

ML LABS INTELLIGENCE

Vendor-Neutral ECG Ingestion Across Device Manufacturers

Every ECG manufacturer encodes data differently. ML LABS built a single ingestion pipeline that parses three major vendors' formats into one clinically precise model.

CASE STUDY

Author Omar Trejo

Date 2026-03-04

ML LABS

mlabs.com/intel/multi-vendor-ecg-ingestion-pipeline

A healthcare AI platform needed to ingest ECG recordings from every major device manufacturer and process them through a unified analysis pipeline. The problem was not connectivity or storage — it was that each of the three major device manufacturers encodes DICOM data differently enough that a generic parser produces clinically incorrect measurements. Trailing zeros in one vendor's amplitude values, renamed classification labels in another's firmware updates, and bitmap-encoded filter settings in a third's proprietary extensions meant that treating all DICOM files identically would silently corrupt the data clinicians rely on.

ML LABS built a vendor-neutral ingestion pipeline with per-manufacturer parsing adapters, measurement normalization that preserves clinical precision, deduplication at the storage layer, and a shared ECG library that provides unified access across all formats — including a research SDK for programmatic data retrieval.

The Problem

The platform had committed to vendor neutrality — accepting ECGs from any supported device without requiring clinics to standardize on a single manufacturer. In theory, the **DICOM standard** should make this straightforward: every compliant device produces structured data that any compliant system can read. In practice, each manufacturer's DICOM implementation diverges enough to break that assumption at the measurement level, the waveform level, and the metadata level.

Three categories of divergence made generic parsing unacceptable:

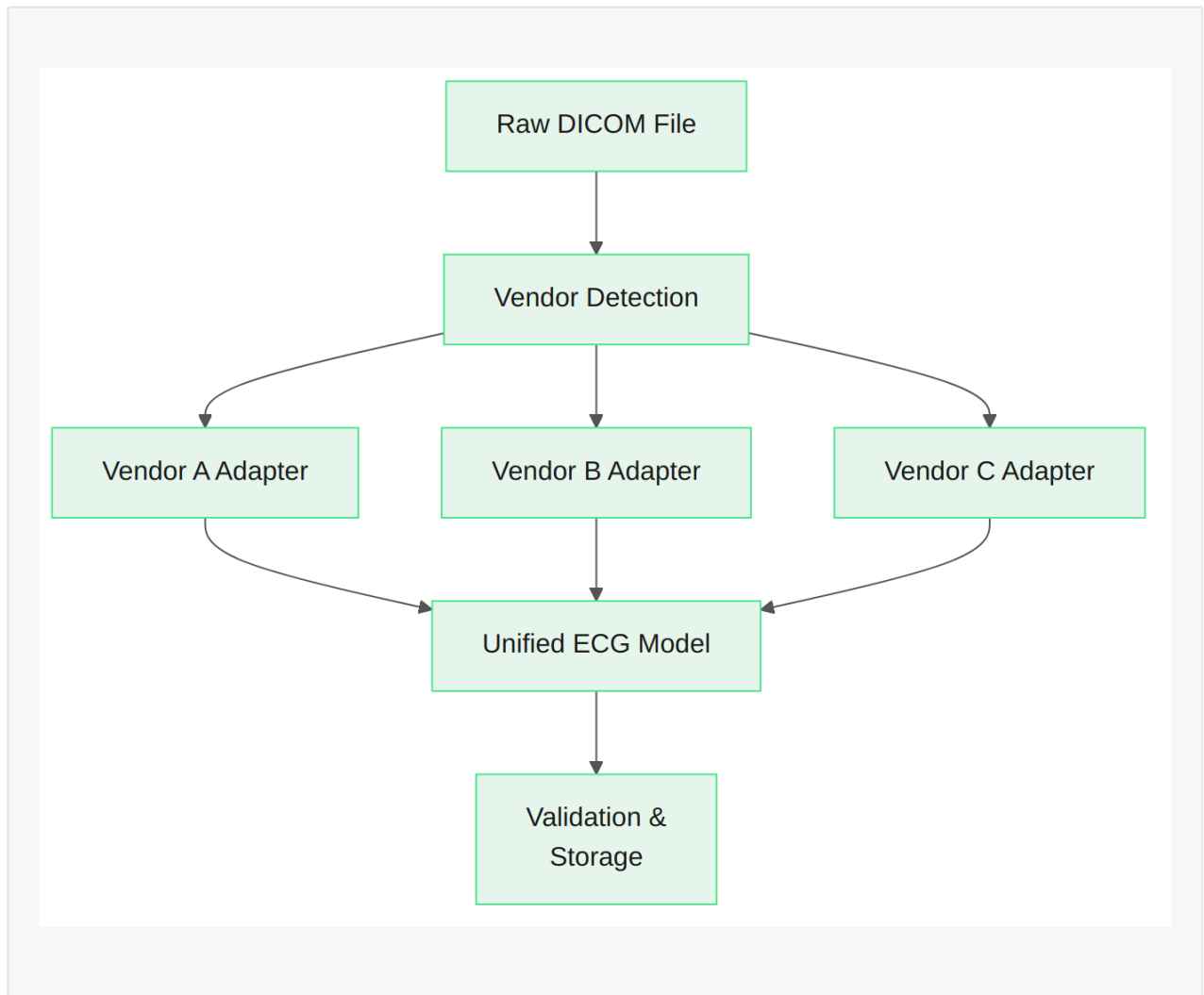
- **Measurement encoding.** One manufacturer stores QTc as an integer in milliseconds, another as a float with trailing zeros, another omits units entirely. These differences compound across dozens of measurement fields per recording.
- **Classification labeling.** Device manufacturers assign proprietary names to ECG classifications. One vendor's firmware update renamed an industry-

standard classification format to a product-specific label — the parser had to recognize both names as equivalent without a breaking change.

- **Filter representation.** ECG devices apply signal filters during acquisition that affect waveform morphology. One vendor encodes filter state in a proprietary bitmap where bit positions map to specific filter types. Another uses enumerated strings. A third omits filter metadata in certain acquisition modes.

Per-Vendor DICOM Parsing

ML LABS built a parsing architecture that treats each manufacturer's DICOM output as a distinct dialect. A shared ECG library provides the common data structures and normalization utilities, while vendor-specific adapters handle the format divergences that matter clinically.



Each incoming file is routed to the correct vendor-specific adapter automatically. This is not optional. Two different vendors' DICOM files containing the same clinical recording differ in measurement encoding, waveform byte ordering, and classification conventions. Running them through the same parsing logic produces values that look plausible but are clinically wrong — the kind of error that surfaces months later when a cardiologist notices a measurement discrepancy between two recordings of the same patient taken on different devices.

Measurement Normalization

ECG measurements — intervals, amplitudes, axes, and derived scores — are the values clinicians read directly on reports and use for diagnostic decisions. The normalization layer solves three distinct problems across vendors.

First, measurement formatting differences: one vendor stores amplitude values as `1.200` while another stores the same value as `1.2`. In a clinical context, trailing zeros can imply measurement precision. The pipeline normalizes all measurement values to a canonical format with explicit precision metadata, so downstream consumers can distinguish genuine precision from formatting artifacts.

Second, classification mapping: device manufacturers assign proprietary names to ECG classifications, and firmware updates can rename standard labels to product-specific ones. The pipeline maps all vendor-specific labels to a unified clinical meaning, so the same condition produces the same classification regardless of which device recorded it — without requiring redeployment when a vendor changes their labeling.



The hardest interoperability problems are not protocol-level. They are the subtle measurement normalization differences between vendors that produce clinically incorrect values if handled generically.

Filter Metadata Normalization

ECG signal filters affect waveform morphology and measurement accuracy. The filter settings recorded at acquisition time travel with the recording and influence how AI models and clinicians interpret the waveform. Getting filter metadata wrong means downstream analysis operates on mischaracterized data.

Each manufacturer represents filter settings in a different proprietary format. The pipeline normalizes all vendor-specific representations into a unified filter record for every ingested ECG: which filters were active, their cutoff frequencies when available, and whether the filter state was read directly from metadata or inferred from the device model's known defaults. This record travels with the waveform through every downstream processing step — AI inference, clinical display, and research export all consume the same normalized filter information.

Waveform Deduplication and Storage

A platform that ingests ECGs through multiple paths — web upload, SMB file share from clinic networks, and EHR integration — will receive the same recording more than once. A clinician uploads a study through the web interface; the same study arrives via automated file share sync; the EHR sends it through an integration trigger. Without deduplication, the platform stores three copies and displays them as three separate studies.

ML LABS implemented deduplication at the storage layer based on clinically meaningful content rather than raw file bytes — so the same recording arriving through different paths with different wrapper metadata resolves to a single stored record. This kept storage costs predictable as ingestion volume scaled across clinical sites and eliminated a class of clinical workflow bugs where duplicate records confused clinicians reviewing a patient's ECG history. The immutable storage layer satisfies **regulatory data integrity requirements** — every access and every derivation traces back to an immutable deduplicated record.

SDK and Research Access

The platform needed to support clinical research alongside clinical operations. Researchers need programmatic access to ECG waveform data across all ingested formats — loading recordings from two different manufacturers of the same patient's ECG and comparing waveform characteristics should not require format-specific handling in the research code.

ML LABS built SDK waveform loading that abstracts vendor format differences behind a unified interface. The SDK loads waveforms from any supported format, applies the same normalization pipeline used in clinical processing, and returns a common data structure with measurement values, filter metadata, and waveform samples. This was extended to include DICOM support in the research SDK specifically, so research workflows could access raw DICOM metadata alongside the normalized representation when studying format-specific artifacts.

Integration tests for the clinical icons pipeline validated that DICOM files from each vendor produced correct classification mappings through the full processing chain — from raw file through vendor detection, classification extraction, and icon assignment.

When This Requires Prior Pattern Recognition

Vendor-neutral parsing is tractable when the format landscape is bounded — a known set of device manufacturers with documented DICOM conformance statements. It becomes categorically harder when a new vendor's implementation deviates significantly from the standard, when proprietary extensions lack documentation, or when firmware updates change encoding conventions without notice.

The difference between a first attempt and an experienced build is not incremental. Teams that have navigated the undocumented behaviors of multiple ECG device manufacturers — the trailing zeros, the renamed classifications, the bitmap filter fields, the private DICOM tags — move through new vendor onboarding at a pace that first-time builders cannot match. Deep expertise paired with AI-augmented execution now compresses what would otherwise be an expensive reverse-engineering effort into targeted adapter development with comprehensive test coverage.

First Steps

1. **Catalog your device landscape.** Inventory every ECG device manufacturer, model, and firmware version across your clinical sites — this defines the parsing scope and identifies the vendor-specific behaviors your pipeline must handle.
2. **Build one vendor adapter end-to-end.** Parse a single vendor's DICOM output into a common ECG model with full measurement normalization and filter handling, then validate against clinician-reviewed reference values before adding a second vendor.
3. **Anchor regression tests to real device output.** Collect anonymized DICOM files from every supported device and firmware version in your network, and test against these real files rather than synthetic data — synthetic files miss the encoding quirks that cause production failures.

Practical Solution Pattern

Build vendor-neutral ECG ingestion by treating each manufacturer's DICOM output as a distinct parsing dialect with a dedicated adapter, while all adapters converge on a single common model with deterministic measurement normalization, explicit filter metadata, and deduplication at the storage layer. Invest in a shared ECG library that encapsulates cross-format handling, and extend it to an SDK so that research workflows get the same normalized access that clinical processing uses — without format-specific code in the consuming application.

This architecture works because it isolates vendor-specific complexity at the adapter boundary. The clinical data model, storage layer, and every downstream consumer — AI inference, clinical reports, research exports — remain stable as new device manufacturers are added. Organizations that ship these platforms fastest concentrate parsing expertise and architectural authority in a single technical operator rather than distributing vendor-specific work across teams that must coordinate on shared data models. The measure of success is not how many adapters exist but whether a clinician can trust that the same patient's ECG produces identical measurements regardless of which device recorded it. If your organization needs to scope a vendor-neutral ingestion architecture, a **Strategic Scoping Session** can pressure-test the parsing requirements and normalization strategy before engineering work begins.

References

1. Cuevas-González, D., et al. **ECG Standards and Formats for Interoperability between mHealth and Healthcare Information Systems: A Scoping Review.** *International Journal of Environmental Research and Public Health*, 2022.
2. NEMA. **DICOM Standard.** *National Electrical Manufacturers Association*, 2024.

3. Persons, K. R., et al. **Interoperability and Considerations for Standards-Based Exchange of Medical Images: HIMSS-SIIM Collaborative White Paper**. *Journal of Digital Imaging*, 2020.
 4. U.S. Food and Drug Administration. **Cybersecurity in Medical Devices**. *Regulatory Reference*, 2025.
 5. U.S. Food and Drug Administration. **Artificial Intelligence-Enabled Medical Devices**. *Regulatory Reference*, 2025.
 6. Giese, D. **Why Does the DICOM Standard Exist?**. *Innolitics*, 2020.
 7. HL7 International. **FHIR Standard**. *HL7*, 2024.
-



ML LABS

Custom AI Systems for High-Value Workflows

mlabs.com