

Juventud, evaluación e IA

02-06-2026



Mg. Hugo Morán Lojano | Co-coordinador EvalYouth Ecuador

www.glocalevalweek.org

El propósito: sociedad civil y juventudes en la evaluación

¿Por qué importa la sociedad civil?

La cultura evaluativa no se construye solo desde el Estado.

En Ecuador los esfuerzos se centran en seguimiento y rendición de cuentas. La evidencia se produce pero rara vez se usa para mejorar políticas.

El reto es generar demanda: que los tomadores de decisión vean la utilidad de la evaluación y la exijan. Cuando la evidencia es útil, los decisores vuelven a pedirla.

El valor de los YEEs



Pasión y compromiso genuino con el cambio



Perspectivas frescas e innovadoras



Destreza tecnológica (tech-savviness)



Liderazgo propositivo, no simbólico

Pasar de participación simbólica → involucramiento significativo y estructurado en la toma de decisiones evaluativa

Diagnóstico: Ecuador y el INCE

Índice INCE 2022-2024

5 dimensiones del INCE — puntaje sobre 10



2022 2024

Lo que el INCE nos dice

✓ Marco e infraestructura: en crecimiento

Estructura y calidad técnica mejoran. Ecuador produce evaluaciones, pero el ciclo no se cierra.

! Uso de evidencia: el mayor déficit

2.82 → 3.50 sobre 10. Las evaluaciones se perciben como trámites, no como insumos para decidir.

→ El reto: generar demanda real

Lograr que los tomadores de decisión vean la utilidad y exijan más evaluaciones. EvalYouth y SEEval son parte de esa palanca.



La metasíntesis EvalYouth Ecuador: el corpus

108

evaluaciones
en el banco

Universo del análisis

29

evaluaciones
seleccionadas

Criterios metodológicos

220

recomendaciones
a analizar

Unidad de análisis

¿Qué pregunta responde esta metasíntesis?

¿Qué nos dicen las 220 recomendaciones sobre el Estado ecuatoriano en materia de políticas públicas?
La sensibilidad a juventud no es un criterio previo: es una dimensión que el análisis mismo revelará.

Cómo usamos la IA: colaboración Humano-IA en la práctica

A

Humano selecciona

El equipo define criterios y elige las 29 evaluaciones del banco

B

IA extrae y codifica

Claude lee los PDFs y clasifica las 220 recomendaciones con el prompt

C

Humano valida

El equipo revisa el 20%. Toda interpretación final: humana

D

Hallazgos de 2° orden

Patrones estructurales que ninguna evaluación individual puede ver

El prompt que usamos con Claude

prompt.txt

Clasifica esta recomendación en 6 dimensiones. Responde SOLO en JSON.

- ① **Tipo:** Normativo · Presupuestario · Operativo · Coordinación
- ② **Innovación:** Correctiva · Adaptativa · Transformadora
- ③ **Factibilidad:** Alta · Media · Baja
- ④ **PND:** Directa · Parcial · Sin alineación
- ⑤ **Juventud:** Explícita · Implícita · No menciona
- ⑥ **Calidad:** ¿responsable? ¿específica? ¿indicador? → Alto/Medio/Bajo

Las 5 primeras dicen QUÉ ES la recomendación · La 6ª evalúa QUÉ TAN BUENA ES · Juntas responden: ¿por qué no se implementan?

Lo que devuelve Claude (JSON)

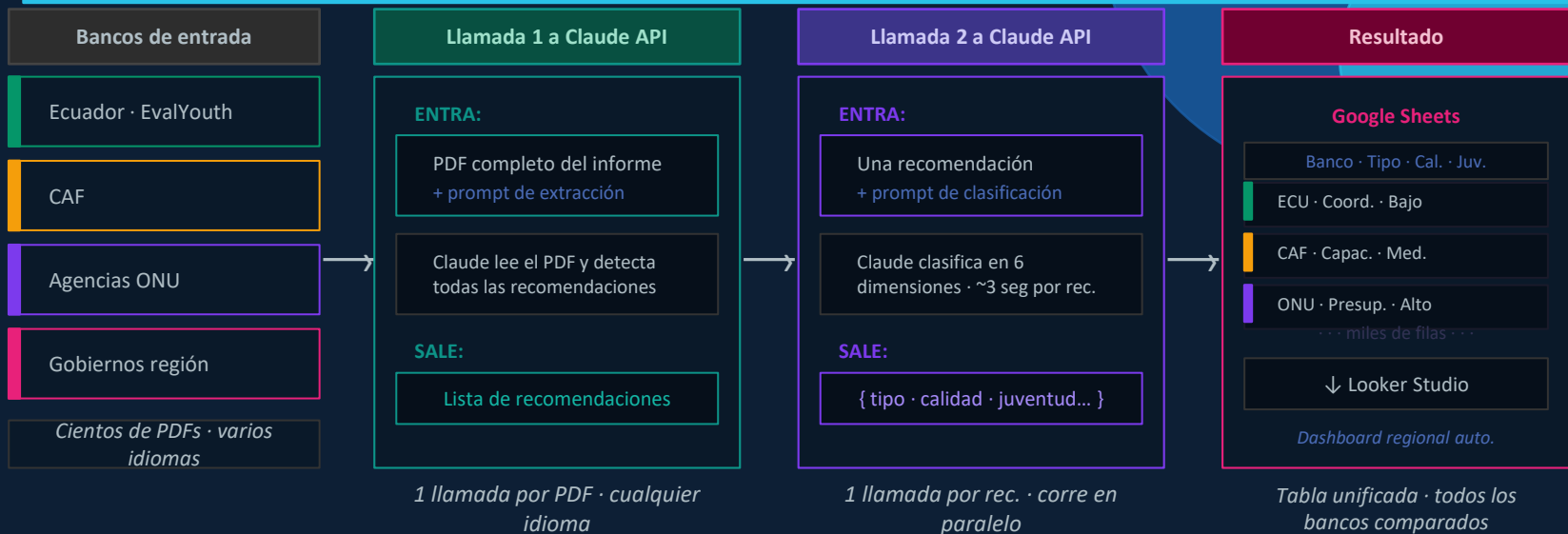
response.json

```
{
  "tipo": "Coordinación",
  "innovacion": "Correctiva",
  "factibilidad": "Media",
  "pnd": "Directa",
  "juventud": "Implícita",
  "calidad_puntaje": "Bajo",
  "responsable": "No",
  "especifica": "Parcial",
  "indicador": "No"
}
```

En Excel: una fila por recomendación · 6 columnas · tablas dinámicas para ver patrones de calidad.

Fase 2 — cómo trabaja Claude API con múltiples bancos

Dos llamadas a la IA · flujo automático · sin intervención humana entre pasos



¿Por qué dos llamadas separadas y no una sola?

Llamada 1 lee el contexto completo del PDF y extrae texto libre — tarea de comprensión documental. **Llamada 2** aplica el mismo criterio de clasificación a cada rec. — tarea de análisis estructurado.

De piloto a escala: mejorar recomendaciones en Ecuador y la región

FASE 1

Piloto Ecuador — EvalYouth · 220 recomendaciones · sin programar

NotebookLM

Sube y extrae

Claude.ai

Clasifica con prompt

Excel

Análisis y patrones

- 1 Extraes las 220 recomendaciones de los 29 informes y las organizas en una tabla
- 2 La IA clasifica cada una en 6 dimensiones — incluyendo si la recomendación es de calidad alta, media o baja
- 3 Identificas patrones: ¿qué % son vagas, sin responsable, sin indicador de seguimiento?
- 4 Produces criterios concretos para que futuras evaluaciones generen mejores recomendaciones

2–3 días · ~USD 5 · Sin instalar nada · Sin saber programar

FASE 2

Fase 2 — Escala regional con un socio técnico

Claude API

Extrae y clasifica solo

Google Sheets

Resultados automáticos

Looker Studio

Dashboard regional



Claude API hace el trabajo pesado

Lee los PDFs, extrae las recomendaciones y las clasifica en las 6 dimensiones.

Todo automático — nadie abre el navegador ni copia nada.

Google Sheets recibe los resultados solo

Cada recomendación clasificada se escribe automáticamente en una fila. 1.000 recomendaciones en una hora, sin intervención manual.

Looker Studio muestra el dashboard en tiempo real

Gráficos que se actualizan solos cuando entran nuevos datos. CAF, agencias ONU y gobiernos ven sus patrones desde el navegador.

El equipo humano solo interpreta

Valida el 20% y toma decisiones. El tiempo que antes se iba en procesar datos ahora se invierte en aprender de ellos.

¿Quién lo construye? Un socio técnico · 3 semanas · una vez · sirve para toda la región

Tenemos el piloto. Necesitamos un socio técnico.

Conectar Claude API + Google Sheets + Looker Studio · 3 semanas de trabajo · sirve para CAF, agencias ONU y gobiernos de la región.

CAF

PNUD

OIT

OPS

FAO

UNICEF

Primeros hallazgos: Fase 1 del análisis

Lo que el corpus ya nos revela

- 108 evaluaciones en el banco · 29 seleccionadas · 220 recomendaciones
- Promedio de 7–8 recomendaciones por evaluación
- Pocas tienen responsable explícito, indicador de seguimiento o plazo
- El doble problema: el Estado no usa la evidencia Y recibe recs. no diseñadas para implementarse
- Calidad y sensibilidad a juventud: dimensiones que el análisis mismo revelará

Las preguntas que guían el análisis

- CALIDAD · ¿Qué % son vagas, sin responsable o sin indicador? ¿Las evaluaciones recientes mejoran en calidad?
- JUVENTUD · ¿Cuántas mencionan a jóvenes como sujetos activos vs. beneficiarios pasivos?
- ESTADO · ¿Qué instituciones concentran las recs. de menor calidad y menor implementación?
- FUTURO · ¿Qué criterios deberían exigir los TdR para que futuras recs. sean implementables?

Prospectiva 2040: tres escenarios para la evaluación juvenil en Ecuador

01



Continuidad

La evidencia sigue archivada

Las recs. siguen siendo correctivas y de baja factibilidad. En 2040 se evalúan los mismos problemas. La IA procesa más rápido pero la evidencia sigue sin usarse.

02



Adaptación

La metasíntesis se institucionaliza

El banco deja de ser archivo y se convierte en memoria activa. Los nuevos TdR se diseñan con aprendizajes previos. La demanda de evaluación crece.

03



Transformación

Los jóvenes co-evalúan el Estado

EvalYouth lidera un sistema donde jóvenes son co-evaluadores. La IA democratiza el acceso. En 2040 cualquier organización juvenil consulta qué dice el Estado.

"No evaluamos para reportar ni para cumplir. Evaluamos para mejorar."

Mejores recomendaciones

Produce criterios concretos para que futuras evaluaciones generen recomendaciones específicas, con responsable claro y con indicador medible.

Demanda activada

Cuando la evidencia es útil y bien formulada, los tomadores de decisión la leen, la exigen y actúan sobre ella.

Modelo regional

Lo que EvalYouth demuestra en Ecuador puede viajar a CAF, agencias ONU y gobiernos de la región. Una vez funciona aquí, escala.