

Rúbrica de Evaluación — CTRL+HACK 2.0

Documento confidencial para jueces. Score total: 0–100 puntos. El equipo con mayor puntuación gana.

Cómo usar esta rúbrica

1. Verificar que el equipo **califica para evaluación** (Gate de Selección)
2. Evaluar cada una de las 5 dimensiones de forma independiente
3. Sumar los puntajes
4. En caso de empate: primero por Complejidad Técnica, luego por puntos totales de agentes seleccionados

Paso 1 — Gate de Selección (¿Califica?)

Antes de evaluar, confirmar que el equipo cumple el umbral mínimo de dificultad.

Modalidad	Puntos mínimos requeridos
Equipo de 3 personas	10 puntos
Participante individual	7 puntos

Tabla de puntos por agente:

Agente	Puntos	Tipo
ERP & Gestión Empresarial	2 pts	Cloud
Producción Avanzada	3 pts	Hybrid
Digital Twin	3 pts	Cloud
Mantenimiento & CMMS	3 pts	Hybrid
Calidad & SPC	3 pts	Hybrid
Energía & Sustentabilidad	3 pts	Hybrid
Cadena de Suministro	3 pts	Cloud

Ciberseguridad Industrial	4 pts	Hybrid
Infraestructura & Edge	4 pts	Hybrid
Control & SCADA	4 pts	Edge
Seguridad Industrial & HSE	4 pts	Hybrid
AI & ML Industrial	5 pts	Hybrid
Visión Artificial Industrial	5 pts	Hybrid

Reglas:

- El mismo agente no puede contarse dos veces
- Si el equipo no llega al umbral → **No califica. Score = 0.**
- Puntos totales del equipo: _____ (registrar aquí)

Paso 2 — Rúbrica de Calidad (0–100 pts)

Dimensión 1: Complejidad Técnica /20

Evalúa si el agente realmente ejecuta su función o es una simulación superficial.

Puntos	Criterio
0–5	Demo con datos hardcoded, sin lógica real, mocks estáticos
6–10	Lógica de tool implementada pero sin separación real Edge/Cloud
11–15	Tools funcionales, separación Edge/Cloud existe aunque sea simulada con claridad
16–20	Tools ejecutan con datos reales o simulados realistas, Edge/Cloud en runtimes separados o claramente diferenciados

Preguntas guía:

- ¿Las tools del agente hacen algo cuando se ejecutan o solo retornan JSON fijo?
- Si el agente es Hybrid, ¿hay una diferencia real entre lo que procesa en Edge vs Cloud?
- ¿El sistema funciona con datos que no son los del ejemplo de la presentación?

Score D1: _____ / 20

Dimensión 2: Cumplimiento IES v2.0 /20

Evalúa si los eventos publicados por el agente cumplen el Industrial Event Standard.

Revisión objetiva — cada campo faltante descuenta puntos:

Campo requerido	Presente (✓/X)
<code>event_id</code>	
<code>timestamp</code> (formato ISO 8601 UTC)	
<code>platform_version: "2.0.0"</code>	
<code>module.id</code> (en snake_case)	
<code>module.version</code>	
<code>asset.asset_id</code>	
<code>event.type</code> (en snake_case)	

- 7/7 campos presentes = 20 pts
- Por cada campo faltante = -3 pts (mínimo 0)
- Si el evento ni siquiera llega al API = 0 pts

Criterios adicionales (hasta +5 pts bonus incluidos en los 20):

- ¿El `event.category` usa alguno de los valores válidos? (`quality` , `productivity` , `maintenance` , `energy` , `safety` , `configuration` , `system`)
- ¿El `event.severity` usa uno de los valores válidos? (`low` , `medium` , `high` , `critical`)
- ¿El agente respeta la lógica de `triggerCondition` (no emite eventos siempre, sino cuando la condición aplica)?

Score D2: _____ / 20

Dimensión 3: Utilidad Industrial Real /20

Evalúa si el agente resuelve un problema que un gerente de planta o supervisor reconocería como valioso.

Puntos	Criterio
0–5	El caso de uso es genérico o no está explicado claramente
6–10	Caso de uso reconocible pero la implementación no demuestra impacto claro
11–15	Problema industrial claro, la demo muestra cómo el agente lo resuelve
16–20	Caso de uso convincente con datos realistas, el juez puede imaginar el agente en producción real

Preguntas guía:

- ¿Los valores de los datos son realistas? (temperatura, vibración, SKUs, volúmenes de stock)
- ¿El agente tiene un nombre de asset_id que sugiere un contexto industrial real? (MOTOR-BANDA-03 vs ASSET-001)
- ¿Podrías presentar esta demo a un cliente industrial sin que se rían?

Score D3: _____ / 20

Dimensión 4: Innovación y Alineación al Reto **/20**

Evalúa si el equipo fue más allá del mínimo requerido.

Puntos	Criterio
0–5	Implementación mínima, sin nada más allá de lo requerido
6–10	Cumple el reto con buen nivel de acabado
11–15	Detalles extra: más de 3 tools, comunicación entre 2 agentes, escenario de fallo/recuperación
16–20	Algo sorprendente: comunicación cruzada entre múltiples agentes, comportamiento emergente, caso de uso no anticipado en las bases

Puntos adicionales a buscar:

- ¿El agente publica eventos que otro agente del equipo consume? (integración cross-agent)
- ¿Hay lógica de fallback cuando la conexión al cloud falla (modo offline)?
- ¿Se implementaron más tools de las mínimas requeridas (>3)?
- ¿El equipo definió sus propios tipos de eventos más allá de los ejemplos?

Score D4: _____ / 20

Dimensión 5: Presentación y UI /20

Evalúa la interfaz gráfica (requerida) y la claridad del demo.

La interfaz gráfica es obligatoria. Un proyecto sin UI no puede obtener más de 10 pts en esta dimensión.

Puntos	Criterio
0–5	Sin interfaz gráfica funcional, o demo incomprensible
6–10	Interfaz básica funcional, demo entendible con ayuda de explicaciones
11–15	Interfaz clara que muestra los eventos y acciones del agente en tiempo real
16–20	Interfaz bien diseñada, intuitiva, el demo habla por sí solo sin necesitar explicación verbal

Preguntas guía:

- ¿La UI muestra los eventos IES en tiempo real o al menos al dispararse la acción?
- ¿Un observador sin contexto técnico puede entender qué está pasando?
- ¿La interfaz refleja el contexto industrial (nombres de plantas, líneas, activos)?

Score D5: _____ / 20

Resumen de evaluación

Dimensión	Score
D1 — Complejidad Técnica	/20

D2 — Cumplimiento IES v2.0	/20
D3 — Utilidad Industrial Real	/20
D4 — Innovación y Alineación	/20
D5 — Presentación y UI	/20
TOTAL	/100

Calificó para evaluación (Gate): Sí / No

Puntos totales de agentes seleccionados: _____ pts

Notas del juez:

(espacio libre para comentarios)

Criterio de desempate

1. Mayor score en Dimensión 1 (Complejidad Técnica)
2. Mayor suma de puntos de agentes seleccionados
3. Decisión del jurado por mayoría