

RADAR ADRES

Información del proyecto

MVP 3 meses: replicación “CoMET-scale” para eventos de salud + SIIFA / FEV-RIPS



Información general

Nombre:	RADAR ADRES — Observatorio de patrones y desviaciones del gasto en salud
Definición	Iniciativa institucional para integrar fuentes clave, normalizar semántica y publicar evidencia agregada y trazable que permita identificar patrones históricos y señales de desviación en el uso de recursos del sistema de salud
Objetivo:	Habilitar toma de decisiones y priorización de acciones (ADRES–MinSalud–Supersalud–INS) mediante indicadores agregados, comparabilidad y trazabilidad , con foco inicial en desviación financiera basada en patrones históricos .
Alcance:	MVP (4–5 meses): modelo base de eventos + semántica + tableros institucionales + señales de desviación financiera (explicables). Fuentes MVP/H1: FEV-RIPS/RIPS, BDUA, RETHUS, REPS, ECAT, MIPRES, SISMED, catálogos CUPS/CUMS y SIIFA* (según disponibilidad y acuerdos). Apertura progresiva: publicación de open data agregado + metodología.
Limitaciones y riesgos	Disponibilidad/tiempos de acceso a fuentes (p. ej., SIIFA). Calidad, completitud y consistencia de datos entre fuentes. Restricción de publicación: sin valores unitarios ni bandas (riesgo de interpretación si el mensaje no es claro). Gobernanza y versionado (modelo, semántica, reglas) como condición para credibilidad pública.
Responsables	Sponsor institucional: Dr. Felix Martinez Liderazgo funcional (negocio/política): Daniel Garavito Liderazgo técnico/arquitectura y datos: Dr. Carlos Arango, Dr. Carlos Augusto Lopez, Dr. Alonso Verdugo, Ing. Santiago Perez Gobernanza y publicación: PM
Áreas involucradas	ADRES · Ministerio de Salud · Superintendencia Nacional de Salud · Instituto Nacional de Salud · (ecosistema: academia/industria mediante open data agregado, progresivo)
Fecha de inicio y fin:	Inicio: Nota Conceptual aprobada + acuerdos de acceso a fuentes del MVP + definición de gobernanza mínima. Fin (MVP): arquitectura operativa + datasets certificados agregados + tableros institucionales + señales de desviación financiera + lineamientos de publicación.
Hitos	1. Acceso a fuentes MVP y acuerdos de intercambio 2. Modelo canónico de eventos v1 + diccionario de datos 3. Semántica CUPS/CUMS v1 (equivalencias/vigencias) 4. Indicadores agregados + linaje/auditoría v1 5. Señales de desviación financiera v1 (explicables, contra patrones históricos) 6. Tableros institucionales v1 + política de publicación (open data agregado)
Enlaces de interés:	Nota Conceptual RADAR ADRES (versión vigente) Diagrama de arquitectura (Draw.io) Lineamientos oficiales de FEV-RIPS (documento MinSalud – versión vigente) Plan Estratégico Institucional ADRES 2023–2026 (versión vigente) Resolución/Marco normativo SIIFA (versión vigente) <i>Incluir los enlaces de notebook, códigos, carpeta Sharepoint institucional, entre otros</i>

Encaje institucional

- **Estrategia:** RADAR ADRES se alinea con el Objetivo Estratégico 2 (Transformación Digital y Arquitectura Empresarial) orientado a modernizar la entidad, mejorar la transparencia y publicidad de la información y fortalecer el seguimiento de los recursos de la salud. (PEI ADRES 2023–2026, p. 38).
- **Gestión pública / enfoque CONPES:** RADAR ADRES se propone como instrumento habilitante de política pública para evidencia comparable, trazable y auditable, que soporte priorización de acciones y mejoras de eficiencia en el uso de recursos (sin reemplazar decisiones regulatorias).
- **Capacidad y articulación interinstitucional:** requiere coordinación operativa entre ADRES, MinSalud, Supersalud e INS, con soporte de áreas de datos/analítica, TI, jurídica, comunicaciones y control interno, y un esquema de gobernanza de publicación (roles, versionado, trazabilidad).
- **Normativa y salvaguardas:** se enmarca en las competencias y lineamientos institucionales vigentes y exige cumplimiento normativo, seguridad y privacidad, control de accesos, y auditoría/linaje. La publicación se limita a indicadores agregados y patrones históricos, sin precios/tarifas unitarias ni bandas.

ADRES. (2025). *Plan Estratégico Institucional ADRES 2023–2026* (p. 38 y p. 13).

Antecedentes del proyecto

¿Qué pasaba antes de este proyecto?

La evidencia relevante para entender **cómo se usan los recursos en salud** existe, pero está **distribuida en múltiples fuentes** (eventos clínicos, afiliación, prestadores, talento humano, medicamentos/tecnologías). Esa fragmentación limita la **comparabilidad** y la **trazabilidad** necesarias para explicar patrones, identificar desviaciones y priorizar acciones interinstitucionales con consistencia.

¿Qué detonó la necesidad de este proyecto hoy?

Dos hitos recientes vuelven viable y urgente organizar la evidencia de manera trazable:

Los lineamientos oficiales para usar **RIPS como soporte de la Factura Electrónica de Venta (FEV) en salud**, que fortalecen la estandarización y la trazabilidad del evento.

La definición de la **estructura del SIIFA**, que formaliza un marco de integración financiera-asistencial.

En este contexto, RADAR ADRES se plantea como una **respuesta institucional lógica**: articular fuentes, normalizar semántica y habilitar evidencia agregada para la toma de decisiones, **sin publicar precios/tarifas unitarias ni bandas**.

Hechos soporte (insumos habilitantes identificados):

Fuentes habilitantes: FEV-RIPS/RIPS, BDUA, RETHUS, REPS, ECAT, MIPRES, SISMED y catálogos CUPS/CUMS (semántica).

Hito normativo (trazabilidad de eventos): lineamientos oficiales FEV-RIPS.

Hito normativo (integración financiera-asistencial): estructura SIIFA (Res. 00001962 de 2025).

Intentos previos / mandato institucional:

Intentos previos documentados: *[POR COMPLETAR CON SOPORTE: acta / proyecto / informe]*

Mandato institucional: *[POR COMPLETAR CON SOPORTE: lineamiento/PEI/acto interno aplicable]*

Ministerio de Salud y Protección Social. (2025). *Lineamientos para la generación, validación y envío del RIPS como soporte de la Factura Electrónica de Venta (FEV) en salud* (Versión 3.2). Bogotá, D.C.

Ministerio de Salud y Protección Social. (2025). *Resolución 00001962 de 2025: Estructura del Sistema Integral de Información Financiera y Asistencial (SIIFA)*. Bogotá, D.C.

Justificación y objetivos del proyecto

Por qué RADAR ADRES:

La evidencia necesaria para decisiones públicas sobre uso de recursos en salud está **dispersa** en múltiples fuentes; esto limita la **comparabilidad**, la **trazabilidad** y la **explicación** de patrones y desviaciones con alcance interinstitucional.

Existen **hitos habilitantes recientes** que fortalecen la trazabilidad y estandarización del evento en salud (**FEV-RIPS**) y la articulación financiero–asistencial (**SIIFA**), creando una ventana institucional para consolidar evidencia.

RADAR ADRES se orienta a **visibilizar patrones históricos e indicadores agregados**, y a generar **señales explicables de desviación**, con una salvaguarda explícita: **no publicar precios/tarifas unitarias ni bandas**.

Incluye un enfoque de **apertura progresiva** mediante **open data agregado** y documentación metodológica, para fortalecer el ecosistema (academia, industria, control social), sin comprometer restricciones normativas ni de publicación.

Objetivo general:

Habilitar evidencia **agregada, trazable y explicable** sobre el uso de recursos en salud para soportar toma de decisiones y priorización de acciones (ADRES–MinSalud–Supersalud–INS), con **foco inicial en desviación financiera** contra **patrones históricos**.

Objetivos específicos (MVP / H1):

Integrar fuentes habilitantes del MVP y asegurar trazabilidad/linaje de la evidencia (datasets certificados agregados).

Establecer semántica y comparabilidad (catálogos CUPS/CUMS y reglas versionadas de agrupación/unidades comparables).

Implementar tableros institucionales y mecanismos de consulta para stakeholders, basados en indicadores agregados y metodología documentada.

Generar señales de desviación financiera explicables (contra patrones históricos) para priorización institucional, sin divulgar valores unitarios ni bandas.

Definir e implementar la gobernanza mínima: control de accesos, auditoría, versionado y criterios de publicación. *scribir brevemente qué se hará con el proyecto y porqué.*

Referencias:

Ministerio de Salud y Protección Social. (2025). *Lineamientos para la generación, validación y envío del RIPS como soporte de la Factura Electrónica de Venta (FEV) en salud* (Versión 3.2). Bogotá, D.C.

Ministerio de Salud y Protección Social. (2025). *Resolución 00001962 de 2025: Estructura del Sistema Integral de Información Financiera y Asistencial (SIIFA)*. Bogotá, D.C.

ADRES. (2025). *Plan Estratégico Institucional ADRES 2023–2026*. Bogotá, D.C.

Impacto del proyecto

Impacto institucional (ADRES / Estado)

Consolida evidencia **agregada, comparable y trazable** para soportar decisiones y coordinación **ADRES–MinSalud–Supersalud–INS**.

Fortalece funciones de **seguimiento, control y rendición** de cuentas mediante linaje/auditoría y metodología documentada.

Reduce ambigüedad institucional al trabajar sobre un **modelo común de eventos + semántica** (CUPS/CUMS) y reglas versionadas.

Impacto tecnológico (Transformación digital)

Instala una “columna vertebral” de datos: **datasets certificados, catálogo/definiciones oficiales, control de versiones** y criterios de publicación.

Habilita reutilización: tableros, consultas institucionales y open data agregado sin reconstrucciones ad-hoc.

Impacto económico (eficiencia / optimización / prevención)

Habilita **identificar y priorizar** señales de desviación financiera **contra patrones históricos** (explicables), para orientar acciones de revisión y control.

Permite medir oportunidades de eficiencia y mejora operativa **sin publicar precios/tarifas unitarias ni bandas** (salvaguarda explícita).

Impacto al actor final (ciudadanía / ecosistema)

Incrementa la **transparencia** mediante publicación **progresiva de open data** agregado y documentación metodológica.

Facilita **análisis independiente** por academia/industria/control social sobre patrones y desviaciones (agregadas), preservando restricciones normativas.

Indicadores de impacto (medibles, para seguimiento CONPES)

Cobertura de fuentes del MVP integradas y certificadas.

Tiempo de respuesta para generar indicadores/tableros institucionales (línea base vs. RADAR).

de indicadores agregados publicados con metodología y linaje.

de señales de desviación generadas y priorizadas para revisión institucional.

Cronograma de actividades

MVP (H1): Mes 1 a Mes 5

Línea de trabajo (MVP)	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5
1. Gobernanza mínima + acuerdos de acceso (fuentes MVP)					
2. Ingesta + cortes + control de calidad (perfilado/validaciones)					
3. Modelo canónico de eventos v1 (diccionario + llaves + versionado)					
4. Semántica v1 (CUPS/CUMS: equivalencias + vigencias + normalización)					
5. Datasets certificados agregados + linaje/auditoría + catálogo					
6. Señales de desviación financiera v1 (explicables, patrones históricos)					
7. Tableros institucionales v1 (MinSalud/ADRES/Supersalud/INS)					
8. Política de publicación + open data agregado v1 + metodología					

Cronograma de actividades

Salida del MVP (fin Mes 4–5):

- Datasets certificados agregados + catálogo/definiciones + linaje/auditoría
- Tableros institucionales v1
- Señales de desviación financiera v1 (explicables, contra patrones históricos)
- Política de publicación (sin precios/tarifas unitarias ni bandas)

Horizonte 2 (Mes 6–18) — Consolidación y expansión (solo títulos, sin prometer de más)

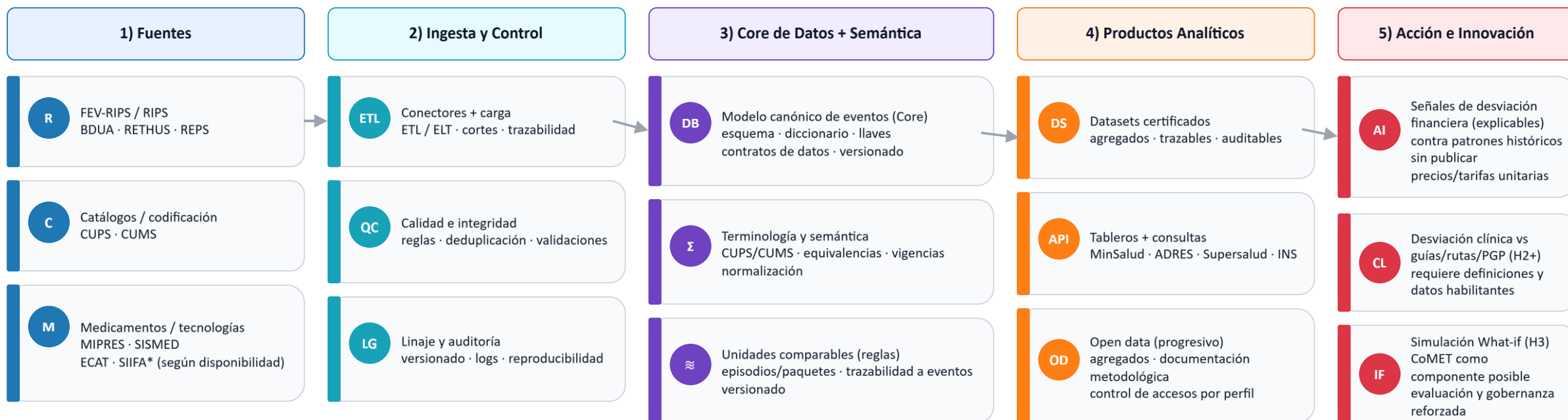
- Ampliación de fuentes (incluye expansión SIIFA y otras según disponibilidad/acuerdos)
- Madurez de calidad, linaje y comparabilidad (reglas versionadas)
- Extensión progresiva a desviación clínica vs guías/rutas/PGP (**H2+**, requiere definiciones habilitantes)
- Mayor apertura de open data agregado y documentación metodológica

Horizonte 3 (Mes 19–36) — Innovación avanzada

- Simulación *what-if* / modelos de eventos (**CoMET como componente posible**)
- Evaluación, monitoreo y gobernanza reforzada para modelos avanzados
- Ampliación de casos de uso y capacidades analíticas según decisiones institucionales

Arquitectura RADAR (Alto Nivel)

Salida pública: indicadores agregados y patrones históricos (sin publicar precios/tarifas unitarias ni bandas)



Gobernanza transversal: seguridad y privacidad · control de accesos · auditoría/linaje · control de versiones · cumplimiento · criterios de publicación (sin precios/tarifas unitarias ni bandas)

* SIIFA: disponibilidad según acuerdos de acceso para el MVP (H1). En H2/H3 se amplía a nuevas fuentes según disponibilidad y acuerdos.

La salud experimenta fuerzas sísmicas

¿**Dónde** se ofrece la atención?

Evolución de la prestación de cuidados

¿**Cómo** se ofrece la atención?

Convergencia de ecosistemas

¿**Quién** recibe la atención?

Transición al Individuo

Personalización y Participación



Preparación de la información y conocimientos accionables

Modernización de la estructura de la cadena de valor

¿Qué es el Modelo CoMET?

Cosmos Medical Event Transformer (Epic / Microsoft, 2025)

Fuente: arXiv:2508.12104v1

La Medicina como Lenguaje

CoMET es un **Modelo Fundacional Generativo** que trata la historia clínica longitudinal del paciente como una secuencia de texto, donde cada evento médico (diagnóstico, medicamento, procedimiento) es una "palabra".

Analogía con ChatGPT



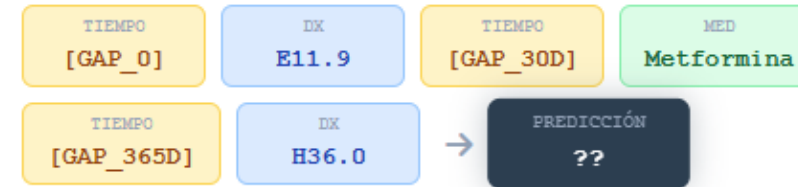
Así como un LLM predice la siguiente **palabra** en una frase, CoMET aprende a predecir el siguiente **evento clínico** en la vida de un paciente.

Objetivo del Modelo

Capturar dependencias temporales complejas a largo plazo (ej. una diabetes diagnosticada hace 10 años) para predecir riesgos futuros sin reglas manuales.

MECANISMO DE "MEDICAL TOKENIZATION"

⚡ POWERED BY GEMINI API



Prueba el "Cerebro" (Simulador CoMET via Gemini):

Diagnóstico Diabetes, Metformina, Visión borrosa

⚡ Predecir

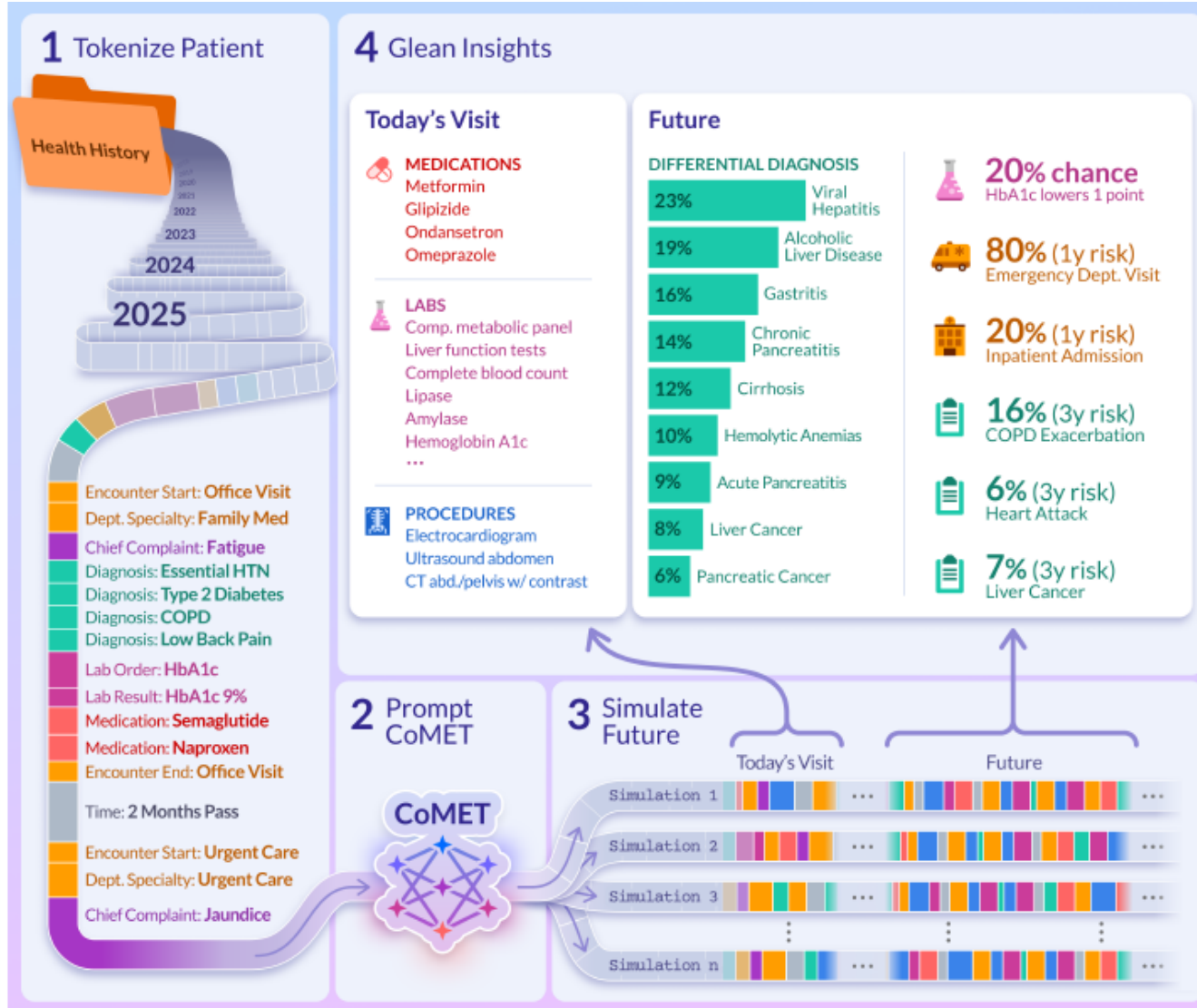
🤖 Esperando historial para predecir siguiente evento...



⚙️ **MOTOR TRANSFORMER**

Procesa patrones ocultos en 115 Billones de eventos clínicos.

¿Qué pasaría si tuviéramos todos los datos del sistema interoperados?



- **Predicción de la ruta de atención integral sin GRDs, diagnostico, manejo y egreso**
- **Transparencia de punta a punta**
- **Evitar dobles o triples gastos**
- **Calidad de Egreso**
- **Gestión de Salud Poblacional**
- **Reduce Fraude al Sistema:** Se reducen las posibilidades para hacerle fraude al sistema.

Fuente: Generative Medical Event Models Improve with Scale, 2025

<https://arxiv.org/abs/2508.12104>

| Cosmos Medical Event Transformer (COMET)

El artículo presenta los modelos **Cosmos Medical Event Transformer (COMET)**, una familia de modelos de inteligencia artificial generativa entrenados en Epic Cosmos, una de las bases de datos de salud desidentificadas más grandes del mundo, que contiene miles de millones de eventos médicos de cientos de millones de pacientes.

El enfoque central del modelo es aprender de las trayectorias de salud longitudinales de los pacientes, representadas como una secuencia cronológica de eventos (diagnósticos, medicamentos, resultados de laboratorio, etc.). A partir del historial de un paciente, COMET puede **simular múltiples futuros posibles**, generando de forma autorregresiva los eventos médicos que podrían ocurrir a continuación.

Los hallazgos clave del estudio son:

1. **Alto Rendimiento sin Entrenamiento Específico (Zero-Shot):** A diferencia de los modelos tradicionales que requieren ser entrenados para cada tarea específica, COMET demostró un rendimiento que iguala o supera a los modelos supervisados en 78 tareas clínicas diversas, incluyendo la predicción de diagnósticos, el pronóstico de enfermedades y la optimización de operaciones hospitalarias, todo ello sin necesidad de reentrenamiento o ajuste fino.
2. **El Rendimiento Mejora con la Escala:** El estudio confirma que, al igual que en los modelos de lenguaje, el rendimiento de COMET mejora de manera predecible a medida que aumentan el tamaño del modelo, la cantidad de datos de entrenamiento y la capacidad de cómputo. Esto sugiere que hay un margen considerable para futuras mejoras a medida que la tecnología y los datos sigan creciendo.
3. **Aplicaciones Operativas:** El modelo demostró una notable precisión en tareas operativas cruciales para los sistemas de salud, como predecir la duración de las estancias hospitalarias y el riesgo de reingreso a 30 días, superando a los modelos de referencia.

En esencia, COMET funciona como un "motor de simulación" para la trayectoria de salud de un paciente, ofreciendo un marco generalizable y flexible para una amplia gama de aplicaciones clínicas y administrativas.

Cosmos Medical Event Transformer (COMET)

Generative Medical Event Models Improve with Scale

Shane Waxler^{*1}, Paul Blazek^{*1}, Davis White^{*1}, Daniel Sneider^{*1},
Kevin Chung¹, Mani Nagarathnam¹, Patrick Williams¹, Hank Voeller¹, Karen Wong¹,
Matthew Swanhorst¹, Sheng Zhang², Naoto Usuyama², Cliff Wong², Tristan Naumann²,
Hoifung Poon², Andrew Loza³, Daniella Meeker^{3, 4}, Seth Hain¹, and Rahul Shah^{†1}

¹Epic Systems

²Microsoft Research

³Yale School of Medicine

⁴Cosmos Governing Council

Abstract

Realizing personalized medicine at scale calls for methods that distill insights from longitudinal patient journeys, which can be viewed as a sequence of medical events. Foundation models pretrained on large-scale medical event data represent a promising direction for scaling real-world evidence generation and generalizing to diverse downstream tasks. Using Epic Cosmos, a dataset with medical events from de-identified longitudinal health records for 16.3 billion encounters over 300 million unique patient records from 310 health systems, we introduce the Cosmos Medical Event Transformer (CoMET) models, a family of decoder-only transformer models pretrained on 118 million patients representing 115 billion discrete medical events (151 billion tokens). We present the largest scaling-law study for medical event data, establishing a methodology for pretraining and revealing power-law scaling relationships for compute, tokens, and model size. Based on this, we pretrained a series of compute-optimal models with up to 1 billion parameters. Conditioned on a patient's real-world history, CoMET autoregressively generates the next medical event, simulating patient health timelines. We studied 78 real-world tasks, including diagnosis prediction, disease prognosis, and healthcare operations. Remarkably for a foundation model with generic pretraining and simulation-based inference, CoMET generally outperformed or matched task-specific supervised models on these tasks, without requiring task-specific fine-tuning or few-shot examples. CoMET's predictive power consistently improves as the model and pretraining scale. Our results show that CoMET, a generative medical event foundation model, can effectively capture complex clinical dynamics, providing an extensible and generalizable framework to support clinical decision-making, streamline healthcare operations, and improve patient outcomes.

1 Introduction

Posibles Beneficios Futuros

Más allá de los resultados ya demostrados, la arquitectura y el enfoque de COMET abren la puerta a una serie de beneficios transformadores para el futuro de la medicina y la gestión de los sistemas de salud:

- **Medicina Preventiva y Proactiva (Razonamiento Contrafáctico):** El artículo sugiere la posibilidad de realizar análisis "what-if" (¿qué pasaría si...?). En el futuro, un médico podría usar el modelo no solo para predecir el riesgo de un evento adverso, sino para simular y comparar los resultados de diferentes intervenciones. Por ejemplo: "¿Cuál es la trayectoria de salud más probable de este paciente si le prescribo el medicamento A, el medicamento B, o si solo recomiendo cambios en el estilo de vida?". Esto transformaría la toma de decisiones de reactiva a proactiva, personalizando la prevención.
- **Optimización de Ensayos Clínicos:** La capacidad de predecir no solo *si* un evento ocurrirá, sino *cuándo*, podría revolucionar el diseño de ensayos clínicos. Se podrían identificar cohortes de pacientes con mayor probabilidad de experimentar un resultado de interés dentro del plazo del estudio, haciendo los ensayos más rápidos, económicos y con mayor probabilidad de éxito.
- **Planificación de Recursos a Nivel de Sistema:** La habilidad del modelo para predecir la utilización de servicios (como ingresos hospitalarios o visitas a urgencias) podría escalarse para la gestión de todo un sistema de salud. Los administradores podrían pronosticar la demanda de camas de UCI, la necesidad de personal o el consumo de medicamentos específicos para los próximos meses, permitiendo una asignación de recursos mucho más eficiente y evitando cuellos de botella.
- **Modelos Especializados por Patología:** Aunque el rendimiento general de COMET es impresionante, el modelo podría ser afinado (*fine-tuned*) con datos de poblaciones específicas para crear versiones expertas. Por ejemplo, un "COMET-Oncología" podría predecir la respuesta a quimioterapias específicas, o un "COMET-Cardiología" podría especializarse en la progresión de la insuficiencia cardíaca, ofreciendo predicciones aún más precisas en áreas de alta complejidad.
- **Integración de Datos Multimodales:** El futuro del modelo pasa por incorporar nuevos tipos de datos, como notas clínicas no estructuradas, imágenes médicas (radiografías, resonancias) y datos genómicos. La integración de esta información crearía una visión verdaderamente holística del paciente, permitiendo al modelo correlacionar un hallazgo genético con una trayectoria de enfermedad o una imagen radiológica con una predicción de resultados a largo plazo, llevando la medicina de precisión a un nuevo nivel.

EL VALOR DEL DATO

De piezas sueltas a decisiones estratégicas

1. DATOS



Piezas sueltas sin conexión.



Recolección

RIPS y BDUA
crudos.

2. CLASIFICADOS



Agrupados por color.



Validación

Limpieza y
consistencia.

3. ORDENADOS



Estructura definida.



Auditoría

Reglas de negocio
Automatización e
Inteligencia Artificial

4. VISUALES



Forma comprensible.



Tableros

Dashboards
Integrados e
Interactivos de
gestión.

5. EXPLICADOS



Historia de valor.



Impacto

Política Pública.
Gestión de Riesgo
por Individuo

Objetivo (3 meses)

Construir un MVP que transforme eventos de prestación y pago (FEV RIPS/SIIFA) en una referencia **tarifaria** y analítica de **transparencia**, usando una metodología tipo CoMET: “secuencia de eventos → modelo generativo → simulación y predicción”.

Por qué ahora

- SIIFA define el canal oficial para registro de contratos, FEV-RIPS, glosas y pagos.
- FEV-RIPS aporta un estándar nacional con validación (CUV) y trazabilidad del flujo.
- ADRES prioriza Transformación Digital e “Información de valor” para decisiones.

Qué entregamos en 12 semanas

- Lakehouse ADRES (Databricks) con pipelines RIPS/FEV + datos maestros.
- Motor “tarifario inteligente”: rangos de referencia por CUPS/tecnología y contexto.
- Asistente normativo (RAG) con citas: SIIFA + FEV-RIPS + manuales.
- Tablero ejecutivo: variación, outliers, glosas, oportunidad de pago, trazabilidad.

Apuesta técnica

- Modelo de eventos (Foundation Model) + LLM para explicación y gobernanza.
- Azure Foundry para orquestación de agentes, evaluación y despliegue.
- FHIR como modelo canónico (cuando aplique) para interoperabilidad.

Decisión requerida: priorizar casos MVP + acceso a datos + cómputo GPU

Lo que nos condiciona (y habilita) para el MVP

SIIFA – canal oficial

- Registro de contratos, facturas, devoluciones, glosas, respuestas y pagos.
- Interoperabilidad y roles; autenticación asociada a SISPRO.

Implicación: trazabilidad end-to-end

Contratos → CUCON

- Registro del acuerdo de voluntades con código único (CUCON).
- CUCON se informa en la FEV para validación en el ecosistema.

Implicación: link contrato ↔ factura

FEV-RIPS (CUV)

- Factura electrónica validada por DIAN + RIPS JSON validado por MinSalud.
- CUV como sello de validación para radicación y trazabilidad.

Implicación: datos confiables para IA

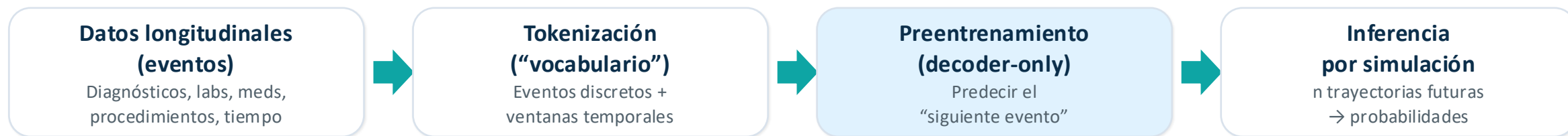
Seguimiento a facturas/pagos

- Devoluciones y glosas (reglas, tiempos y respuestas).
- Registro de pagos y transferencias (incl. giro directo).

Implicación: analítica de oportunidad

Consecuencia para la plataforma tarifaria: el “evento” no es solo clínico (CUPS/CIE10), también contractual-financiero (CUCON, glosa, pago, ajuste).

Idea central: “paciente = secuencia de eventos” → modelo generativo para simular futuros



Qué nos llevamos de CoMET

- Modelos generativos de eventos clínicos escalan con más tokens/modelo (leyes de escala).
- Inferencia “sin fine-tuning” por simulación: 1 modelo → muchas tareas.
- Evaluación multi-tarea: diagnóstico, pronóstico y operaciones; mejora con escala.

Cómo lo “traducimos” a ADRES

- “Paciente” y “prestación” como journeys trazables: RIPS/FEV + pagos/glosas.
- Tokenización basada en CUPS/CIE10/CUM/ATC + atributos (edad, región, complejidad).
- Outputs: distribución de costos esperados, outliers, probabilidad de glosa, etc.

Modelo de eventos híbrido: clínico + administrativo-financiero

Eventos clínicos (desde RIPS)

- Encounter / Atención (tipo, servicio, sede, fecha).
- Diagnóstico (CIE-10) y motivo/causa (cuando exista).
- Procedimiento / Servicio (CUPS) con cantidades.
- Medicamentos / tecnologías (CUM/ATC u otros catálogos).
- Laboratorios / resultados (si están disponibles en RIPS).
- Ventanas de tiempo (tokens) para longitudinalidad.

Eventos financieros / proceso (SIIFA/FEV)

- Contrato y CUCON (acuerdo de voluntades).
- Factura / Nota crédito / Nota débito (DIAN) + CUV (MinSalud).
- Radicación, devolución, glosa, respuesta, ajustes.
- Pago (fecha, valor, mecanismo, giro directo).
- Reglas de oportunidad (SLA) y trazabilidad.

Resultado: secuencias de eventos para predecir costos, rangos tarifarios y señales de riesgo

Estrategia: normalizar datos para analítica y IA sin pelearse con los formatos fuente

Mapeo conceptual

- RIPS → Encounter, Condition, Procedure, Observation, MedicationRequest/Dispense.
- FEV (facturación) → Claim / Invoice (según alcance), más metadatos de radicación.
- SIIFA (proceso) → ClaimResponse / PaymentNotice + “eventos de proceso”.
- Mantener “lineage” (CUV, CUCON, IDs técnicos) para auditoría.

Opciones de implementación

A) “FHIR-lite” en lakehouse (MVP)

- Tablas Delta con estructura inspirada en FHIR.
- Rápido, bajo costo.
- Suficiente para el MVP tarifario.

B) FHIR Service (Azure HDS)

- API FHIR gestionada para intercambio/serving.
- Útil si se abren APIs a terceros.
- Se integra con Entra ID y analítica.

C) Híbrido

- Lakehouse como “sistema de análisis”.
- FHIR Service para interoperabilidad externa.
- Compartir vocabularios y trazabilidad.

Lakehouse + MLOps + agentes: una plataforma, múltiples productos

Fuentes

RIPS (JSON) • FEV/NC/ND (XML) • Contratos (CUCON) • Pagos/Glosas (SIIFA) • Maestros: REPS, BDUA, catálogos

Ingesta & Calidad

Pipelines batch/stream • Validaciones (CUV/CUCON) • Normalización de códigos • Pseudonimización / de-identificación

Lakehouse (Databricks)

Delta tables • Medallion (Bronze/Silver/Gold) • Unity Catalog (gobernanza) • Feature store / embeddings

Modelado & MLOps

Entrenamiento (GPU) • Registry • Evaluación y auditoría • Monitoreo drift/calidad

Serving & Apps

Azure Foundry (agentes) • APIs / dashboards • Asistente normativo (RAG) • Motor análisis de comportamiento tarifario



Tres caminos posibles (recomendación: híbrido)

Dimensión	A) Event-FM first	B) LLM-first (RAG)	C) Híbrido (recomendado)
	Mayor valor a largo plazo	Más rápido, menos profundo	Velocidad + escalabilidad
Qué construye	Modelo generativo de eventos (tipo CoMET) para costos/uso.	Asistente con RAG (normas + catálogos + FAQ) y prompts.	RAG + modelo de eventos mínimo viable (small) y “motor tarifario”.
Ventaja	Generaliza a muchas tareas; mejora con escala.	Time-to-value rápido; bajo costo inicial.	Rápido + prepara el camino al CoMET-scale completo.
Riesgo	Mayor demanda de datos limpios + GPU; más ingeniería inicial.	No capta dinámicas longitudinales (solo “consulta inteligente”).	Gestionar coherencia entre ambos componentes.
MVP 12 semanas	Posible si el vocabulario es acotado y el modelo es pequeño.	Muy viable (y necesario para normatividad).	Óptimo: entregar valor ejecutivo y dejar base para escalar.

Recomendación: C (Híbrido) para dar valor rápido y construir el “core” de largo plazo

Pipeline de modelado (recomendado): 2 modelos + 1 capa de evaluación

Modelo 1: Foundation Model de eventos (tipo CoMET)

- 1) Definir vocabulario mínimo (CUPS/CIE10/CUM + tokens de tiempo + contexto).
- 2) Construir secuencias por afiliado/episodio y “eventos de proceso” (radicación, glosa, pago).
- 3) Preentrenar modelo pequeño (p.ej., 100–300M params) para validar señal.
- 4) Evaluación: tareas proxy (predicción de próximo CUPS, probabilidad de glosa, costo esperado).
- 5) Ajuste por iteraciones: ampliar vocabulario, mejorar discretización, calibración.
- 6) Escalado: aplicar leyes de escala (tokens/compute) para decidir el siguiente tamaño.

Modelo 2: LLM + RAG normativo (Foundry)

- Ingesta de normas (SIIFA, FEV-RIPS, manuales) y catálogos.
- Chunking + embeddings + búsqueda semántica.
- Respuestas con citas (auditoría) + plantillas operativas.
- Herramientas: extracción documental, validaciones, generación de reportes.

Capa 3: evaluación, riesgos y gobierno

- Conjunto de “golden tasks” y métricas (exactitud, calibración, sesgo, deriva).
- Trazabilidad de datos: lineage, CUCON/CUV, control de versiones.
- Privacidad: minimización, seudonimización, acceso por roles.
- Aprobación clínica/financiera: comité de validación y “human-in-the-loop”.

Tres casos de uso para demostrar valor ejecutivo y técnico

Caso 1: Corridores tarifarios

- Top N procedimientos CUPS (p.ej., 200) y tecnologías.
- Rangos por región, complejidad, modalidad de pago.
- Descomposición: costo clínico vs. fricción administrativa.

Caso 2: Outliers y señales de riesgo

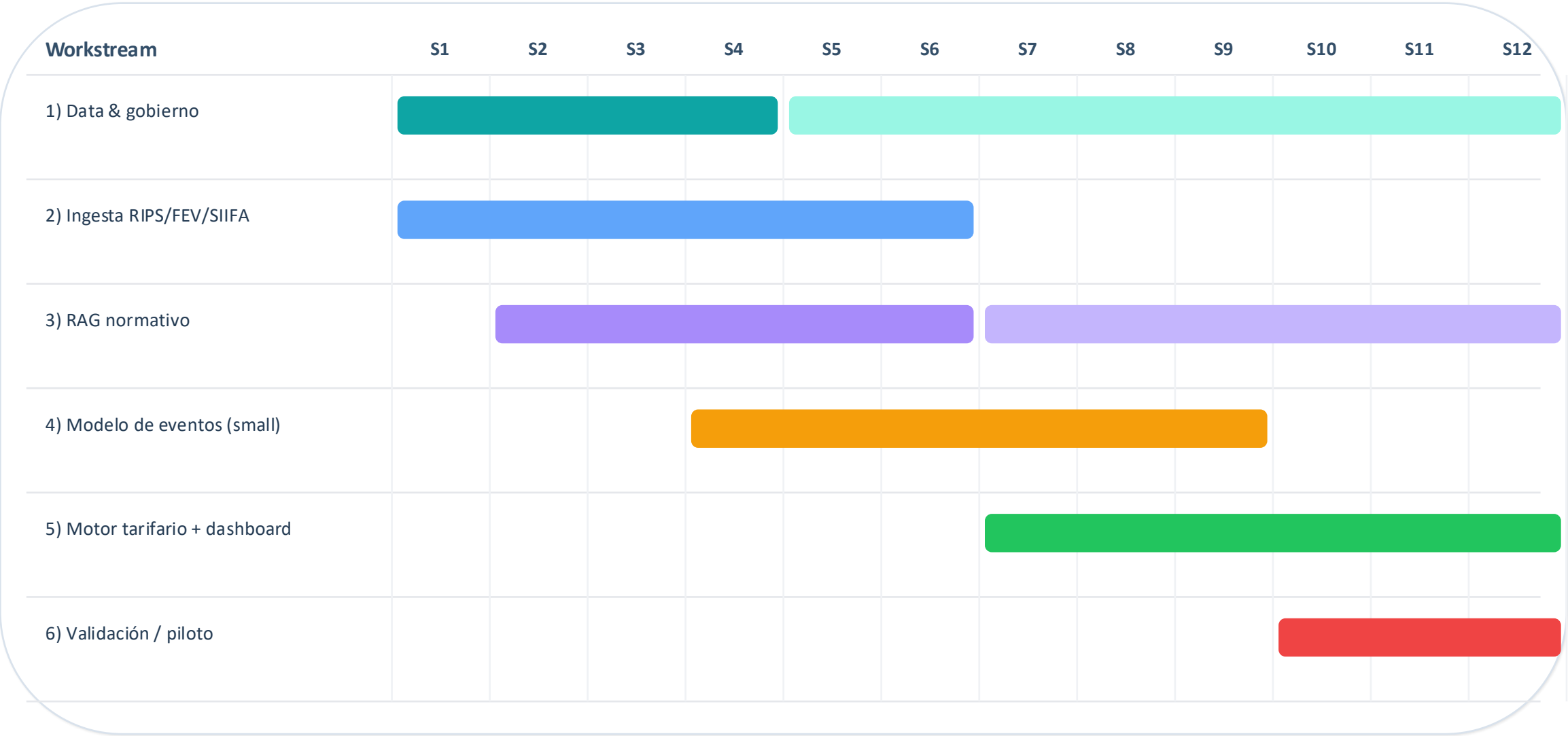
- Detección de valores atípicos por proveedor/territorio.
- Predicción temprana de probabilidad de glosa/devolución.
- Ranking de “drivers” (variables que más explican el outlier).

Caso 3: Asistente SIIFA/FEV-RIPS

- Preguntas normativas con respuesta citada.
- Checklist de cumplimiento y tiempos (SLA).
- Generación de reportes ejecutivos y soporte a auditoría.

Plan de trabajo detallado (12 semanas)

Objetivo: entregar MVP operativo + base para escalar (CoMET-Col-scale)



Hitos: S4 “Datos listos + pipelines básicos” • S8 “Modelo small + primeras métricas” • S12 “Demo director + plan de escalado”

Riesgos técnicos, operativos y de cumplimiento (con mitigaciones concretas)

Calidad y consistencia de datos

- Validaciones por capa (Bronze→Silver→Gold).
- Catálogos y mapeos versionados.
- Monitoreo de “missingness” y drift.

Riesgo de privacidad / reidentificación

- Minimización y seudonimización temprana.
- Acceso por roles (least privilege).
- Ambientes segregados y auditoría.

Sesgos y uso indebido

- Evaluación por subpoblaciones (región, régimen, complejidad).
- Human-in-the-loop para decisiones de política.
- Explicabilidad basada en evidencias.

Costo/tiempo de entrenamiento

- Modelo “small” en MVP; escalado por leyes de escala.
- Entrenamiento incremental y reuso de features.
- Priorización de vocabulario.

Dependencia normativa cambiante

- RAG con versionamiento de documentos.
- Tests de regresión normativos.
- Gobierno con comité técnico.

Adopción interna

- Dashboard ejecutivo + historias de uso.
- Capacitación y guías operativas.
- Pilotos con líderes de auditoría/pagos.

Para arrancar mañana: 6 decisiones de dirección

Decisiones

- 1) Casos MVP: (a) corredores tarifarios, (b) outliers, (c) asistente normativo.
- 2) Universo inicial: N CUPS/tecnologías + periodo histórico (p.ej., 24–36 meses).
- 3) Accesos: RIPS/FEV, pagos, contratos; catálogos maestros (calidad y versiones).
- 4) Cómputo: presupuesto GPU para entrenamiento “small” y evaluación.
- 5) Gobierno: comité clínico-financiero + seguridad + TI (aprobación de uso).
- 6) Piloto: 1–2 líneas de servicio / territorios para demo con impacto real.

Entregables al Director (Semana 12)

- Demo en vivo (dashboard + asistente + motor tarifario).
- Documento de arquitectura y seguridad.
- Plan de escalado: de “small” a “CoMET-scale” (datos, compute, roadmap).
- Propuesta de indicadores institucionales alineados al PEI.

Próxima reunión (1–2 horas)

- Alinear alcance de MVP y dataset mínimo.
- Nombrar responsables por workstream.
- Cerrar lista de decisiones y fechas (S1–S2).

Mensaje clave: empezamos pequeño, medimos, y escalamos con evidencia (leyes de escala + valor institucional)

3) Cómo lo logramos: COMET + Arquitectura para producción (DGTIC)

COMET = ciclo corto con evidencias y transferencia (de PoC a producción)

C

Capturar
caso de uso

dolor + KPI

O

Orquestar
datos

calidad + linaje

M

Modelar /
Experimentar

sandbox + guardrails

E

Entregar
producto

tablero/API

T

Transicionar
+ Medir

DGTIC + SLO

Rol de Arquitectura (TOGAF) para acelerar

- Gobierno mínimo viable: **decisiones, responsables y cadencia**.
- Arquitectura de información: **dominios, linaje, calidad** y seguridad.
- Arquitectura de aplicaciones/tecnología: **patrones y referencia**.
- Roadmap + “**paquetes de transferencia**” para DETIC (runbooks, SLO, monitoreo).

Puente a producción con DGTIC

- **Checklist de salida:** datos listos, modelo versionado, seguridad, costos.
- **Artefactos:** ADRs, diagramas, contratos API, plan de pruebas.
- **Operación:** monitoreo, alertas, métricas de valor, soporte.
- **Cierre:** **lecciones aprendidas y reutilización** (gestión del conocimiento).

La Tecnología Habilita la atención, la promoción y la prevención de enfermedades



Mejor salud

Mejorar la salud de la población



Mejor atención

Mejorar la experiencia de atención



Menor costo

Reducir el costo per cápita de la atención



Experiencia clínica

Mejore la eficiencia del flujo de trabajo de los médicos



Equidad en salud

Comprender los determinantes sociales de la salud

Transformación digital en salud

Arquitectura Macro

Datos

Semántica

Análisis

Aplicación



