

RURAL DATA INFRASTRUCTURE IN WEEKS, NOT QUARTERS

Hitting Your RHTP F.2 Milestones Before the 2027 Deadline

Executive Summary

You have less than 17 months to obligate your RHTP Year 1 dollars. Every year after, CMS recalculates your workload-fund allocation against a technical score that includes Factor F.2 (Data Infrastructure). The state Health Information Exchange (HIE) is running. Reports it feeds are still written by hand. The bottleneck is translating what each field means across four different EHRs. Automate that, and rural data infrastructure runs in weeks instead of quarters.

- **Transport is solved. Semantics is not.** HIEs and Qualified Health Information Networks (QHINs) deliver C-CDA and Fast Healthcare Interoperability Resources (FHIR) bundles between facilities. They do not define what “A1c” or “PHQ-9” means when the bundle lands across four different EHRs and a pharmacy system.
- **There is no one left in the building to hand-map this.** Forty percent of local health departments have no dedicated informatics staff (NACCHO, 2024). HRSA books Uniform Data System (UDS) reporting alone at 238 hours per FQHC per year, about six full-time weeks. A 25-FQHC network carries nearly three full-time analysts of UDS work it does not have on staff.
- **The scoring clock runs every year.** CMS recalculates your workload-fund technical score on your annual report from FY26 through FY30. What you miss in Year 1 shrinks your allocation in Years 2 through 5.

With the semantic layer automated, federal reporting stops being a fire drill, new facilities onboard in days, and your F.2 report is assembled from schemas you already own. MTN Data Foundry works on metadata, not patient data, and sits between your HIE and your analytics platform. Thirty days on three facilities, with no PHI leaving the building, is enough for your state to decide.

The record moves through the HIE. The meaning does not travel with it.

Intended Audience (Executive Edition)

Target Audience: State Health IT Directors and CIOs, Medicaid Directors, Public Health Officers, Office of Rural Health Directors, DoH leadership.

Program Focus: Year 1 RHTP obligation (Sept 30, 2027) and annual Technical Factor F.2 score recalculation.

Structural Constraint: 5% cap on EHR replacement; Rural Tech Catalyst Fund up to 10% / \$20M per period.

Posture: Metadata-only semantic layer, downstream of HIE/QHIN, upstream of analytics. No PHI crosses boundaries.

Closing the Semantic Gap in Rural Health Data

A state rural health network is not one organization with one IT department. It is a constellation of independent facilities: Critical Access Hospitals (CAHs), federally qualified health centers (FQHCs), rural health clinics (RHCs), community mental health centers, community pharmacies, tribal health systems, and county public health departments. Each runs a different EHR, maintains a different data dictionary, documents care in a different idiom, and reports to a different set of federal and state programs. No state-level platform can ignore the long tail. TruBridge alone serves more than 1,500 rural facilities; Meditech Expanse dominates rural inpatient; Epic Community Connect, athenahealth, eClinicalWorks, NextGen, Azalea, and Juno split rural ambulatory. A typical state rural network spans four to six EHR vendors, while most urban integrated delivery networks run one.

States need unified visibility across this network for three purposes: population health measurement, care coordination, and federal reporting (UDS, the Medicare Beneficiary Quality Improvement Project (MBQIP), Transformed Medicaid Statistical Information System (T-MSIS), electronic Clinical Quality Measures (eCQMs), electronic case reporting (eCR)). What they encounter is fragmentation. HRSA's Paperwork Reduction Act estimate puts UDS at 238 hours per health center per year (OMB Control 0915-0193). NACCHO's 2024 profile reports that 40 percent of local health departments have no dedicated informatics staff, and 66 percent cite funding as the primary barrier (NACCHO, 2024). The Office of the National Coordinator's (ONC) most recent hospital brief puts it plainly: 71 percent of hospitals report routine access to necessary clinical information from outside providers, but fewer than half (42 percent) say clinicians routinely use it (ONC, 2023).

This is not a technology adoption problem. It is the structural gap between transport (moving data between organizations) and semantics (making the data mean the same thing once it arrives). It is the exact gap Technical Factor F.2 asks states to close.

A CLINICAL SCENARIO: CHRONIC DISEASE ACROSS INDEPENDENT RURAL PROVIDERS

Consider a common patient in a rural county, receiving care across four independent organizations: a primary care provider at an FQHC (diabetes management, labs, medications), a community mental health provider (depression screening and therapy, often in a 42 CFR Part 2 context), a community pharmacy (dispensing and refill adherence), and a Critical Access Hospital emergency department (acute episodes, labs, discharge summaries).

In an integrated delivery system, a quality analyst runs a report. In a rural network, the report is a manual chart review reconstructed from four different EHRs, one pharmacy dispensing system, and whatever the local public health department can extract for electronic case reporting.

Each source system uses its own patient identifier, its own clinical terminology, and its own documentation convention. The record traverses the state HIE or TEFCA QHIN without difficulty. The meaning does not travel with it.

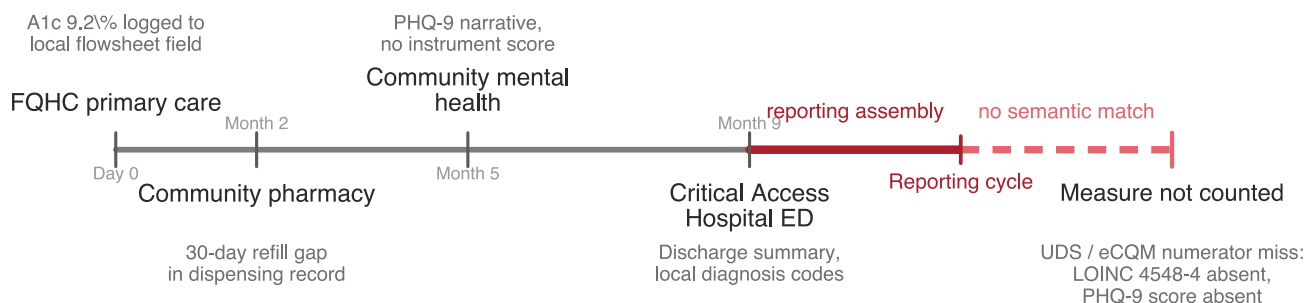
Where it breaks

Two concrete mismatches, both drawn from common rural workflows:

What the source documents	What the reporting specification requires	Result
FQHC: "A1c 9.2%, drawn 3/15"	UDS Table 7: LOINC 4548-4 with result and date in the measurement period	Result exists but is not coded as required; excluded from numerator
Behavioral health provider: narrative depression assessment	eCQM CMS2v13: PHQ-9 score 10 or greater with documented follow-up plan	Narrative without instrument score; measure not met even when follow-up care occurred

The Medicaid and CHIP Payment and Access Commission's 2022 analysis of T-MSIS rated many states' race and ethnicity data "high concern" or "unusable" (MACPAC, 2022). A Journal of the American Medical Informatics Association (JAMIA) study of 1,246 hospitals found that CPSI-based hospitals scored worse than Epic-based hospitals on 8 of 11 process measures even after adjustment (Everson and others, 2022), which is

One patient. Four independent rural facilities. One year. Measure not counted.



The record moved through the HIE. The meaning did not travel with it.

Figure 1. One rural patient’s chronic-disease care over twelve months traverses four independent facilities: an FQHC, a community pharmacy, a community mental health provider, and a Critical Access Hospital emergency department. Each encounter is documented in its own idiom and code system. When the annual reporting cycle arrives, the record has moved through the state HIE without difficulty but key elements (LOINC-coded A1c, PHQ-9 instrument score) are absent in the form the specification requires, and the measure is not counted. Illustrative scenario, not a claimed MTN result.

partly a measurement artifact of vendor heterogeneity rather than a care-quality difference. The measure “not counting” does not mean the care did not happen. It means the record and the specification do not share a language.

Why this problem persists

Transport solves a different problem. State HIEs, TEFCA QHINs (11 designated as of November 2025), Carequality, and CommonWell are built for transport and identity matching. HIE legal agreements typically prohibit altering message content in flight, so receiving systems get payloads in whatever non-standard codes the sender happened to use.

Code-set heterogeneity is structural. International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM), Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT), Logical Observation Identifiers Names and Codes (LOINC), RxNorm, Current Procedural Terminology (CPT), and Healthcare Common Procedure Coding System (HCPCS) all coexist with local proprietary codes. Only about 40 percent of HIE organizations routinely send data aligned to the United States Core Data for Interoperability (USCDI) specification (ONC, 2023).

Specifications drift every year. Annual eCQM value-set updates, annual UDS Program Assistance Letter changes, quarterly ICD-10 updates, biannual SNOMED and LOINC releases, and T-MSIS specification revisions produce continuous downstream breakage. The 2026 UDS restructure alone eliminates Table 4 and redesigns Tables 5, 6A, and 8A; HRSA has called it one of the largest restructurings in decades. Silent breakage surfaces as reporting failure weeks later.

Rural labor markets cannot absorb this. Forty percent of local health departments have no informatics staff; the Council of State and Territorial Epidemiologists estimates that reaching target need would require a 656 percent increase in mental-health epidemiologists alone (CSTE, 2021). The UNC Sheps Center tracks 338 additional rural hospitals at risk of closure on top of the 195-plus that have already closed since 2005 (UNC Sheps Center, 2025). Any solution that requires hiring rural informaticists is not a solution.

MTN Data Foundry: How It Works

MTN Data Foundry automates schema mapping and keeps mappings current as specs change. It turns mismatched source schemas into the formats your reporting programs require. It is not an HIE, not an

EHR, not an analytics platform, and not a data lake. It is the semantic meaning that makes the data your transport layer already moves usable when it arrives.

Operates on metadata, not patient data. The engine works at the schema level: field names, data types, API documentation, internal data dictionaries. It does not require access to specific patient records, which means no PHI crosses facility boundaries. For a State Health IT Director whose top priority is cybersecurity posture, and whose top fear is a ransomware event at a CAH subrecipient, this is a structural rather than policy guarantee.

Automated mapping with confidence scoring. The platform reads database schemas and the documentation that describes them to identify what each field means and how fields relate across facilities and to target reporting schemas (UDS, MBQIP, T-MSIS, eCQM value sets). High-confidence mappings are accepted automatically; low-confidence mappings are flagged for review with context: the source field, candidate concepts, alternative interpretations, and reasoning behind each score.

Schema drift detection. When an upstream system is upgraded, a field is renamed, or a new federal specification is released, the platform detects the deviation, proposes an updated mapping, and alerts the team. Drift detection is what protects the annual RHTP technical-score recalculation without hiring new informaticists.

Rapid facility onboarding. Because onboarding operates on metadata rather than patient data, each new facility joins against an established semantic layer rather than triggering a fresh months-long integration project. This compresses timelines from months to days and makes voluntary participation by independent rural facilities realistic inside the Year 1 obligation window.

WHERE IT FITS IN THE STATE STACK

Data Foundry is an explicit complement to the investments a state has already made or is about to make. It sits downstream of the transport layer and upstream of the analytics layer.

Layer	Examples	What it does	Data Foundry posture
Transport and identity	State HIE, TEFCA QHIN, Carequality, CommonWell	Moves records between organizations; identity matching	Complement; Data Foundry receives payloads after transport
Semantic normalization	MTN Data Foundry	Normalizes what the data means across vendors, code sets, and reporting specifications	This paper
Analytics and reporting	Innovaccer, Arcadia, Azara, Health Catalyst; Snowflake, Databricks, Microsoft Fabric	Population health analytics, quality reporting, data warehousing	Complement; Data Foundry supplies normalized input

State HIEs (for example Big Sky Care Connect, CRISP, MiHIN, Contexture, Connie, IHIE, SHIN-NY, Health-connect Texas) are political incumbents. The correct posture is never to replace them. The correct posture is to make the data they move usable once it arrives. The same is true of the analytics vendors states have already chosen: Innovaccer, Arcadia, Azara, Health Catalyst, and the cloud data platforms need normalized input, and that input is what Data Foundry produces.

A PRACTICAL DEPLOYMENT PATH

Phase	Timeline	What Happens
Schema ingestion	Days 1 to 5	Ingest schemas, data dictionaries, and API documentation from participating facilities. No data migration. No PHI required.
Automated annotation and mapping	Week 1 to 2	Automated concept mapping across facility sources and target reporting schemas; confidence scoring flags what needs review.
Targeted review	Week 2 to 3	State staff and clinical subject matter experts review flagged mappings and validate reporting semantics.
Ongoing maintenance	Continuous	Schema drift detection and re-mapping as systems and reporting specifications evolve.

Traditional integration scales linearly: each new facility requires a new engineering team, a new timeline, and a new set of brittle pipelines. Data Foundry scales sublinearly. Once the semantic layer exists, each new facility maps against established concepts.

WHY THE BURDEN COMPOUNDS

The 238-hour HRSA UDS anchor is not a speculative number. It is the federal government’s own Paperwork Reduction Act estimate, published under OMB Control 0915-0193 and binding on HRSA reporting guidance. A state with 25 FQHCs is already carrying roughly 6,000 hours per year of UDS reporting alone at that baseline, before MBQIP quarterly abstraction, T-MSIS submissions, or eCR onboarding. At a 2,080-hour analyst-year, that is nearly three full-time analysts of manual reporting a state does not have on staff.

The snapshot understates the problem. Traditional integration grows with every new facility onboarded

(because schemas drift relative to each other) and every annual specification change (eCQM, UDS, SNOMED, LOINC, T-MSIS). Data Foundry’s mapping-artifact architecture grows with neither. The effective economic question for a state is not “how much does manual reporting cost today,” but “how much faster does it grow than the informatics labor market can absorb.” The NACCHO 40-percent-of-LHDs-with-no-informaticists number is the ceiling the labor market puts on the traditional curve. The 2026 UDS restructure is one of many drift events the curve has to absorb. Any portion of that compounding burden converted from manual reconciliation to automated semantic mapping pays back the cost of the platform, and it does so every reporting cycle and every specification change.

TARGETED VENDOR QUESTIONS

“Are you another data-lake vendor?” No. Data Foundry sits upstream of the analytics stack (Snowflake, Databricks, Microsoft Fabric, Health Catalyst, Arcadia, Innovaccer). It provides the semantic mapping layer those tools do not automate and that customers otherwise build with consulting teams.

“Is this another vendor lock-in risk?” Data Foundry produces mapping artifacts (metadata and transformation logic) that the state owns and can export. The state retains the right to run the same mappings in another environment if it changes vendors later.

“Does this require broadband at every facility?” Because Data Foundry operates on metadata rather than streaming PHI, its bandwidth requirements are a fraction of transport-based integration. This matters in low-broadband rural settings.

“How does this work with our existing HIE or QHIN?” Complement, not replace. The HIE or QHIN gets the C-CDA or FHIR bundle from point A to point B. Data Foundry normalizes what the fields mean once the bundle arrives.

“What about HIPAA and 42 CFR Part 2?” Metadata-only architecture means Data Foundry is not a traditional PHI processor. The platform never sees patient records, which structurally eases Part 2 segregation and re-disclosure tracking. Business Associate Agreements, SOC 2 Type 2, and HITRUST are standard.

“Can we sole-source this via the Rural Tech Catalyst

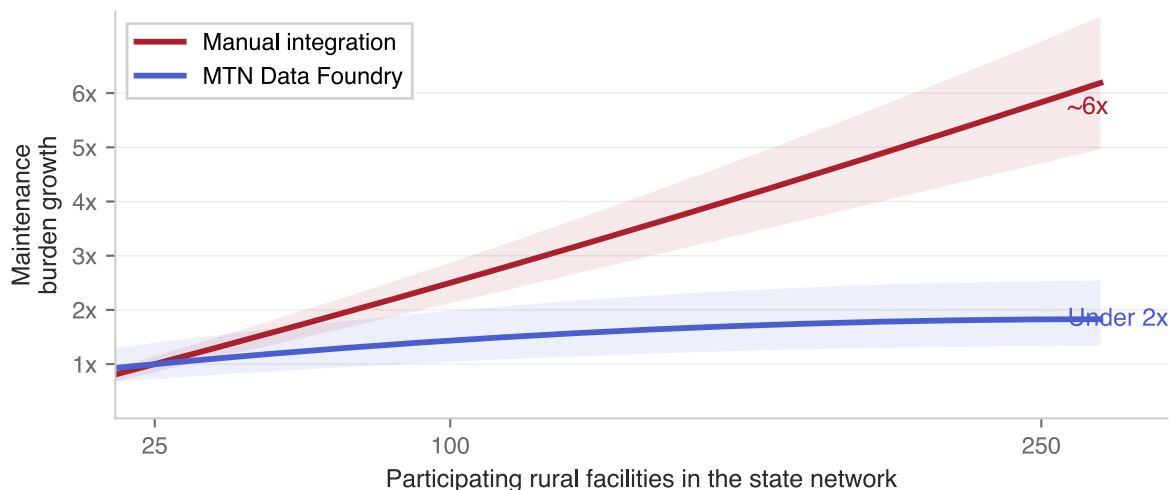


Figure 2. Maintenance burden compounds with every new facility onboarded and every annual specification change. Traditional manual integration grows superlinearly; MTN Data Foundry’s shared-semantic-layer approach stays nearly flat. Indexed to the smallest state network (25 facilities = 1x) so the figure reads as growth shape, not absolute hours. Shaded regions indicate ranges based on network complexity and the volatility of annual eCQM, UDS, SNOMED, LOINC, and T-MSIS spec changes. Illustrative growth, not a claimed MTN result.

Fund?™ Yes. Up to 10 percent of the award (or \$20 million per budget period, whichever is less) is designated for catalytic technology investments. NASPO ValuePoint and state cooperative master contracts are the faster alternative.

Where We Are

MTN’s founding team has built semantic mapping systems before, across healthcare data infrastructure, machine learning, and clinical informatics. Active research at the Medical Machine Intelligence Lab at the University of Utah extends the platform toward clinical decision support.

Whether you are standing up a Rural Health Data Hub, building a health data utility, or connecting a statewide network of independent rural facilities, the RHTP questions are the same. How many of your participating facilities are technically connected but semantically incompatible? What fraction of your UDS, MBQIP, and T-MSIS reporting is still manual reconciliation? What is your plan to protect the annual Technical Factor F.2 score recalculation in a rural labor market where 40 percent of local health departments have no informatics staff?

PROCUREMENT PATHWAY

Data Foundry fits NOFO Category F (significant information technology advances) and supports Technical Factor F.2 (Data Infrastructure) progress. The Rural Tech Catalyst Fund, up to 10 percent of the award or \$20 million per budget period (whichever is less), is explicitly designated for catalytic technology investments. NASPO ValuePoint and state cooperative master contracts are the preferred pre-competed vehicles for shortcutting the RFP cycle inside the Year 1 obligation window.

THIRTY-DAY PROOF OF VALUE

We offer a thirty-day proof of value scoped to three participating facilities in your network. It maps the semantic gap between what those facilities document and what your target reporting specifications (UDS, MBQIP, T-MSIS, eCQM) require, quantifies the hours of manual reconciliation being carried today, and demonstrates the drift-resilient mapping pipeline against a real specification change. No PHI is required to run it. To schedule, contact warren@themtn.ai.

TECHNICAL LEADERSHIP

Our team’s work has been published in Nature journals, PNAS, JMIR, Chest, PLoS Computational Biology, The

Royal Society, and other leading venues.



Warren Pettine, MD, Co-Founder and CEO. Assistant Professor at the University of Utah where he leads the Medical Machine Intelligence (M²Int) Lab. Trained in machine learning research at Harvard, Stanford, NYU, and Yale. Prior health policy experience in the U.S. Congress and service on the University of Utah Institutional Review Board ground MTN's approach in policy and regulatory expertise.



Matthias Christenson, PhD, AI Architect. Investigator with the M²Int Lab. PhD and postdoctoral research at Columbia University in computational ML, with prior industry experience as a Deep Learning Research Engineer at DeepLife training foundational models on genomic and biometric data. Leads MTN's technical architecture design and data model development.



Brian Locke, MD, MSCI, Clinical AI Lead. Investigator with the M²Int Lab. Active ICU physician and Assistant Professor at Intermountain Healthcare, bringing firsthand understanding of clinical workflows across academic medical centers and integrated delivery networks. Provides the methodological rigor for the clinical and operational implications of MTN's technology.



Samuel Wecker, Lead Systems Engineer. Over twelve years building and scaling production software, including as a founding engineer at a startup that grew to a billion-dollar platform. Specializes in unifying disparate systems and data sources at scale. Leads Data Foundry's core platform development.

Contact:

Warren Pettine, MD, CEO of MTN

warren@themtn.ai

<https://www.themtn.ai>