

# Búsqueda trazable de evidencia con IA: Tarjeta-Prompt y regla anti-alucinaciones en Ciencias de la Salud



**Prof. Francisco José García González, PhD**

Facultad de Enfermería, Fisioterapia y Podología · Universidad Complutense de Madrid

01 Contexto

02 Metodología

03 Prompt en acción


04 Rúbrica

05 Resultados

06 Replicar

# La IA generativa en educación sanitaria: oportunidad con riesgos reales


## EL CONTEXTO

 Asignatura optativa Grado Enfermería UCM: Iniciación a la investigación Clínica en enfermería

 36 alumnos · 3º y 4º curso

 Uso libre de IA generativa (ChatGPT, Claude...)

 Bases: PubMed · CINAHL · Cochrane · Scopus

 Objetivo: buscar evidencia clínica con IA de forma segura y verificable

## LOS 3 RIESGOS CLAVE



### Alucinaciones

DOIs falsos · estudios inexistentes · fechas inventadas



### Irrastreabilidad

Sin strings, sin filtros · imposible replicar ni auditar



### Riesgo clínico

Decisiones asistenciales sobre evidencia no verificada

# Tarjeta-Prompt → Trazabilidad → Verificación doble

01



## Tarjeta-Prompt

Plantilla CO-STAR  
Umbral t=0,90  
Tabla 6 columnas estandarizada

02



## Trazabilidad

PICO definida  
Strings de búsqueda  
Checklist + DOIs verificables

03



## Verificación Doble

GPT docente → pre-evaluación  
Profesor → check coherencia  
Rúbrica 0-10 · 5 criterios

## ¿Qué cambia al usar la Tarjeta-Prompt?



### SIN TARJETA-PROMPT

#### PROMPT

"Dime qué evidencia hay sobre úlceras por presión en UCI"

#### RESPUESTA

"Según un estudio de 2021 publicado en Lancet (DOI:10.1016/xxx)" → DOI inventado

#### PROBLEMA

No hay tabla · No hay nivel de evidencia · No hay fuente real

#### RESULTADO

El alumno no puede verificar nada · Riesgo de alucinación no detectada



### CON TARJETA-PROMPT

#### PROMPT

"Contexto: Enfermería UCI. Estilo: tabla 6 col. Audiencia: asistencial. Regla t=0,90..."

#### SALIDA

Tabla: afirmación · nivel GRADE · tipo fuente · fecha · DOI real · notas/limitaciones

#### REGLA t

"No lo sé" generado cuando la evidencia no supera el umbral de confianza 0,90

#### RESULTADO

El alumno verifica cada DOI · detecta lagunas · aplica AMSTAR-2/CASPe/PEDro

# CO-STAR + regla anti-alucinación t=0,90: la plantilla completa

## TARJETA-PROMPT ESTANDARIZADA

<b>Contexto:</b>	[disciplina/unidad] + tema específico
<b>Objetivo:</b>	síntesis: qué sabemos y qué NO sabemos según evidencia actual
<b>Estilo:</b>	tabla 6 col: Afirmación   Nivel evid.   Tipo   Fecha   DOI/URL   Notas
<b>Tono:</b>	Registro lingüístico apropiado (clínico, divulgativo, técnico)
<b>Audiencia:</b>	Quién va a leer la información: equipo asistencial / academia
<b>Salida:</b>	solo tabla · sin divagaciones · sin citas no verificables

## CO-STAR + CLEAR

C · Contexto clínico específico	+	C · Conciso
O · Objetivo concreto		L · Lógico
S · Estilo de salida (tabla)		E · Explicito
T · Tono		A · Adaptado
A · Audiencia asistencial		R · Reflexivo
R · Respuesta estructurada		

## ⚡ REGLA ANTI-ALUCINACIÓN

### t = 0,90

«Responde solo si tu probabilidad interna de acierto es > t; acierto = +1; error = -t/(1-t); IDK = 0. Si no superas el umbral o la evidencia es insuficiente, responde exactamente "No lo sé" e indica qué información faltaría para responder con seguridad.»

# Output: tabla de 6 columnas · Tema: analgesia no farmacológica en neonatos

Afirmación	Nivel evid.	Tipo fuente	Fecha	DOI / URL	Notas / Limitaciones
La succión no nutritiva (SNS) reduce el dolor en punción de talón neonatal	Alta (GRADE)	RS Cochrane	2023	10.1002/14651858.CD013253	Solo neonatos a término; I <sup>2</sup> =48%
La sacarosa oral reduce la respuesta al dolor en RN ≥32 semanas	Alta (GRADE)	Meta-análisis ECA	2022	10.1542/peds.2021-055033	Dosis óptima no establecida; ausencia datos >28 días
El contacto piel con piel reduce PIPP-R durante punción de talón	Moderada	ECA (n=120)	2021	10.1016/j.pain.2020.09.014	Muestra pequeña; no cegamiento · PEDro=6/10

⚠ "No lo sé" → Guías ≤5 años sobre piel con piel en prematuros <32 sem: evidencia insuficiente. Faltaría: ECA en esta población específica.

✅ El alumno verifica cada DOI en Crossref · asigna herramienta de calidad (AMSTAR-2/CASPe/PEDro) · elimina entradas no accesibles · completa Notas

# Verificación doble: GPT docente (pre-filtro) + Profesor (check final)

⚠ Penalizaciones: -0,5 pt por DOI roto sin sustitución · -1 pt sin registro de filtros · -1 pt si no se identifican lagunas cuando la evidencia es insuficiente



## Pregunta y estrategia PICO

2pt

- Claridad de la pregunta PICO/PiCo
- Adecuación de términos MeSH/DeCS
- Filtros: años, tipo estudio, idioma



## Ejecución Tarjeta-Prompt

2pt

- Salida en tabla correcta, sin divagaciones
- 6 columnas completas y coherentes
- Regla  $t=0,90$  aplicada



## Trazabilidad

2pt

- Checklist completo
- DOI/URL válidos y clicables
- Strings + recuentos por base



## Verificación crítica

3pt

- AMSTAR-2 / CASPe / PEDro / AGREE-II
- Notas/limitaciones argumentadas
- Lagunas ('No lo sé') identificadas



## Presentación y citación

1pt

- Formato limpio y legible
- Estilo cita consistente: Vancouver / APA 7ª

# GPT Docente + Profesor: corrección escalable sin perder calidad



## Feedback estandarizado

El GPT aplica la misma rúbrica para todos, eliminando variabilidad en la evaluación inicial



## Prioriza fallos frecuentes

Detecta automáticamente: DOI roto, búsqueda no replicable, PICO vaga, filtro incoherente



## Libera tiempo docente

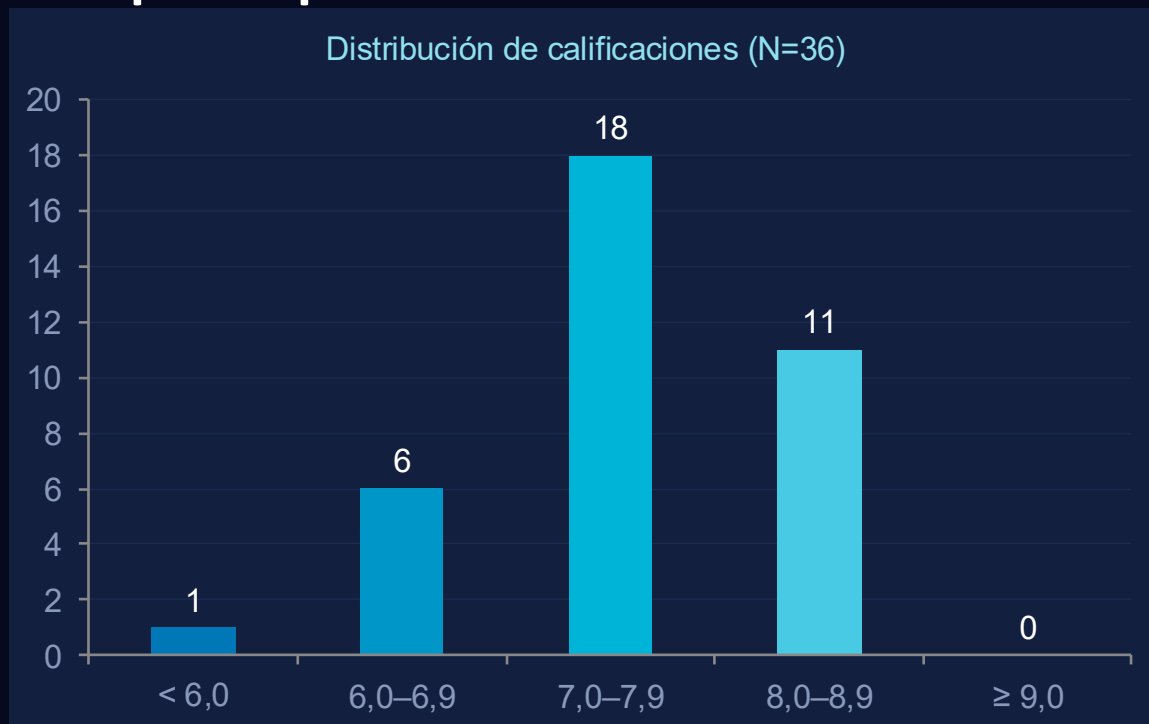
El profesor reserva su juicio experto para la calidad clínica y metodológica de nivel alto



## No sustituye al experto

El GPT pre-filtra; el profesor confirma coherencia PICO-evidencia y calidad crítica final

# N=36 · Concentración en 7,0–8,9 · Rango 5,2–8,5 · Ningún suspenso por alucinación



**83%** aprobados  $\geq 7,0$

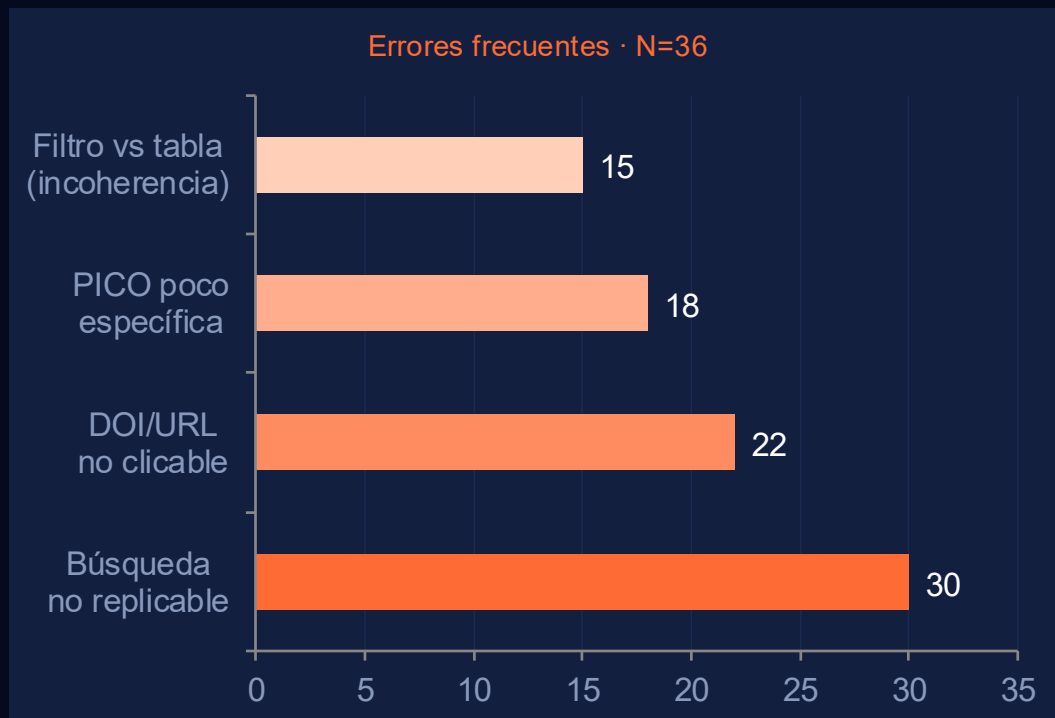
**7,0–8,9** concentración de notas

**5,2** nota mínima

**8,5** nota máxima

💡 La regla  $t=0,90$  funcionó como primera red de seguridad: ningún alumno entregó DOIs inventados sin corregir

# ¿Dónde falla el alumnado? Patrones detectados y soluciones



30

## Sin string/fecha/recuento

→ Obligar campo 'estrategia de búsqueda' en el registro de trazabilidad

22

## DOI/URL no clicable

→ Penalización -0,5 pt · verificación en Crossref obligatoria

18

## PICO poco específica

→ Taller previo de formulación PICO con ejemplos de la disciplina

15

## Filtros ≠ tabla generada

→ Checklist cierra tabla: incluidos + excluidos con motivo explícito



La verificación doble GPT + profesor depuró el 100% de referencias inventadas y detectó todas las incoherencias de filtros

# De Enfermería a cualquier disciplina sanitaria: implementación inmediata

01



Descarga la Tarjeta-Prompt

Disponible en campus Moodle. Adapta 'Contexto' a tu disciplina. Copia literalmente el umbral  $t=0,90$  y la regla de salida estructurada. 5 minutos de adaptación.

02



Diseña el registro de trazabilidad

Checklist incluido. Exige: PICO · strings + fechas · bases consultadas · filtros · DOIs verificados · herramienta de calidad (AMSTAR - 2/CASPe/PEDro/AGREE-II).

03



Configura un GPT docente (opcional)

ChatGPT Custom GPT o Claude Project con la rúbrica como system prompt. Pre-evalúa y genera feedback estandarizado. El profesor hace el check final de calidad.

04



Aplica la rúbrica de 5 criterios

PICO/estrategia (2) · Tarjeta-Prompt (2) · Trazabilidad (2) · Verificación crítica (3) · Presentación (1). Penaliza DOIs rotos y lagunas no identificadas.

✓ Transferible a: Enfermería · Fisioterapia · Podología · Medicina · Farmacia · cualquier Grado en Ciencias de la Salud

## Impacto y transferibilidad



La Tarjeta-Prompt + trazabilidad convierte la búsqueda con IA en un proceso replicable y auditable



La regla  $t=0,90$  funciona: los 'No lo sé' documentados son tan valiosos como las respuestas con evidencia



Mejora la alfabetización informacional y el pensamiento crítico del alumnado de Ciencias de la Salud



El GPT docente estandariza el feedback y libera al profesor para valorar la calidad clínica de nivel alto



Modelo transferible a Fisioterapia, Podología y cualquier disciplina sanitaria con mínima adaptación



# La IA no sustituye el pensamiento crítico.

Lo potencia — si le ponemos las reglas correctas.



## Tarjeta-Prompt

campus virtual (Moodle)



## Trazabilidad

replicable · auditable



## GPT + Profesor

verificación doble



## Transferible

cualquier disciplina sanitaria



UNIVERSIDAD COMPLUTENSE DE MADRID  
Grado en Enfermería · Fac. Enfermería, Fisioterapia y Podología

**Prof. Francisco José García González, PhD**

Materiales disponibles: Tarjeta-Prompt · Rúbrica completa · Registro trazabilidad · GPT Docente → solicita en la sesión