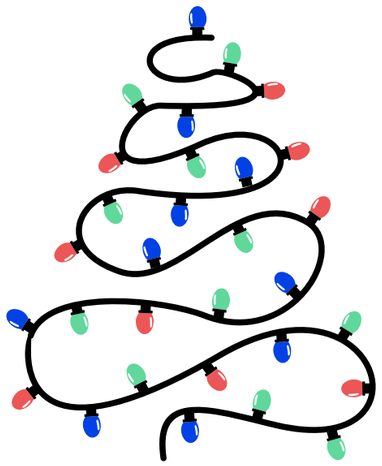
data rabbit

In 2025, presents are brought by rabbits...



Your tech guide into 2026

Before Holiday season...

We've taken the time to analyze the upcoming changes and opportunities within our field of expertise. During this holiday season, we'd like to share a few things worth keeping an eye on in the year ahead. **So let's unwrap your gifts and get ready for new challenges.**

For ways to increase revenue in 2026 go to page 3

For exclusive analysis - From siloed systems to unified, searchable medical data: a reference architecture - check out page 6

Compliance changes - your mandatory 2026 checklist on page 8

Ways to increase revenue in 2026

1. Monetising medical data

The demand for high-quality, real-world medical data to train next-generation AI models is peaking. This is a high-margin, scalable revenue stream that leverages your existing data assets. The EHDS framework and evolving data-sharing models in the USA create a legal and commercial pathway for this secondary use.

Actionable Tip: Do not sell raw data. Instead, focus on creating curated, obviously anonymised, and annotated data sets that are immediately usable by AI developers. This value-added service commands a premium.

Revenue Driver	Actionable Steps for 2026
Data Brokerage	Define internal policies for data anonymisation and consent management that align with EHDS/GDPR (EMEA) and HIPAA/CCPA (USA).
Curated Data Sets	Identify your most valuable data assets (e.g., rare disease imaging, longitudinal patient records) and package them as commercial products.
Strategic Partnerships	Form exclusive partnerships with large pharmaceutical companies and AI foundation model developers who require your specific data niche.

Ways to increase revenue in 2026

2. Cost optimisation (with the help of AI)
The fastest way to increase profit is to reduce operational friction. AI is no longer a futuristic investment; it is a proven tool for cutting high-overhead costs and accelerating the time it takes to get paid. Focus on the highest-impact areas where AI can free up high-cost human labour.

Actionable Tip: Target the administrative ”swivel-chair” processes that require high-cost human labour and cause payment delays. A 10% reduction in administrative overhead can translate directly into millions in profit.

Cost Optimization Area	Actionable Steps for 2026
Prior Authorization (AI-PA)	Deploy AI tools to automate the submission, tracking, and follow-up of prior authorisation requests.
Clinical Documentation	Deploy clinical AI assistants. Use AI for note-taking, patient history summarisation, and administrative tasks to free up high-value clinical staff time.
Claims Processing	Automate denial management. Use machine learning to predict and automatically appeal claim denials, identifying root causes for systemic fixes.

Ways to increase revenue in 2026

3. Personalised patient care

The future of healthcare is moving beyond generalised care to personalised patient journeys enabled by multimodal data and intelligent automation. This strategy creates value by improving patient outcomes, increasing loyalty, and reducing the cost of support.

Actionable Tip: Shift your focus from treating populations to optimizing the experience and outcome for the individual patient.

Revenue Driver	Actionable Steps for 2026
Precision health analytics	Combine EHR, genomics, imaging, and real-time wearable data to create a 360-degree patient profile.
Personalised care pathways	Deploy machine learning models to predict individual patient risk (e.g., readmission, adverse events) and automatically offer personalised care interventions.
Support automation	Implement voice bots for high-volume, low-complexity tasks such as clinic queue management, after-care follow-ups, and providing medical knowledge assistance.

From siloed systems to unified, searchable medical data: a reference architecture

Collecting terabytes is easy; unification, searchability and governance make it valuable.

Healthcare organizations rarely lack data. In fact, it's quite the opposite (figures vary, but some reports estimate healthcare for even 30% of global data production).
But they lack data that is discoverable, trustworthy, and safe and easy to use at scale – especially once you're joining multiple modalities (EHR events, labs, notes, imaging, waveforms, claims) and start to operate in multi-terabyte estates.

If any of this sounds familiar, you're already feeling the problem:

- "We can't answer **key operational questions** without a multi-week effort."
- "We spend **more time reconciling data and definitions** across sources than analyzing outcomes."
- "AI/analytics **is blocked** because we can't guarantee PHI won't leak"
- "When someone asks 'why is this number correct?', we **can't answer it defensibly**."
- "Audit questions turn into cross-system log **archaeology even Indiana Jones wouldn't be able to handle.**"

So, what can a platform look like when it actually fixes these failure modes?
The first principle is unifying data in one place. But the goal isn't just to land terabytes in one place – it's to make data findable, safe, and easy to use across teams. Here is a reference architecture, plus the tips and best practices that we distilled from delivering multiple real-world implementations

End Users

Administrators

Discovery & Access

Web UI / data portal
(business owners, clinical ops, research ops)

API / SDK
(data/AI teams, engineers, product teams)

LLM assisted query builder
(business users + analysts; also helps data/AI teams speed up exploration)

Global Supporting Services

Identity & Policy
(IAM / roles / access control)

Audit & Compliance
(access/audit trail)

Observability
(logging / monitoring)

Security Posture

Multimodal data storage

RAW storage (immutable)
- structured
- documents
- imaging (DICOM)
- signals (waveforms/time series)
- omics
- ...

Metadata / catalog
(searchable)

Lifecycle / tiering

Ingest Layer

Metadata extraction

Data validation
(deduplication, format fixes (if possible), etc.)

Privacy transformations
((pseudo)anonymization, etc.)

Connectors / Adapters

"Raw" data sources

Clinical systems
- EHR/EMR
- ADT/MPI

Diagnostics
- Lab resultss
- Imaging
- Pathology
- Omics

Signals & externals
- Devices & waveforms
-Wearables / RPM
-Claims / payer feeds

Do’s and dont’s in building medical data platforms

✅ Best practices we’ve seen work (at scale)

- **Multimodality is a multiplier.** While you might not have all the types of data from the diagram (either at all, or at least at the beginning) – combining modalities enables joins, cohorts, and longitudinal views, with value much greater than the sum of its parts.
- **Batch by default.** Most updates don’t need real-time; batch reduces cost and operational burden. Further optimize the cost by leveraging cost-efficient compute (e.g. spot instances if using the cloud).
- **Repeatable ingest.** Ingestion should be both incremental + idempotent, while also handling validation + quarantine so reprocessing doesn’t create duplicates or ambiguity.
- **Make metadata the product.** Extract all possible fields/tags + provenance/lineage to a dedicated base, and store pointers to raw. If something is not in metadata, it’s effectively invisible.
- **Prioritize pseudonymization.** Pseudonymization allows joins – utilize it whenever there might ever be a need for some entity-level linkage; reserve irreversible anonymization only for categorically non-linkable use cases (or when compliance demands it).
- **Cater to different users.** API/SDK for tech customers, intuitive and comprehensive UI for others. LLM-assisted query building (NL → structured metadata query) is a new possibility that greatly simplifies adoption for all audiences.
- **Separate hot search from cold raw.** Search in metadata; keep raw immutable and cost-optimized (tiering/retention).
- **Make your API boring.** Utilize industry standards (especially for data formats). Contracts should be stable, and with whatever developers would expect (pagination, versioning, throttling, etc.).
- **Ultra-granular permissions.** Not only design everything, with the principle of least privilege, but also mix ABAC + RBAC to make it as granular as possible (down to the single operations and object level where required)

⚠️ Watch outs/anti-patterns

- **“We’ll add (some) metadata later.”** No, you won’t – and search will be limited. Someone will need it, sooner or later, so if it’s somewhere – make sure to add it to your metadata storage.
- **No idempotency/dedup.** Conflicting counts = no trust.
- **Don’t fight standards.** Keep DICOM/FHIR/BAM/VCF – avoid inventing new formats. It leads to immediate integration debts and kills the adoption.
- **Avoid generic LLMs for anonymization.** It’s both costly, and accuracy is not there yet – utilize task-specific models instead.
- **Audit as app logs.** Audit trail should be handled from day 1. Use structured audit events (who/what/when/why + actions context).

Building a high-volume, multimodal medical data platform that is truly searchable and provides value, while also being governed and compliant isn’t a small feat. The points above are the core patterns that consistently work. But there’s more nuance behind each layer that doesn’t fit on one page. If you’re building something similar and want to sanity-check a couple of design choices, feel free to reach out – we’re happy to help and share knowledge.

Compliance changes — your mandatory 2026 checklist - EMEA region

The EU AI Act

The EU Artificial Intelligence Act is the world's first comprehensive legal framework for AI. For the health sector, the good news is that the AI Act largely builds upon existing regulatory foundations (European MDR, American HIPAA, etc.). This means much of the compliance is a matter of convergence and gap analysis rather than starting from scratch. However, the Act introduces stringent requirements for High-Risk AI Systems (e.g., those used for diagnosis or treatment), with the majority of the Act becoming fully applicable on August 2, 2026.

Crucial Change: The focus shifts to operationalizing compliance. You must ensure your AI development and deployment pipelines are fully integrated with a robust QMS that specifically addresses the AI Act's unique requirements for:

- **Data Governance** (strict standards for training data),
- **Technical Documentation** (comprehensive records demonstrating compliance),
- **Human Oversight** (mechanisms to ensure that human operators can effectively monitor and intervene in the AI system's operation).

European Health Data Space (EHDS)

The **European Health Data Space** (EHDS) will fundamentally change how health data is accessed and used, facilitating both primary use (patient access) and secondary use (research and innovation). While full secondary use is expected to mature in 2027, 2026 is the critical preparation year.

Crucial Change: The EHDS will mandate a high degree of interoperability for Electronic Health Record (EHR) systems across the EU. Crucially, this creates an opportunity for the opening up of certain legacy systems (e.g., Hospital Information Systems - HISs), which were previously closed data silos.

Regulatory Convergence and Quality Standards

The global regulatory landscape is converging on a high standard of quality management. The core **MDR/IVDR** remain in force, but the industry is pushing for a more functional and sustainable regulatory system through a targeted revision proposal expected to be finalised in 2026.

Furthermore, the FDA's **Quality Management System Regulation** (QMSR), which aligns the US system with ISO 13485:2016, becomes effective on **February 2, 2026**. This makes robust compliance with ISO 13485 a de facto requirement for global market access. The full implementation of ISO 27001:2022 is also essential, as the transition deadline has passed.

Compliance changes — your mandatory 2026 checklist - North America

Action Item	Critical Deadline / Milestone	Check box
Integrate AI Act requirements into QMS for all High-Risk AI systems.	August 2, 2026 (Full Applicability)	
Ensure QMS is fully compliant with ISO 13485 for global market access.	February 2, 2026 (FDA QMSR Effective)	
Verify full maturity and audit readiness of ISMS against ISO 27001:2022.	Ongoing 2026 (Full maturity expected post-Oct 2025 deadline)	
Assess and upgrade EHR/data systems to meet EHDS interoperability standards.	Ongoing 2026 (Preparation for 2027 EHDS data access)	
Review final MDR/IVDR targeted revision and adjust QMS accordingly.	Ongoing 2026 (Anticipate finalization and adjust QMS)	

Save this for the future!

Compliance changes — your mandatory 2026 checklist - North America

FDA QMSR

FDA is moving the device quality framework to the **Quality Management System Regulation** (QMSR) (ISO 13485:2016 incorporated by reference) and will start enforcing it on Feb 2, 2026—with a new inspection process (QSIT withdrawn).

Crucial change:
FDA is making "quality evidence" inspection-real.
If your AI is part of a regulated device/SaMD, you need systems that can produce proof of how each model version was built, tested, released, and monitored – fast. Importantly:

- **Mature, audit-ready MLOps (Machine Learning Operations) are must:** version + retain dataset lineage, models versions, evaluation results, approval – so you can answer "what shipped and why" on demand. Post-deployment monitoring and drift detection should also be in place.
- **Treat model updates as controlled releases:** retraining, threshold changes, prompt/config changes must go through defined impact assessment + verification/validation + sign-off (not exactly "new," but it's now much harder to rely on informal/undocumented workflows).
- **Vendor/model supply chain is now in the blast radius:** FDA explicitly says it can inspect records like supplier audits, quality audits, management review that used to be exempt – so third-party model/data/tooling governance must be documented and retrievable.

HIPAA Security Rule NPRM

HHS/OCR published a **proposed** HIPAA Security Rule overhaul aimed at stronger cybersecurity for ePHI, including removing the "addressable vs required" distinction and introducing more prescriptive requirements.
While it's still proposed, current regulatory agenda targets final action around May 2026 and the direction is clear – the security baseline for platforms handling ePHI, including data/AI ones, is rising.

Crucial change: HIPAA security shifts from "we did what's reasonable" to "implement these controls and prove it" - directly hits data lakes, analytics stacks, and MLOps environments.

- **Map and inventory ePHI across the data/AI stack:** proposed requirement for a technology asset inventory + network map showing movement of ePHI - that includes feature stores, training data buckets, model registries, logs, BI layers, ETL jobs, etc.
- **Baseline controls become universal:** proposed encryption at rest/in transit, MFA, network segmentation, plus routine vulnerability scans (≥ every 6 months) and pen tests (≥ annually). That causes "shadow" ML environments and ad-hoc pipelines to become compliance liabilities.
- **Operational rigor + vendor pressure:** proposed requirements include regular compliance audits, tighter incident/contingency expectations (e.g., restore within 72 hours), and BA notification timelines. Platforms need mature observability, incident response hooks, and vendor controls.

Compliance changes — your mandatory 2026 checklist - EMEA region

Action Item	Critical Deadline / Milestone	Check box
Ensure QMS is fully compliant with ISO 13485 – by having audit-ready MLOps with evidence trail + controlled model updates + supplier/AI-vendor audit records retrievable	February 2, 2026 (FDA QMSR Effective)	
HIPAA Security Rule (NPRM): map where ePHI lives and moves across the data/AI stack + run a gap assessment (roadmap vs proposed controls)	May 2026 is the target final action (regulatory agenda; may shift)	
HIPAA Security Rule (Final): roll out the required controls + documentation (audit-ready)	Within 180 days after the Final Rule's effective date (effective date is typically 60 days after publication)	

Check all of the boxes!

From concept to reality:
Intelligent solutions for life sciences

Thank you for 2025
& see you again in 2026

If you'd like similar insights and the latest healthcare & life sciences trends, subscribe to CliniSignal by datarabbit here.

Contact:
contact@datarabbit.ai
+ 48 796 750 136
www.datarabbit.ai