The USB-C Port for Healthcare AI: Wh MCP Is the Protocol That Actually Matters Right Now

MAR 14, 2026 • PAID

6 1 Share

Table of Contents

Abstract

- MCP, open-sourced by Anthropic in late 2024 and donated to the Linux Found in Dec 2025, solves the classic M x N integration problem for AI agents connecti enterprise systems

- Athenahealth announced Aug 2025 an industry-first MCP server on athenaOne framing it as the connective tissue for their AI-native platform serving 160,000+ providers

- The protocol dramatically lowers the cost of building agentic health tech, but introduces serious HIPAA/PHI risk surface that founders and investors cannot i

- FHIR + MCP is the architectural stack that will define the next generation of c workflow tools

- Investment implications span ambient documentation, prior auth automation, clinical decision support, and interop infrastructure

What MCP Actually Is (and Why Everyo Keeps Explaining It Badly)

There's a particular flavor of tech explanation that shows up in healthcare confe where someone puts up a slide with a bunch of boxes and arrows and calls it "interoperability." MCP has started getting that treatment. So let's skip the slide version.

MCP, or Model Context Protocol, was originally released by Anthropic in late 2(an open standard, and then donated to the Linux Foundation's Agentic AI Found in December 2025. The spec has since attracted formal backing from OpenAI, G DeepMind, Microsoft, and AWS, which is about as close as you get to a consensu standard in AI infrastructure. As of mid-2025, there were reportedly over 5,000 MCP servers listed in the Glama MCP Server Directory, with more than 115 production-grade vendor implementations.

The core problem MCP solves is what computer scientists call the M x N integra problem. If you have five AI models and five enterprise systems, you don't have t integrations, you have twenty-five. Each model needs custom glue code for each system. Every time the model updates or the API version changes, something bre Multiply that across a health system's tech stack, which might include a primary a PACS system, a lab information system, a scheduling platform, and several reve cycle tools, and the engineering cost of connecting AI agents to real clinical data enormous. MCP flattens that matrix. Instead of point-to-point custom integratic every model plugs into MCP and every system plugs into MCP. The analogy that stuck, because it's genuinely accurate, is USB-C. One port, multiple devices, standardized handshake.

Technically, MCP operates over a lightweight JSON-RPC layer and defines three building blocks for how AI agents interact with external systems: action tools (th the agent can do), read-only resources (data the agent can pull), and reusable pro templates. The agent doesn't need to know the underlying schema of your EHR o your billing system. It expresses a need, the MCP server handles the translation, applies access controls, and returns a structured response. The host layer can be to encrypt in transit, log every call, and enforce least-privilege access. Done righ actually a more auditable architecture than a lot of the bespoke integrations cur running in production across health systems.

The reason this matters specifically in healthcare is that health data is probably most structurally complex regulated data domain that exists. FHIR, HL7, DICOI LOINC, SNOMED, RxNorm, ICD-10, prior auth formats, payer-specific EDI, sta level HIE feeds, the list goes on. Before MCP, connecting an LLM to even a subs this required domain-specific engineering that only a handful of teams really go right. What MCP does is provide the standardized discovery and communicatior that makes it possible for a well-built agent to navigate that complexity without custom hardcoding. When paired with FHIR R4, which is the current gold stand for structured clinical data exchange, you start to get a stack that can actually ex the full context of a patient encounter to an AI model in a controlled, auditable

The Healthcare Interoperability Proble MCP Was Born to Solve

To understand why MCP landed with such traction in health tech specifically, yo have to appreciate how bad the status quo is. Not bad in a generic "healthcare is behind technology" sense, but structurally, mechanically broken in ways that ha resisted decades of regulatory pressure and investment.

The 21st Century Cures Act and its ONC information blocking rules were suppo to liberate clinical data. TEFCA was supposed to create the trusted exchange framework that made nationwide interoperability real. FHIR mandates on payer providers under CMS interoperability rules were supposed to mean that, by now

patient's data would flow reasonably well between their primary care doc, their specialist, their hospital, and their payer. The reality is that compliance happene the minimum viable level. Most EHR vendors built FHIR endpoints that technic pass certification tests but return incomplete records, require friction-heavy OA flows, and don't actually surface the clinical context that makes the data useful. Athenahealth, to their credit, was the first company to implement TEFCA acros eligible customers at scale, reportedly connecting over 100,000 providers to the national exchange framework. That's a real differentiator in the ambulatory mar and provides meaningful context for why they'd be first to layer MCP on top of i

The deeper problem is that even when data moves between systems, it moves as documents or discrete fields stripped of context. A CCD (Continuity of Care Document) arriving in an EHR inbox might contain a patient's med list and prob list, but none of the clinical reasoning behind it. Medications get duplicated. Al conflict. Lab trends that span multiple care settings don't reconcile automaticall Clinicians end up being the integration layer, spending cognitive energy resolvir data conflicts instead of doing clinical work. Studies consistently show that phys spend somewhere between one-third and one-half of their working hours on documentation and administrative tasks. At 160,000+ providers on the athenaOr platform, even small efficiency gains multiply fast.

What MCP changes in this context is that it makes it possible for an AI agent to just receive a document dump from a FHIR endpoint but to interact with clinica systems dynamically. An agent can request the patient's most recent hemoglobin then conditionally query the medication history for the relevant diabetic agents, pull the care gap data for that patient from the quality program registry, and the surface a synthesized clinical summary at the moment of the encounter, all throu standardized protocol layer without the physician opening multiple tabs or port That's not hypothetical. The architecture to do this exists now. The question is v builds the MCP servers, who governs the data access, and who is accountable wh something goes wrong.

Athena's Big Bet: First Mover in the EH Space

In August 2025, athenahealth announced what they described as an industry-firs MCP server pilot on athenaOne platform APIs, framed as the technical foundati what their CPO Paul Brient called "breaking down the walled gardens that have constrained independent practices for more than a decade." CEO Bob Segert wa characteristically direct about the strategic logic, explaining that the MCP serve would allow both athena's own customers and their network of over 500 API-connected partners to build their own agents with full access to the underlying d those agents need.

Read that again, because it's worth unpacking. Athena is not just building their AI features on top of MCP. They are explicitly enabling third-party developers a partners to build agentic AI capabilities that tap into athenaOne data via the MC layer. That is a platform bet, not just a feature bet. It's the difference between building an app and building an app store. For a company that has historically competed heavily on its interoperability story and its open API ecosystem, this m makes strategic sense. Their existing marketplace of 500+ partners becomes the distribution network for the next generation of ambient, agentic, and decision-support tools. The MCP server becomes the connective tissue that makes those integrations dramatically cheaper to build and maintain.

The specific features athena announced alongside the MCP pilot fill in the prod picture. Next Generation Document Services uses machine learning to process o billion pages of faxes annually received by practices, automatically labeling clini imaging, and administrative documents and extracting discrete data elements th be trended and reported against for quality programs. Next Generation ChartSy uses AI to reconcile and digest information arriving from national exchange netv placing data directly in provider workflows rather than requiring manual review Intelligent Summaries, which entered alpha testing at time of announcement, ge AI-powered overviews of diverse clinical data including social history and medic history to reduce the time clinicians spend pulling context before an encounter.

Assist, also in alpha, functions as an in-workflow AI assistant for physicians. On ambient side, athena's integrations with Abridge, Suki AI, and iScribe AI cover t voice-to-text clinical documentation layer that has arguably been the fastest-ado AI use case in ambulatory care to date.

The cloud-native architecture of athenaOne matters here in a way that is easy to underestimate. Because the platform was built on a single shared cloud instance rather than the on-premise or federated model that many legacy EHRs use, athe deploy AI feature updates across their entire customer base essentially simultaneously. Epic and Oracle Health have enormous installed bases, but the heterogeneity of their on-premise deployments means AI features roll out slowly unevenly. Athena can, in theory, go from pilot to full network deployment of an enabled feature in weeks rather than quarters. For investors and founders evalua the EHR market, that architectural advantage is real and material.

The Real Risk Surface: HIPAA, PHI, and the Confused Deputy

Here's where most MCP coverage in health tech gets politely hand-wavy. The pr is genuinely exciting infrastructure, and the risks are genuinely serious, and bot things are true at the same time.

The core security problem with MCP in a healthcare context comes from what security researchers call the "confused deputy" problem. An AI agent operating MCP acts on behalf of a user, but it may have access privileges that exceed any individual user's actual authorization. If an agent is tricked, by a prompt injecti patient document, a maliciously crafted fax, or a subtly manipulated clinical not calling a high-privilege function it wasn't supposed to call, it executes that unauthorized action without any apparent malice or intent. In a regulated clinic environment, this is not a theoretical edge case. It's an attack vector that bad act will eventually exploit, and one that security teams need to model before deploy

The audit trail problem is related but distinct. MCP tool calls, in a vanilla implementation, may not be logged in enterprise-grade systems of record. If an

agent queries a LIMS system, modifies a clinical note, or pulls PHI from a payer source via MCP, and that action bypasses the standard user audit logs, you've los traceability that HIPAA, 21 CFR Part 11, and standard BAA frameworks depend For any health tech company building on MCP, this means the compliance architecture can't be bolted on after the fact. It has to be designed into the MCP server layer from the start.

Data exfiltration risk is the third major concern. LLMs that have access to PHI v MCP can, in theory, summarize, synthesize, or transform that data and return it obfuscated formats. Without output sanitization and context-aware filtering, thi real HIPAA exposure. The business associate agreement question alone, specific who is the covered entity's BAA counterparty when an MCP server is sitting bet a clinician's AI agent and an EHR's FHIR endpoint, has not been cleanly resolve across the industry as of this writing.

Several approaches to healthcare-specific MCP compliance are emerging. The Healthcare Model Context Protocol (HMCP) is a healthcare-specific profile of th MCP spec that includes FHIR U.S. Core alignment, terminology normalization a SNOMED, LOINC, and RxNorm, risk scoring to block unsafe requests in real ti and event-based audit trails that timestamp every agent interaction with user co agent ID, and access purpose. Layered encryption and access control that separa PHI data zones from synthetic test data and agent logs is also part of the HMCP Early adopters of HMCP-based implementations report integration cycles meas in days rather than months, which tracks with the broader MCP value propositi

The practical implication for founders is that building an MCP-native health tec product without baking in HIPAA-grade access control, RBAC, OAuth2 with SM on FHIR scoping, and a full audit logging stack is not an option in the regulated space. It's also, notably, a potential competitive moat. The teams that get the compliance architecture right early will be very hard to displace once they are embedded in a clinical workflow, because ripping out a BAA-covered, HIPAA-au AI integration is expensive and operationally risky for health systems.

Where MCP Belongs in Healthcare (and Where It Doesn’t)

The most valuable use cases for MCP in healthcare cluster around four areas, at in the current generation of deployments.

Ambient clinical documentation is the most mature. Ambient scribes that use vo to-text to capture patient-clinician dialogue, retrieve lab results and prior notes FHIR MCP calls, and generate structured SOAP notes in real time are already in production at several health systems. The documentation reduction numbers tha have been cited in pilots, some reporting roughly 40% reductions in after-hours documentation time, are consistent with the general trajectory of ambient AI in ambulatory care. Because the MCP layer handles the data retrieval, the ambient doesn’t need to be pre-trained on a specific EHR’s schema. It can dynamically pu context it needs for each encounter.

Prior authorization is the second high-value use case, and one with a regulatory tailwind that is worth paying attention to. CMS finalized rules in late 2024 that require payers to support FHIR-based prior authorization APIs starting in 2026. creates a standard data pathway that an MCP-integrated agent can use to compil relevant ICD-10 codes, imaging reports, clinical notes, and policy language, subr the prior auth electronically, and track its status, replacing a process that curren consumes enormous amounts of clinical staff time. For independent ambulatory practices, which are athena’s core customer segment, prior auth burden is consis one of the top-cited operational pain points.

Clinical decision support is the third domain, and arguably the one with the hig ceiling and the longest timeline to full deployment. An MCP-native CDS agent c pull real-time labs, medication history, vitals, and care gap data to surface actior clinical insights at the point of care, something that legacy CDS tools based on s rule engines have failed to do well for decades. The architectural shift matters b it allows the CDS logic to reason over the full patient context rather than firing on incomplete data. The open-source MCP-FHIR framework published in 2025, on the SMART Health IT sandbox with FHIR R4, demonstrated role-based clini

summarization for clinicians, caregivers, and patients with genuinely different c windows for each persona. That multi-persona capability is actually more intere than it sounds, because payer AI agents, care coordinator tools, and patient-faci apps can all run on the same underlying data fabric with appropriately scoped a controls.

Revenue cycle and administrative automation is the fourth bucket. Patient-facin assistants handling appointment scheduling, insurance verification, and bill pay queries are projected to save providers roughly 12 billion dollars annually in the by 2027, according to projections circulating in the space. The MCP architectur makes those agents meaningfully more capable because they can access real pati data in real time rather than relying on generic scripts or requiring staff escalati anything nontrivial.

Where MCP does not belong yet, or at least where it demands extreme caution, i any workflow where autonomous AI action, rather than AI-assisted human decis making, could directly affect a clinical outcome. Medication ordering, diagnostic coding that drives treatment pathways, or any loop that writes back to the EHR without human review in the middle is not a safe MCP deployment in the curren state of LLM reliability. The human-in-the-loop principle is not just a regulator preference, it's an engineering requirement until the models prove themselves ir specific clinical domains with real-world outcome data.

Investment Thesis: What This Means fo Founders and Angels

The MCP-native health tech opportunity is real, but it's not evenly distributed a the stack. For angels and seed investors evaluating opportunities, here's how the landscape looks right now.

The infrastructure layer is largely taken. The core MCP spec is open-source and major AI labs are building MCP support natively into their model offerings. Healthcare-specific MCP server implementations, including HMCP-compliant frameworks, open-source FHIR MCP servers from companies like Momentum a

WSO2, and vendor implementations from players like Innovaccer, are proliferati quickly. The winners in infrastructure will be the companies that get deeply embedded as the compliance and governance layer, specifically the BAA-covered HIPAA-audited MCP host layer that enterprise health systems can actually depl That's a real business but it's also getting crowded fast.

The application layer, on the other hand, is wide open for domain-specific workf The ambient documentation space is crowded with well-funded players, Abridge Suki, and others, so greenfield there is limited. But the intersection of MCP-ena data access with domain-specific clinical workflows, think oncology decision sup that synthesizes genomics data with real-time trial enrollment feeds, or value-ba care population health tools that combine EHR data with social determinants an claims, is largely unbuilt. These are workflows that have historically been stymie the cost of data integration. MCP dramatically lowers that cost. For a founder w deep clinical domain expertise, the ability to build an agentic tool that reads fro FHIR MCP endpoint on athenaOne, Oracle Health, or Epic's platform without building bespoke connectors is a meaningful change to the build vs. buy math.

The partnership and channel dynamics around athena's MCP announcement des particular attention. Athena has explicitly stated they want their 500-plus API-connected marketplace partners to build agents on the MCP layer. That is a structured distribution opportunity for the right startup. Getting into athena's partner ecosystem has historically required navigating their marketplace progra meeting technical certification requirements. But for a company that builds a hi value, MCP-native workflow tool for ambulatory practices, that channel is now explicitly open and officially encouraged at the CEO level. That's not a trivial commercial signal.

The regulatory environment is broadly supportive in ways that weren't true two ago. The TEFCA implementation push, the FHIR prior auth mandates, the Whit House executive order on health data sharing, and the generally favorable postur both the current FDA and ONC on AI in clinical workflows, all create a policy backdrop that reduces friction for MCP-native deployments. That doesn't mean compliance work is easy, it most definitely isn't, but the direction of regulatory t

is clearly toward more open, standardized data access, which is exactly what MC enables.

For angels specifically, the risk profile of MCP-native health tech startups looks this: the technology risk is lower than it was twelve months ago because the prot is stable, the major EHR vendors are building toward it, and the FHIR data layer increasingly available. The execution risk is where deals will be won or lost. Buil a HIPAA-compliant, BAA-covered, audit-logged MCP implementation that a hea system's security team will actually approve takes real compliance depth. Hiring healthcare attorney who understands AI governance, a security architect with regulated data experience, and clinical advisors who can validate that the workfl actually improves care delivery rather than adding complexity, those are the earl hires that separate the companies that get deployed from the ones that get stuck pilot purgatory.

The MCP moment in healthcare is real. The question isn't whether the protocol matter, it's whether the right founders with the right domain expertise and compliance architecture will get to scale before the EHR vendors, who now have same protocol, build the features themselves. Athena's move makes that competi clock more legible. When a major EHR vendor opens their platform and says "bu your agents here," the window for independent startups to capture the most valu workflow positions gets shorter with every quarter. The time to build is now, bu building it right is not optional.

6 Likes • 1 Restack

Previous

Nex

Discussion about this post

Comments Restacks

Write a comment...

© 2026 Thoughts on Healthcare · Privacy • Terms • Collection notice
Substack is the home for great culture