AEGIS CIPHER

AI Data Governance in Healthcare

Existing Regulatory Frameworks
HIPAA Minimum Necessary, Secondary Use, Model Training, and Data Sovereignty

Melissa Gaffney | Managing Partner & Chief AI Officer | Aegis Cipher LLC

1. TECHNOLOGY IS NOT A REGULATED SUBJECT; USE IS.

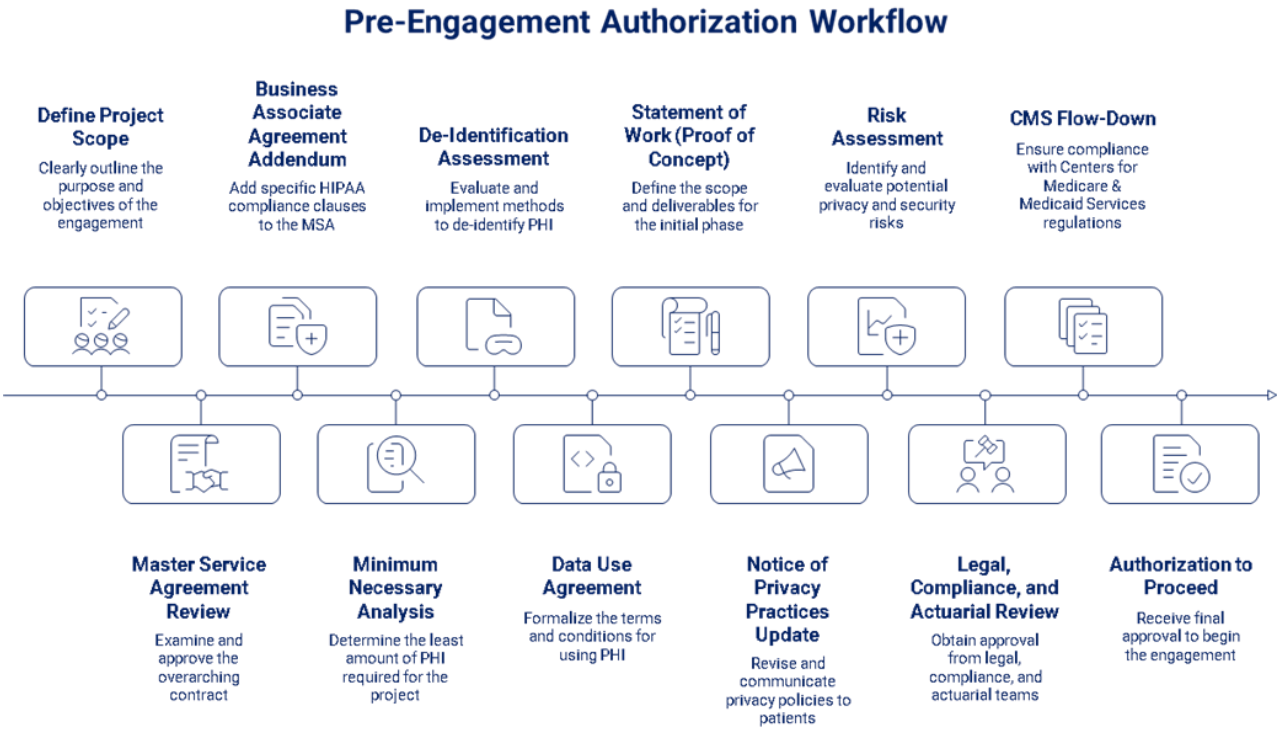
Healthcare data governance operates through four interlocking instruments. Each narrows the permissible universe of data use established by the layer above it. A vendor cannot acquire rights at a lower layer that were not granted at a higher one.

Five Industries comprise 80% of AI/ML/LLM use and have used it for a decade or more. The recent advancements of model training have put us in a global frenzy. It is exciting. Success and market leadership will depend on:

1) proactively address creative destruction principles that are needed to sustain innovation
2) leverage domain experts, behavioral scientists, and master prompt communication nuances
3) understand and act on the fact that AI can be both not new and emerging

This narrative is part of a series to help those entering the technology or domain of Healthcare, Life Sciences, and Insurance get a lay of the land with the intent to move a few "unknown, unknowns to known, unknowns and known, unknowns to knowns. In another write up, I focus on governance, standard covenants and existing regulatory and legal frameworks.

This narrative expands on that but has emphasis on the "Minimum Necessary" provision in HIPAA and the additional risks and complexity when introducing AI features and enablement.



2. REGULATIONS ARE TECHNOLOGY-AGNOSTIC

HIPAA's framework is technology-neutral by design. The Office for Civil Rights (OCR) has confirmed this explicitly. The **Privacy Rule at 45 CFR §164.502** does not distinguish between a human reviewing a medical record and an algorithm processing the same record. It governs *use* and *disclosure*, not the mechanism by which use or disclosure occurs. The Security Rule at 45 CFR Part 164, Subpart C, requires administrative, physical, and technical safeguards for electronic PHI regardless of the system architecture that stores or processes it. **A neural network does not receive an exemption that a relational database does not.**

AI's capacity to infer, generalize, retain, and **recombine creates risk exposure *within* existing frameworks.** When an AI model memorizes training data, that is a retention issue governed by existing data retention and minimum necessary provisions.

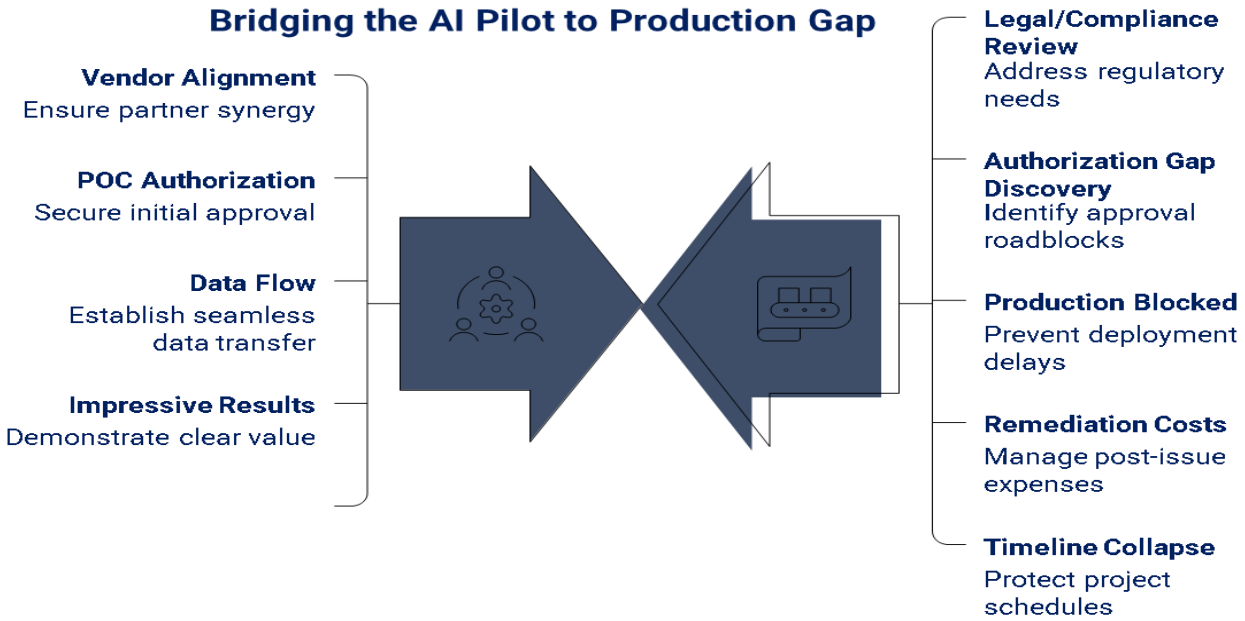
When a model's output can be reversed engineered to identify individuals, that is risk governed by the Breach Notification Rule at 45 CFR Part 164, Subpart D. When a vendor trains a model on a covered entity's data and deploys that model for other clients, that is an unauthorized disclosure.

De-identification is a process, not a permanent state.

Safe Harbor methodology removes 18 specific identifiers per 45 CFR §164.514(b)(2), but it does not account for re-identification risk through AI inference, small cell sizes, or combination with external datasets. Model outputs can re-identify individuals even when trained on data that passed Safe Harbor review. Expert Determination under §164.514(a) is required for each processing modality including AI model training because the risk profile changes with each new use. ***The regulations are not missing. They have not applied. There is a difference and the distinction carries legal consequence.***

3. THE REAL REASON AI PILOTS FAIL

Here is the standard scenario, and it plays out with remarkable consistency across the industry:



4. HIPAA MINIMUM NECESSARY AND THE TRAINING PAYMENT OR OPERATIONS (TPO) QUESTION

The minimum necessary standard at 45 CFR §164.502(b) is one of the most consequential and most consistently underestimated provisions in HIPAA's Privacy Rule. It requires covered entities and business associates to make reasonable efforts to limit the use, disclosure of, and requests for protected health information to the minimum necessary to accomplish the intended purpose. This is not a suggestion.

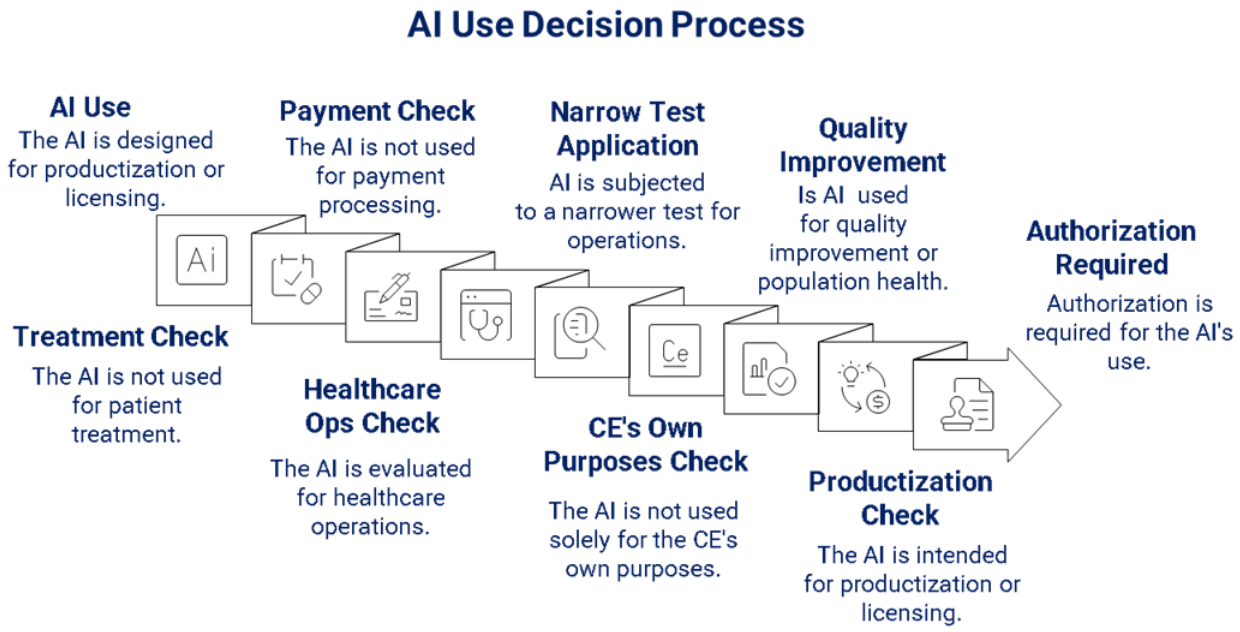
It is a standard, and **it applies to every use and disclosure except those made for treatment, those made to the individual, those made pursuant to an authorization, those required for HIPAA compliance**, and those required by other laws. **When the intended purpose is training an AI model,** the minimum necessary analysis becomes significantly more complex.

What specific data elements are required for model training? Is the full longitudinal record necessary, or would a subset suffice? Are demographic fields required for the model's function, or are they being included by default? Has the volume of records been scoped to the minimum population necessary to achieve statistical validity? These questions must be answered *before* data flows, and the answers must be documented.

4.1 USE VS. DISCLOSURE

HIPAA draws a critical distinction between *use* (sharing, employing, applying, utilizing, examining, or analyzing PHI within the entity that maintains it) and *disclosure* (releasing, transferring, providing access to, or divulging PHI). AI model training can constitute either or both, depending on where the training occurs, who controls the training environment, and what happens to the resulting model.

The threshold question is whether the AI use qualifies as **Treatment, Payment, or Healthcare Operations (TPO) under 45 CFR §164.501** because TPO uses do not require individual authorization. The answer, for most AI model training scenarios, is **no**.



AI model training is not treatment, if it can qualify as operations is very narrow and likely only in the event of internal processes or clean delegation. **Healthcare operations defined under §164.501** include quality assessment and improvement activities, care coordination, clinical staff training, population-based activities relating to improving health or reducing healthcare costs, and certain administrative functions, *for the covered entity's own purposes*. Training a model that will be productized, licensed, or sold even if derived from a health plan's enterprise data

warehouse (EDW) is not healthcare operations for the originating covered entity. It is product development. It is commercial activity. It requires authorization.

Training a vendor's model using a covered entity's data is almost certainly not within the business associate agreement's authorized scope unless that scope has been explicitly and narrowly defined to include AI model training as a permitted use.

Healthcare operations under HIPAA §164.501 authorize certain population-based activities *for the covered entity's own purposes*. The moment the output; a model, a set of weights, an algorithm, is designed to be productized, licensed, deployed for other clients, or retained by the vendor as intellectual property, the activity ceases to qualify as the covered entity's healthcare operations. This is not healthcare operations. This is product development using someone else's PHI.

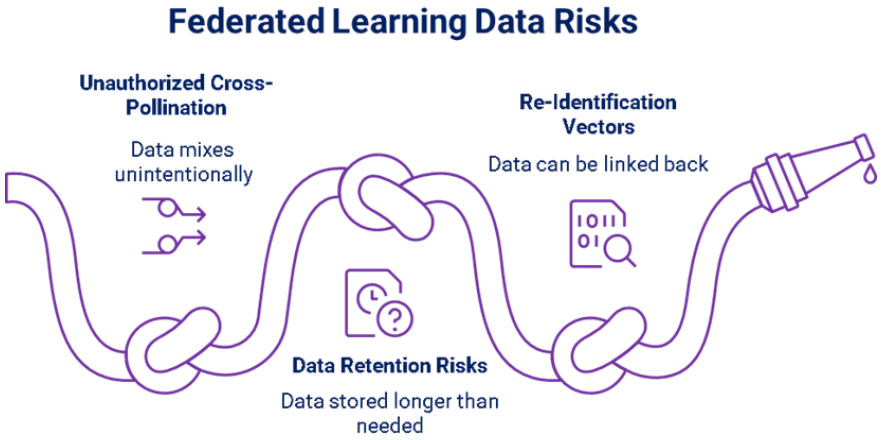
5. THE BAA PROBLEM

Most Business Associate Agreements in effect today were drafted before generative AI existed as a commercially viable technology. They are not consent instruments nor can they be expanded without a formal change. **They are policy and procedure documents required under 45 CFR §164.504(e)** that specify the permitted and required uses and disclosures of PHI by the business associate.

Warning: Retroactive Amendments are Privacy Incidents

Some organizations attempt to cure unauthorized AI data use by executing a retroactive BAA amendment. A retroactive amendment cannot cure a violation that has already occurred. Aamendments governs future use; it does not retroactively authorize past use. **The breach analysis obligation under 45 CFR §164.402** is triggered at the time of the unauthorized use, not at the time an amendment is signed.

6. FEDERATED TRAINING MISCONCEPTIONS

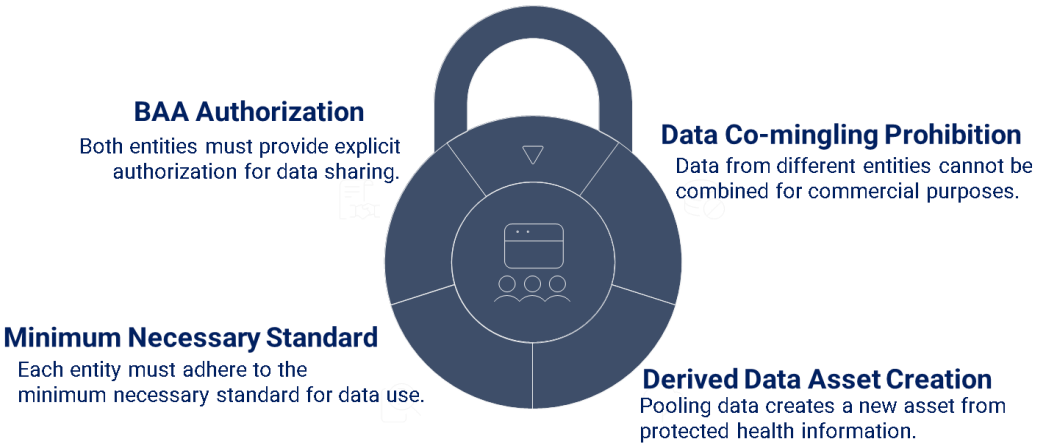


De-identification under HIPAA is a process specifically, it is a determination that health information does not identify an individual and that there is no reasonable basis to believe it can be used to identify an individual **(45 CFR §164.514(a)). Safe Harbor methodology removes 18 specific identifiers. AI brings a clear and present danger for re-identification, at all times.**

Expert Determination under §164.514(a) requires a qualified statistical or scientific expert to determine that the risk of identifying any individual is "very small." This determination must be made for each processing modality and AI model training is a materially different processing modality than business intelligence reporting.

The second error is the assumption that because data resides **in a covered entity's own enterprise data warehouse, the covered entity has unlimited discretion over its use**. It does not. The covered entity is a custodian, not an unrestricted owner. The third error and the most consequential involves **federated learning and multi-client model training**.

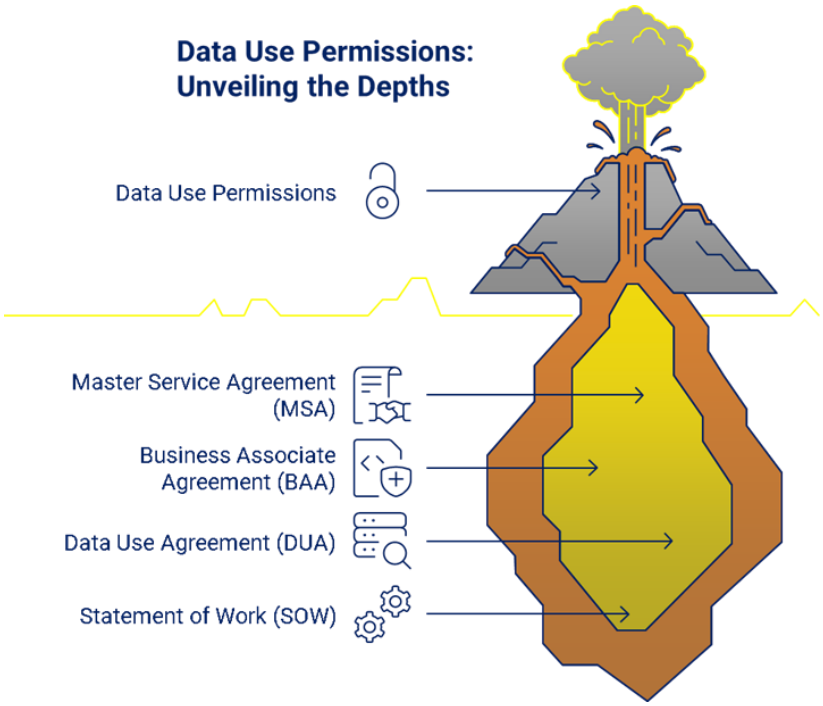
No single entity can authorize a vendor to commingle their data with another covered entity's data for model training without *all* covered entities' explicit BAA authorization. A vendor cannot:

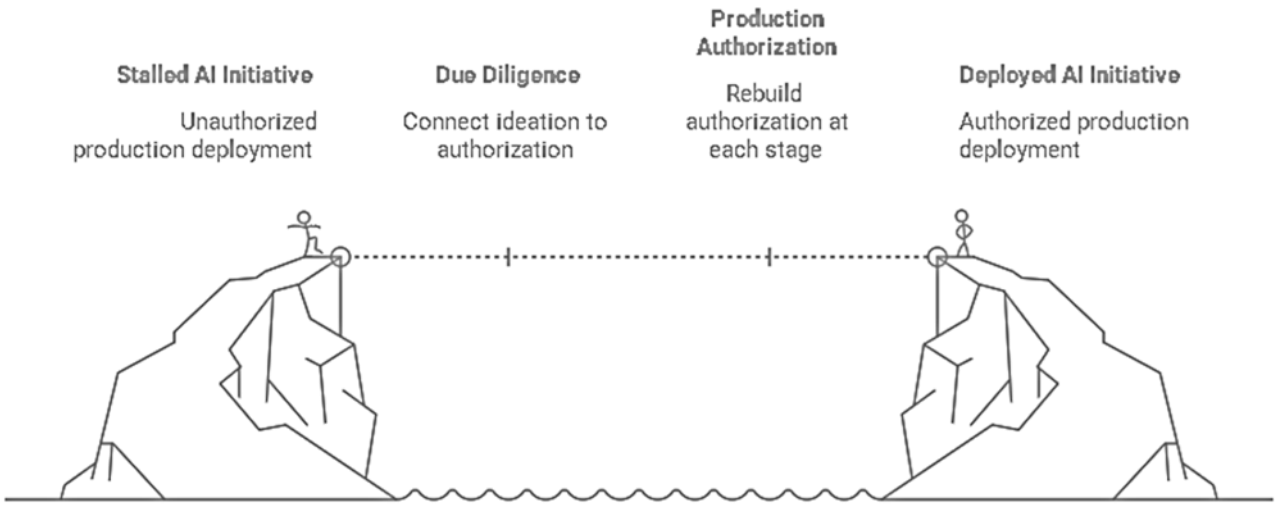


7. THE COVENANT CHAIN — MSA, BAA, DUA, SOW

Healthcare data governance does not operate through a single instrument. It operates through four interlocking contractual covenants, each of which narrows the permissible universe of data use established by the layer above it. These instruments are hierarchical. They are sequential.

And they are non-negotiable in the sense that a vendor cannot acquire rights at a lower layer that were not granted at a higher one. ***"A vendor cannot acquire rights at a lower layer that were not granted at a higher one***





The AI readiness assessment for any healthcare data engagement begins at the MSA layer and works downward. Each instrument must be reviewed for AI-specific adequacy, and are sequential:

Does the MSA establish clear data ownership, IP rights, and secondary use prohibitions for AI/ML?

Does the BAA explicitly authorize (or prohibit) the specific data processing modalities that AI requires; including model training, weight retention, and inference deployment?

Is a dataset-specific DUA in place that defines the permitted use at the granularity required for AI, not just "analytics" or "reporting," but specifically "training a [model type] for [purpose] using [data elements]"?

Has an engagement-specific SOW been executed that scopes the AI work, defines model ownership, establishes production transition requirements, and distinguishes between POC and production?

If the answer to any of these questions is "no," the AI engagement is not authorized. It is not a question of risk tolerance. It is a question of compliance. The authorization chain must be complete before data flows.

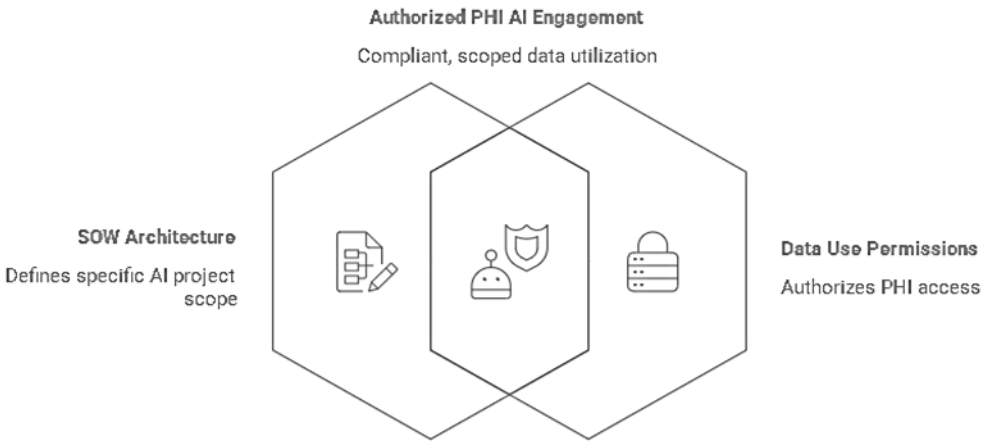
8. DATA SOVEREIGNTY HIERARCHY

Data sovereignty in healthcare is not a metaphor. Rights flow downward through a delegation chain, and they cannot be expanded at lower levels. Each layer in the hierarchy holds a specific relationship to the data: some hold fundamental rights, some hold custodial obligations, and some hold delegated, scoped, purpose-limited permissions that expire when the purpose is fulfilled or the contract terminates.

Contracts do not nullify individual rights. State and federal laws are sovereign. The highest level of data sovereignty is the patient or member. The second is the provider or plan. Everyone else; every vendor, every subcontractor, every AI platform holds a custodial role that is bound by the rights and the laws that supersede all contracts.

9. NOTIFICATION, NPP, AND OPT-OUT OBLIGATIONS

Introducing artificial intelligence particularly for secondary uses, model training, or product development is a material change in how a covered entity uses protected health information. It is not a refinement of existing operations or evolution of business intelligence. It is a new processing purpose that individuals were not notified of when their data was originally collected. And under HIPAA, material changes in data use trigger notification obligations that cannot be satisfied by silence, by assumption, or by burying new language in a sixty-page privacy notice.



9.1 HIPAA NOTICE OF PRIVACY PRACTICES (45 CFR §164.520)

The Notice of Privacy Practices (NPP) is the primary instrument through which covered entities inform individuals about how their PHI will be used and disclosed. When a covered entity introduces AI-based processing, the NPP must be updated to describe these uses with reasonable specificity.

For health plans, the NPP must be provided at enrollment and upon material revision. Introducing AI-based data processing is a material revision. The revised NPP must be distributed, not merely posted on a website. **Per §164.520(c)(1)(v), for health plans, the revised notice must be provided within 60 days of a material revision to any individual then covered by the plan.**

9.2 VENDOR AND SAAS CLIENT NOTIFICATION

Separate from individual notice, covered entity clients must be notified before a vendor expands data use to include AI processing. This notification must be specific: what data will be used, for what purpose, to train what type of model, with what retention period, and under what ownership terms. The covered entity client must have a meaningful opportunity to object before the use begins not after the model has already been trained on their data. Under the BAA, the business associate's permitted uses are defined and bounded. Expanding those uses without a covered entity's informed consent is unauthorized use of PHI, a potential breach in the Breach Notification Rule.

9.3 SECURITY AND RISK NOTICE UPDATES

The HIPAA Security Rule requires covered entities and business associates to conduct risk assessments. When processing modalities change and introducing AI is a fundamental change in processing modality risk assessments must be updated. New threat vectors specific to AI must be assessed, including:

9.4 OPT-IN VS. OPT-OUT FRAMEWORKS

For secondary uses of PHI, which is uses beyond the original purpose for which the data was collected, opt-in is the appropriate default. Individuals should affirmatively consent to the use of their health information for AI model training, product development, or research purposes that were not disclosed at the time of data collection.

Where opt-out mechanisms are used, they must be operationally real, not theoretical. An opt-out that requires a patient to navigate a multi-step web portal, submit a written request, and wait 30 days for processing is not a meaningful opt-out. It is a friction barrier designed to suppress exercise of rights. Meaningful opt-out requires clear communication, accessible mechanisms, and prompt effectuation.

9.5 THE "PUSH" OBLIGATION AND RETROACTIVE RATIFICATION

Posting a revised NPP on a website is insufficient for health plans. The obligation is affirmative delivery a "push" obligation that requires the covered entity to deliver the notice to the individual, not merely to make it available.

For health plans, this means mailing or electronically delivering the revised notice to all current enrollees. Finally, and critically: retroactive ratification does not cure violations. If PHI was used beyond the authorized scope. The amendment governs future use. The past exposure remains, and the breach analysis obligation under 45 CFR §164.402 applies to the period of unauthorized use.

10. CMS AND FDR REQUIREMENTS

For Medicare Advantage (MA) organizations, the regulatory obligations extend beyond HIPAA. The contract between an MA organization and the Centers for Medicare & Medicaid Services (CMS) imposes specific requirements under 42 CFR §422.504 that must flow down to all First Tier, Downstream, and Related Entities (FDRs). These are mandatory floor requirements. They are not negotiable. They cannot be modified by contract between the MA organization and its vendor. They cannot be waived. And non-compliance by an FDR is the MA plan's non-compliance with CMS — which can result in sanctions, civil money penalties, intermediate sanctions under §422.752, or contract termination.

When an AI vendor operates as an FDR, which it does whenever it processes Medicare Advantage enrollee data as a downstream entity, it inherits the full complement of CMS-mandated obligations. These are not optional add-ons. They are conditions of participation in the Medicare program, and they apply to the AI vendor with the same force they apply to the MA organization itself.

10.1 KEY MANDATORY CMS FLOW-DOWN PROVISIONS

The CMS regulatory overlay is not parallel to HIPAA it is additive. An AI vendor processing Medicare Advantage data must comply with *both* HIPAA and CMS requirements simultaneously. The failure to flow CMS requirements to an AI FDR is not merely a contractual gap , it is a regulatory violation by the MA organization, attributable to the organization regardless of whether the FDR was aware of the requirements. However, FDR are required to conduct annual training in this space.

Requirement	Description and AI Implications
HHS/Comptroller Audit Rights	HHS and designers have the right to audit, evaluate, and inspect any books, contracts, records and includes training data provenance, model architecture documentation, validation results, audit logs.
Confidentiality and Accuracy	Must maintain the confidentiality and accuracy of enrollee records. AI models that generate or modify enrollee-facing outputs (coverage, risk scores), errors must be correctable.
Delegation Monitoring	MA org must specify what functions are delegated and are subject to ongoing performance monitoring , not merely initial validation & report model performance, drift, adverse outcomes.
Enrollee Protection	Comply with enrollee protection requirements. AI-driven decisions that affect enrollee access to benefits, coverage determinations, or care coordination must include human oversight and appeal rights.
Compliance Program Reqs	FDRs must participate in the MA orgs compliance program. AI vendors must implement compliance training, reporting mechanisms, and corrective action procedures specific to AI-related risks.
Training and Screening	Fraud, Waste, and Abuse (FWA) training and must screen employees /subcontractors against the OIG exclusion + GSA debarment list. Vendor employees w access to enrollee data must be screened.
Offshore Entity Disclosure	FDRs must disclose offshore subcontractors. If AI model training or data processing occurs offshore including cloud computing in foreign data centers this must be disclosed and may require CMS approval.
Subcontractor Flow-Down	All CMS-mandated provisions must flow down to subcontractors. If the AI vendor uses sub-processors (cloud providers, annotation services, model validation firms), sub-processors inherit same obligations.
Annual Attestation	FDRs must provide annual attestation of compliance. AI vendors must attest annually to their compliance with all applicable CMS requirements, including data handling, security, enrollee protection obligations.

11. LEGAL, COMPLIANCE, AND ACTUARIAL

Introducing or replacing technology particularly technology that changes how data is processed, how decisions are informed, or how risk is assessed changes the organization's risk posture. This change triggers review across

three functions, each of which must independently assess the implications. These reviews are not sequential gates to be passed through on the way to deployment. They are concurrent, substantive analyses that may each independently halt an AI initiative if the governance prerequisites are not met.

11.1 LEGAL REVIEW

The legal review must answer a threshold question: Does the existing contractual architecture — MSA, BAA, DUA, and SOW — authorize the new technology's data processing modality? If the existing BAA authorizes "analytics in support of health plan operations," does that authorization extend to training a machine learning model? In most cases, the answer is no. The legal review must also assess intellectual property ownership implications for models trained on client data, indemnification adequacy for AI-specific risks (model error, bias, discrimination), and the applicability of emerging state AI laws to the specific engagement.

11.2 COMPLIANCE REVIEW MUST INCLUDE:

HIPAA minimum necessary analysis specific to the AI use case, not recycled analysis

De-identification adequacy assessment recognizing that prior Safe Harbor or Expert Determination certifications do not auto apply to a new processing modality with a different re-identification risk profile

Access controls and audit logging for AI systems, including logging of what data was accessed for training, who authorized it, and what model artifacts resulted

Breach risk assessment that accounts for AI-specific threat vectors: model inversion, membership inference, training data extraction, and adversarial manipulation

State law overlay analysis particularly for states with comprehensive privacy or AI-specific legislation, including Colorado, Illinois (BIPA implications for biometric data), Texas, New York, and others

11.3 ACTUARIAL REVIEW, OVERLOOKED TRIGGER

The actuarial review is the review that almost no one performs and it is the one with the most significant downstream consequences for regulated health plans.

AI-driven decision support in utilization management (UM), risk stratification, claims adjudication, or care management changes the actuarial basis of the insurance product. This is not an abstract concern. If an AI model changes denial rates, even marginally, pricing assumptions built on historical denial patterns may no longer hold.

If a model changes approval patterns for high-cost procedures, the medical loss ratio projection shifts. If a model changes risk scores for a Medicare Advantage population, the bid assumptions filed with CMS may be inaccurate. The implications are concrete:

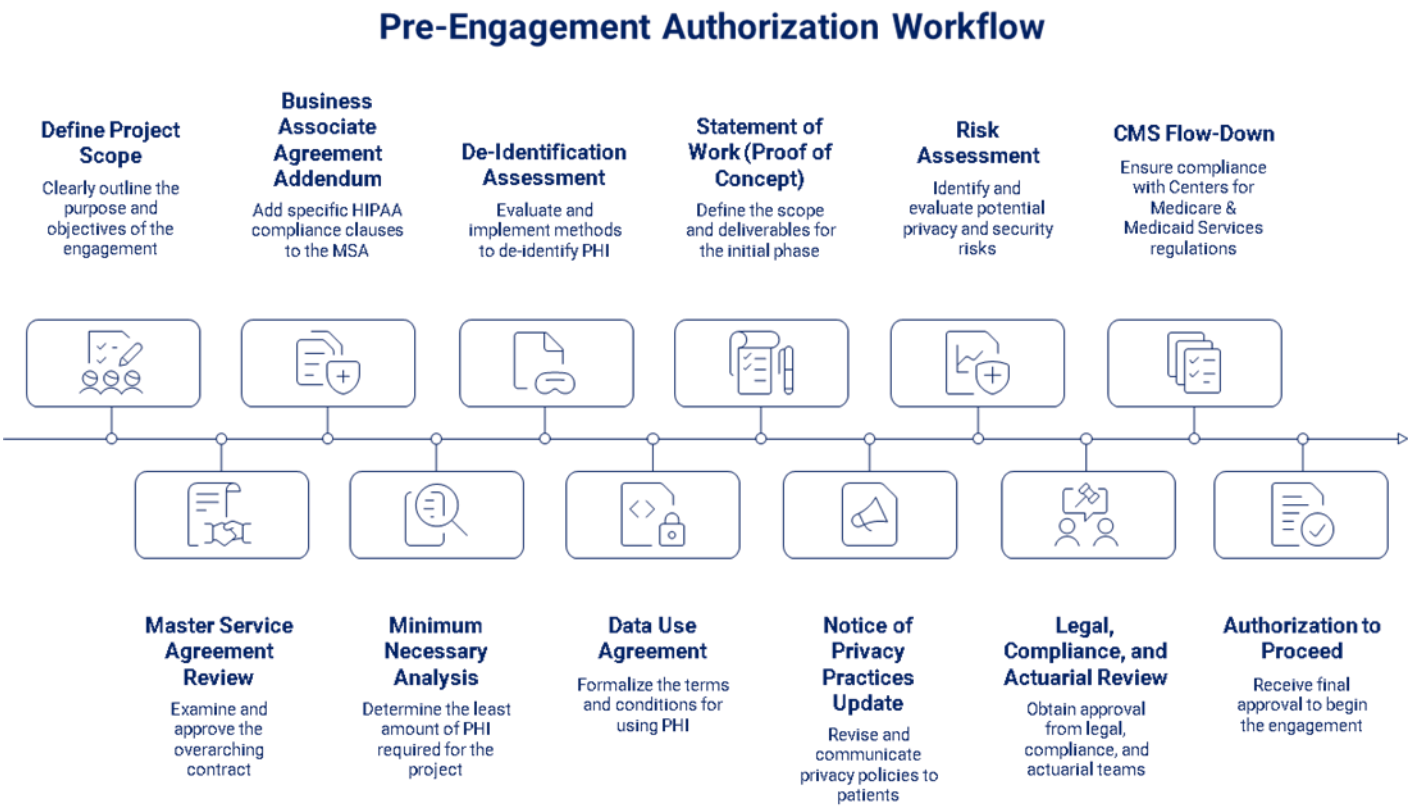
11.4 THE ACTUARIAL BLIND SPOT

An AI model that improves UM efficiency by 15% is not just a technology success story. It is a change in the actuarial basis of the product. If that change was not disclosed in the rate filing, the rate filing may be inaccurate. If the rate filing is inaccurate, the product may be mispriced. If the product is mispriced, the plan faces regulatory exposure from both CMS and state insurance regulators. The technology team that celebrates a 15% improvement without notifying the actuarial team has created a compliance problem, not a competitive advantage.

12. PRE-ENGAGEMENT AUTHORIZATION SEQUENCE AND CONCLUSION

The authorization sequence that must be completed before any AI use of healthcare data begins is not a compliance formality. It is not a bureaucratic obstacle to innovation. It is the mechanism that protects initiative

investments and enables proofs of concept to move to production. Every AI initiative that skips this sequence risks not only regulatory exposure but the loss of the initiative itself — and with it, the organizational credibility required to pursue the next one. The complete pre-engagement authorization sequence is as follows:



13 CONCLUSION

The regulatory framework for AI in healthcare is not missing. It is not ambiguous. It is not waiting for Congress to act or for HHS to issue a new rulemaking. It exists. It has existed for over two decades. It was designed to be technology-neutral precisely because the drafters understood that technology would evolve in ways they could not predict and that the principles of privacy, minimum necessary use, purpose limitation, and individual rights would remain constant regardless of the processing modality.

The industry's task is not to create new guardrails. It is to apply the ones already in place with the rigor that the sensitivity of the data and the stakes of the decisions demand.

Every AI initiative that fails in healthcare fails for the same reason: the authorization that should have preceded the data use was deferred, assumed, or ignored. Technology worked. The governance did not. And the cost of governance failure in dollars, in timelines, in organizational credibility, and in the erosion of the trust that patients and members place in the institutions that hold their most sensitive information is always greater than the cost of doing it correctly from the start.

"The industry does not need new guardrails. It needs to apply the ones already in place — with the rigor that the sensitivity of the data and the stakes of the decisions demand."