DIARIO DIGITAL DEV / RECURSO DE LECTURA

Cuánto cuesta la IA: medir el costo por tarea resuelta

El consumo de un modelo explica una parte del gasto. Una comparación útil también cuenta los intentos fallidos, las herramientas y el trabajo necesario para aceptar el resultado.

Ideas clave del artículo

- Define una tarea aceptada y conserva en el costo los fallos, reintentos y recursos compartidos atribuibles.

- Acompaña el costo unitario con calidad, volumen pendiente, tiempo de respuesta y revisión humana.

- Prueba caché y lotes según el comportamiento real del flujo; un descuento técnico no demuestra ahorro total.

Notas para tu equipo

_____

_____

_____

_____

Artículo y fuentes

https://www.digital-dev.cl/blog/costos-ia-empresas-tarea-resuelta/

FinOps Foundation: Effect of Optimization on AI Forecasting
https://www.finops.org/wg/effect-of-optimization-on-ai-forecasting/

Anthropic — Claude Platform Docs: Pricing
https://platform.claude.com/docs/en/about-claude/pricing

Anthropic — Claude Platform Docs: Prompt caching
https://platform.claude.com/docs/en/build-with-claude/prompt-caching

Anthropic — Claude Platform Docs: Batch processing
https://platform.claude.com/docs/en/build-with-claude/batch-processing