

PAQUETE PARA EDUCADORES

Kit de evaluación Imágenes con IA

Instrumentos diagnósticos, tareas de desempeño, rúbricas y protocolos de calibración para valorar decisiones responsables.

MATERIAL DE TRABAJO | AIIMG-2026.1

Evaluar criterio, no intuición visual

AUDIENCIA

Educadores y responsables de formación

DURACIÓN

45–90 minutos de evaluación

USO

Diagnóstico y evidencia final

Afirmación de validez

Los instrumentos valoran si una persona puede separar observación de inferencia, verificar afirmaciones, considerar consentimiento e impacto, elegir una respuesta proporcional y documentar límites. No miden la capacidad de identificar el modelo que produjo una imagen.

Mapa de evidencias

Resultado	Evidencia	Instrumento
Distingue observación, inferencia y afirmación.	Respuesta escrita	Diagnóstico 1–4; tarea A.
Verifica fuente, contexto y procedencia.	Plan de comprobación	Diagnóstico 5–8; tarea A.
Reconoce consentimiento, identidad, privacidad y sesgo.	Justificación	Diagnóstico 9–12; tarea A y C.
Responde ante daño sin amplificarlo.	Desempeño observado	Tarea B.
Decide y documenta con responsabilidad humana.	Memorando editorial	Tarea C.

Uso recomendado

1. Aplicar el diagnóstico antes del taller, sin discusión previa.
2. Seleccionar al menos una tarea de desempeño; la tarea B requiere el acuerdo de cuidado.
3. Calificar con la rúbrica común y una rúbrica específica.
4. Calibrar dos muestras antes de puntuar todo el grupo.
5. Reaplicar los ítems paralelos y comparar razonamiento, no solo porcentaje.

Límite

Estos instrumentos no sustituyen una investigación de protección, disciplina o cumplimiento. Una revelación real detiene la evaluación y activa la ruta institucional correspondiente.

Formulario previo · 12 ítems

Datos

Código:

Fecha:

Grupo:

Seleccione la mejor respuesta. En los ítems 4, 8 y 12 añada una frase de justificación.

1. Una imagen realista muestra seis dedos. La conclusión más sólida es:
A) es IA; **B)** es falsa; **C)** existe una inconsistencia que merece revisión; **D)** la persona tiene seis dedos.
2. Una imagen sin defectos visibles:
A) es auténtica; **B)** puede seguir siendo generada, editada o descontextualizada; **C)** no necesita fuente; **D)** es segura para publicar.
3. La frase "ganó el premio" debajo de una foto con un trofeo es:
A) observación; **B)** metadato; **C)** afirmación que requiere evidencia; **D)** texto alternativo.
4. ¿Qué comprobarías primero para una imagen de una emergencia escolar?
A) número de dedos; **B)** canal oficial independiente; **C)** estilo fotográfico; **D)** cantidad de reacciones.
Razón:
5. Una visualización de un edificio futuro se publica como "nuestro edificio hoy". El problema principal es:
A) resolución; **B)** color; **C)** afirmación temporal engañosa; **D)** tamaño del archivo.
6. Tener permiso para tomar una fotografía significa que:
A) cualquier edición está permitida; **B)** debe comprobarse si el permiso cubre transformación, finalidad, audiencia y plazo; **C)** la institución posee la identidad; **D)** no hace falta procedencia.
7. Antes de cargar una foto de aula en una herramienta se debe revisar:
A) solo el rostro principal; **B)** fondo, pantallas, horarios, gafetes y metadatos; **C)** solo el nombre del archivo; **D)** únicamente la iluminación.
8. Un LLM encuentra que la corbata no coincide con la especificación. La mejor acción es:
A) acusar fraude; **B)** sancionar; **C)** tratarlo como señal y compararlo humanamente con referencias; **D)** publicar.
Razón:
9. Una galería asigna repetidamente liderazgo a un solo grupo. Esto sugiere:
A) un patrón que debe auditarse en el conjunto; **B)** que cada retrato es falso; **C)** que la diversidad se adivina por apariencia; **D)** que basta añadir una persona.
10. Recibes un meme humillante sobre alguien reconocible. La primera acción más segura es:
A) reenviarlo a amistades; **B)** responder públicamente; **C)** no amplificar, ofrecer apoyo privado y usar el canal de reporte; **D)** buscar quién lo creó.
11. Una captura anuncia cancelación de clases. Falta la URL. Debes:
A) reenviar por precaución; **B)** verificar en un canal oficial independiente; **C)** votar en el chat; **D)** aceptar el diseño como evidencia.
12. Al responder a una persona afectada por una imagen manipulada conviene decir:
A) "muéstramela"; **B)** "seguro se borra"; **C)** "gracias por contarlo; no es tu culpa y no necesitas volver a mostrarla"; **D)** "ignórala".
Razón:

CLAVE DEL DIAGNÓSTICO

Respuestas y lectura

Ítem	Resp.	Principio
1	C	Un artefacto inicia revisión; no prueba origen ni falsedad completa.
2	B	La apariencia impecable no autentica fuente o contexto.
3	C	El objeto visible no demuestra la historia atribuida.
4	B	La urgencia exige confirmación independiente, no análisis estético.
5	C	Una visualización futura puede engañar si se presenta como presente.
6	B	Consentimiento específico, informado y revocable según el contexto.
7	B	La privacidad incluye el fondo y el archivo, no solo el sujeto.
8	C	El modelo aporta señales; una persona conserva la decisión.
9	A	El sesgo puede ser distributivo y requiere auditoría del conjunto.
10	C	Contener y apoyar reduce exposición y devuelve agencia.
11	B	Una interfaz familiar no acredita el origen.
12	C	La respuesta inicial evita culpa, repetición y promesas imposibles.

Puntuación diagnóstica

Otorgue un punto por cada opción correcta. Para 4, 8 y 12 añada un punto si la justificación nombra respectivamente *fuentes independiente, revisión humana y bienestar / no exposición*. Total: 15 puntos.

0-6	Requiere andamiaje explícito y casos de baja sensibilidad.
7-11	Comprensión emergente; necesita practicar transferencia y justificación.
12-15	Preparado para tareas complejas; no equivale todavía a desempeño competente.

No usar como etiqueta

El diagnóstico orienta la enseñanza. No debe convertirse en una clasificación permanente ni en evidencia disciplinaria.

Memorando de verificación visual · 20 puntos

REFERENCIA · PATIO ACTUAL



VISUALIZACIÓN IA · PROYECTO FUTURO



Encargo

Redacte un memorando de 250–400 palabras para decidir si la segunda imagen puede publicarse en: **a)** una consulta sobre el proyecto futuro y **b)** el informe anual sobre instalaciones actuales.

El memorando debe incluir

1. dos observaciones y dos afirmaciones que requieren respaldo;
2. comparación de propósito, audiencia y canal;
3. evidencia disponible, evidencia faltante y una incertidumbre;
4. consentimiento / procedencia y una revisión de privacidad;
5. decisión para cada canal, etiqueta o pie propuesto y condición de retiro.

Registro del evaluador

Observación / afirmación	/4
Verificación y evidencia	/4
Personas, consentimiento y riesgo	/4
Decisión proporcional	/4
Transparencia y responsabilidad	/4
Total	/20

Respuesta en los primeros cinco minutos · 20 puntos

Escenario

Una persona comunica en privado que circula una imagen manipulada sobre ella. No quiere volver a verla, teme represalias y pregunta si la escuela puede garantizar que desaparecerá. Quien responde no necesita ver el contenido para iniciar apoyo.

Organización

Triadas: persona que reporta, persona que responde y observador/a. Dos minutos de preparación, cinco de actuación y cinco de debriefing. Cambiar roles solo si el grupo está cómodo y el tiempo lo permite.

OK	Criterio	Evidencia observable	/4
☐	Recepción y agencia	Agradece, evita culpa, permite pausar y no exige ver.	
☐	Seguridad y bienestar	Pregunta solo lo necesario sobre amenaza, chantaje y apoyo.	
☐	Contención	Reduce exposición y no pide reenvío, descarga o búsqueda.	
☐	Activación	Nombra responsable, canal y escalamiento proporcional.	
☐	Seguimiento	No promete borrado; acuerda próxima actualización y apoyo.	

Notas del observador

Una frase que devolvió control:

Una pregunta que podía omitirse:

Un siguiente paso claro:

Total

/20

Condición de seguridad

Si se menciona un caso real, amenaza inmediata o contenido sexualizado de menores, termine la simulación. No recoja detalles frente al grupo y active la ruta local.

Consejo editorial · 24 puntos

Expediente

El equipo de admisiones propone un retrato generado a partir de tres fotografías autorizadas de una persona adulta ficticia. La finalidad original era un boletín interno. La versión lleva uniforme institucional, cambia la expresión y se publicaría en una campaña abierta. Un LLM señaló que la corbata y la insignia no coinciden con la especificación. El lanzamiento vence hoy.

Producto

Prepare una decisión de una página y una presentación de tres minutos. Elija **PUBLICAR, CORREGIR, RESTRINGIR o RECHAZAR**. Incluya:

- finalidad, audiencia y afirmación institucional;
- alcance del consentimiento y posible semejanza / dignidad;
- comparación humana con la especificación de uniforme;
- tratamiento de la salida del LLM como señal, no veredicto;
- transparencia, texto alternativo, procedencia y condición de retiro;
- nombre del rol que conserva aprobación y seguimiento.

Puntuación

Propósito y afirmaciones	/4
Identidad y consentimiento	/4
Revisión de especificaciones	/4
Uso responsable del LLM	/4
Publicación accesible y transparente	/4
Propiedad de la decisión y retiro	/4
Total	/24

Cuatro niveles de desempeño

Aplique seis criterios de 1 a 4 puntos. En blanco no equivale automáticamente a 1: permita completar si la instrucción fue ambigua o inaccesible.

Criterio	4 · Transferible	3 · Competente	2 · Emergente	1 · Inicial
Observación y afirmación	Separa capas y detecta afirmaciones implícitas.	Separa observación e inferencia con un error menor.	Mezcla capas, pero corrige con apoyo.	Trata apariencia o pie como hecho.
Evidencia y procedencia	Prioriza fuentes independientes y documenta límites.	Propone comprobaciones pertinentes y registra fuente.	Propone una comprobación parcial o dependiente.	Confía en artefactos, reacciones o autoridad aparente.
Consentimiento y privacidad	Examina finalidad, transformación, audiencia, tiempo y datos mínimos.	Identifica permiso y datos principales.	Reconoce el tema sin precisar alcance.	Supone permiso amplio o ignora datos del fondo.

Criterio	4 · Transferible	3 · Competente	2 · Emergente	1 · Inicial
Impacto y protección	Prioriza bienestar, agencia y contención; anticipa recirculación.	Evita amplificar y activa apoyo adecuado.	Reconoce daño pero propone exposición innecesaria.	Culpa, minimiza, investiga públicamente o reenvía.
Decisión proporcional	Elige acción, condiciones y alternativa; conserva incertidumbre.	Elige una acción defendible con razón.	La acción es vaga o excesiva.	La acción depende de certeza inexistente o de presión.
Responsabilidad y acceso	Nombra responsable, transparencia, equivalencia accesible y retiro.	Incluye responsable y divulgación básica.	Incluye uno de los controles.	Delega a la herramienta o no deja seguimiento.

Umbral sugerido

Desempeño competente: al menos 18 de 24 puntos, sin nivel 1 en *impacto y protección*. Ajuste el umbral a su contexto y ofrezca una nueva oportunidad con un escenario equivalente.

Puntuar la misma calidad con el mismo criterio

Protocolo de 20 minutos

1. Lea la tarea y subraye la evidencia exigida, sin mirar todavía respuestas.
2. Puntúe individualmente una muestra de referencia.
3. Compare criterios con una diferencia de dos o más puntos.
4. Cite palabras o acciones observables; no infiera intención o conocimiento.
5. Acuerde un ancla de nivel 2, una de 3 y una de 4.
6. Vuelva a puntuar una segunda muestra y registre el acuerdo.

Registro de calibración

Tarea:

Fecha:

Evaladores:

Criterio con mayor diferencia:

Evidencia textual / observable acordada:

Cambio en la interpretación de la rúbrica:

Muestras breves para conversar

Muestra	Respuesta	Conversación
A	"Tiene seis dedos, por tanto es falsa y la retiro."	¿Qué inferencia no está sustentada?
B	"No puedo confirmar el aviso. Verifico el canal oficial y no reenvío la captura."	¿Qué falta para nivel 4?
C	"Corrijo el uniforme porque el LLM lo rechazó."	¿Quién decide y qué evidencia compara?

Transferencia a un caso nuevo

Caso paralelo

Una imagen convincente muestra a un grupo usando un nuevo espacio deportivo. El archivo llegó por correo sin fuente. En el fondo se ve un horario. El pie propuesto dice "Ya está abierto para toda la comunidad". Un detector asigna 82 % de probabilidad de IA.

Respuesta individual · 12 minutos

Dos observaciones:

Dos afirmaciones que necesitan evidencia:

La fuente o canal independiente que consultarías:

Un riesgo de privacidad:

Qué significa y qué no significa el 82 %:

Decisión: ☐ publicar ☐ corregir ☐ restringir ☐ rechazar

Justificación, divulgación y responsable:

Clave de puntuación · 12 puntos

Dos puntos por cada elemento: observaciones precisas; afirmaciones (apertura actual y acceso comunitario); verificación independiente; minimización del horario; interpretación probabilística no concluyente; decisión documentada con responsable.

De resultados a mejoras

Resumen del grupo

Participantes:

Fecha:

Promedio diagnóstico /15:

Promedio post /12:

Criterio más sólido:

Criterio que requiere reenseñanza:

Diferencia notable por formato o necesidad de acceso:

.....

Acción de seguimiento y fecha:

.....

Patrón observado	Respuesta didáctica
Se confía en defectos visuales.	Volver a imagen–afirmación–evidencia con una imagen impecable.
Se ignora el alcance del permiso.	Comparar cuatro finalidades usando la misma referencia.
Se reenvía “para reportar”.	Practicar registro mínimo y reporte sin adjuntar.
Se delega al LLM o detector.	Pedir citas a una rúbrica y una decisión humana separada.
La decisión carece de seguimiento.	Añadir propietario, hora de actualización y condición de retiro.

Informe responsable

Reporte patrones agregados y necesidades de apoyo. No publique respuestas sensibles, relatos personales ni clasificaciones individuales innecesarias.

Imágenes con intención. Procesos con responsabilidad.



Una guía para crear, examinar y publicar imágenes con IA.
Criterio educativo, cuidado por las personas y transparencia institucional.