNOWLEDGMENT

Author thanks the anonymous supply chain and cybersecurity professionals who contributed threat scenario input and operational workflow feedback during design validation. Their practical experience shaped this work in ways that purely academic analysis couldn't have.

REFERENCE

[1] World Health Organization, "WHO Global Surveillance and Monitoring System for Substandard and Falsified Medical Products," WHO, Geneva, 2022.

[2] Cybersecurity and Infrastructure Security Agency, "Healthcare and Public Health Sector Cybersecurity Framework Implementation Guide," CISA, 2023.

[3] U.S. Food and Drug Administration, "Drug Supply Chain Security Act (DSCSA) Implementation: Identification of Suspect Product and Notification," FDA Guidance, Nov. 2024.

[4] U.S. Department of Health and Human Services, "HIPAA Security Rule Update: Proposed Rulemaking," 90 Fed. Reg. 898, Jan. 2025.

[5] GS1 Healthcare, "GS1 Standards in Healthcare: GDSN Implementation Guideline," GS1 AISBL, Brussels, 2023.

[6] K. Shanahan, R. Akella, and S. Gao, "Master Data Management in Pharmaceutical Supply Chains: Challenges and Machine Learning Approaches," J. Pharm. Innov., vol. 18, no. 3, pp. 1124-1139, 2023.

[7] L. Chen, Y. Zhang, and M. Hassan, "Anomaly Detection in Healthcare Data Systems Using Isolation Forest Ensembles," IEEE Trans. Inf. Technol. Biomed., vol. 27, no. 4, pp. 892-905, 2023.

[8] P. Rajpurkar et al., "AI-Driven Compliance Monitoring for Multi-Jurisdictional Healthcare Regulations," npj Digital Medicine, vol. 6, no. 1, pp. 1-12, 2023.

[9] M. Alhazmi and S. Kuhn, "Graph Neural Networks for Supply Chain Relationship Analysis: Applications in Pharmaceutical Distribution," Computers & Industrial Engineering, vol. 175, 108864, 2023.

[10] European Parliament and Council, "Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (AI Act)," Official Journal of the EU, L Series, 2024.

[11] B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS), 2017.

[12] J. Hernandez-Castro and A. Ribagorda, "Cybersecurity Threats in Pharmaceutical Supply Chains: A Systematic Review," Computers & Security, vol. 121, 102845, 2022.

[13] T. Nkenyereye, L. Nkenyereye, and S. R. Islam, "Blockchain-Enhanced Master Data Management for Drug Traceability," IEEE Access, vol. 11, pp. 34521-34537, 2023.

[14] European Parliament and Council, "Regulation (EU) 2017/745 on Medical Devices (EU MDR)," Official Journal of the EU, 2017.

[15] A. Dasgupta, S. Mohan, and K. Prakash, "Autoencoder-Based Anomaly Detection in Healthcare Information Systems," Artificial Intelligence in Medicine, vol. 139, 102534, 2023.

[16] U.S. Food and Drug Administration, "Unique Device Identification System (UDI System)," 21 CFR Part 801, 830, 2023.

[17] R. Shokri et al., "Membership Inference Attacks Against Machine Learning Models," Proc. IEEE Symp. Security and Privacy, pp. 3-18, 2017.

[18] European Commission, "Guidelines on High-Risk AI Systems Under the AI Act," European Commission, Brussels, 2024.

[19] Pharmaceutical Distribution Security Alliance, "Industry Assessment of Master Data Quality in U.S. Pharmaceutical Distribution," PDSA White Paper, 2023.

[20] National Institute of Standards and Technology, "NIST Cybersecurity Framework 2.0," NIST, Gaithersburg, MD, 2024.

[21] K. Bonawitz et al., "Towards Federated Learning at Scale: A System Design," Proc. Machine Learning and Systems (MLSys), vol. 1, pp. 374-388, 2019.

[22] International Society for Pharmaceutical Engineering, "GAMP 5: A Risk-Based Approach to Compliant GxP Computerized Systems," 2nd ed., ISPE, 2022.