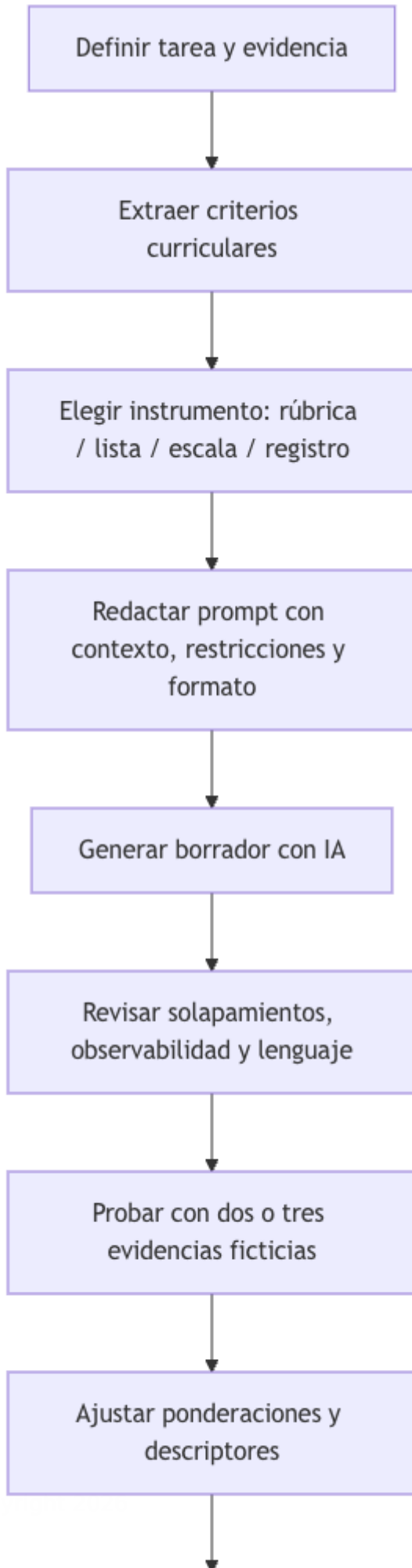


3.1. Consejos para crear buenos prompts

El mejor uso inicial de la IA en evaluación no es la corrección, sino la **creación de instrumentos**. El **orden** correcto es este: **primero se fijan competencias específicas, criterios de evaluación, evidencias y nivel esperado; después se usa la IA para convertir eso en formatos operativos**. Si se trabaja al revés —pidiéndole a un chatbot “hazme una rúbrica” sin criterio previo— lo habitual es obtener textos aparentemente elegantes, pero genéricos, redundantes o mal alineados con lo que realmente se pretende valorar. Suelen ser buenos borradores iniciales, pero requieren revisión y adaptación docente sustantiva.





Aránzazu Cortés + Chat GPT

La fase crítica no es la generación, sino la **prueba con evidencias ficticias o modelos de respuesta**, porque ahí se detecta si la rúbrica distingue de verdad entre niveles. El uso de few-shot y salidas estructuradas mejora mucho esa iteración.

Por lo tanto, cuando pedimos a una IA que genere un instrumento de evaluación, el resultado depende en gran medida de la calidad del prompt. Un buen prompt no debe limitarse a “hazme una rúbrica”, sino que debe contextualizar la tarea, el nivel educativo, la materia, el tipo de evidencia y el uso que se dará al instrumento. Cuanta más información útil reciba la IA, más ajustada será la propuesta.

Conviene incorporar cinco elementos: **contexto didáctico** (etapa educativa, materia, nivel del alumnado), **criterios o competencias, tipo de evidencia a valorar, formato del instrumento y restricciones**. También es útil indicar el **tono** si el instrumento será **para alumnado, profesorado o ambos**, porque el lenguaje cambia mucho según el destinatario. Si el docente no acota, la IA tiende a generar instrumentos genéricos, correctos en apariencia pero poco útiles para una situación concreta.

Cuanto más precisa sea la instrucción, más útil será la propuesta de la IA. También conviene pedir siempre que el resultado sea revisable y que no invente criterios fuera del currículo.

La fórmula más fiable para profesorado suele ser esta:

Rol + tarea + currículo + evidencia + restricciones + formato + control de calidad.

Ejemplo de patrón:

“ Actúa como especialista en evaluación educativa. Diseña una rúbrica analítica para 2.º ESO, Geografía e Historia, para evaluar un podcast de 3 minutos. Usa 4 criterios alineados con [criterios del currículo], 4 niveles de desempeño, lenguaje comprensible para alumnado y profesorado, descriptores observables y no redundantes. No inventes criterios ajenos al currículo. Devuelve una tabla y un JSON.

Estructura útil de prompt

- Contexto: curso, materia y producto final.
- Objetivo de evaluación: qué se quiere valorar.
- Criterios o competencias.



- Formato deseado: rúbrica, lista de cotejo, escala, registro.
- Nivel de detalle.
- Restricciones: lenguaje simple, sin datos personales, alineado con criterios curriculares.

Ejemplos de prompts bien formulados:

“ Crea una rúbrica analítica para 3.º ESO de Lengua Castellana para evaluar una exposición oral de 3 minutos. Incluye 4 criterios, 4 niveles de desempeño, lenguaje claro para alumnado, y relación con comunicación oral, organización de ideas y uso de recursos expresivos.

“ Genera una rúbrica analítica para 4.º de ESO de Biología para evaluar un informe de laboratorio. Debe incluir 4 criterios, 4 niveles de logro, lenguaje claro para alumnado, y centrarse en presentación de datos, interpretación, vocabulario científico y conclusión.

“ Crea una lista de cotejo para Primaria, 5.º curso, para valorar una exposición oral sobre animales vertebrados. Debe ser breve, observable y útil para observación rápida en clase.

La ventaja de este enfoque es que la IA no improvisa, sino que trabaja dentro de un marco pedagógico definido por el docente. Esto favorece la coherencia curricular y evita instrumentos demasiado abstractos o desalineados con la tarea real.

La IA en educación debe usarse con supervisión docente, con protección de datos y dentro de un enfoque ético y pedagógico. Además, en el ámbito de la evaluación educativa los sistemas de IA pueden ser especialmente sensibles cuando afectan a resultados de aprendizaje o decisiones sobre alumnado.

Una regla muy simple y muy técnica a la vez: **identidad, instrucciones, ejemplos y contexto.**

Continuando con la creación de prompts eficaces, además, se recomienda estructurarlos con secciones claras, usar encabezados o etiquetas, incluir ejemplos diversos de entrada/salida y



añadir el contexto relevante al final; además, cuando el resultado tiene que reutilizarse, conviene pedir la salida en formato estructurado, por ejemplo JSON o tabla.

Un prompt eficaz no dice solo “créame una rúbrica”, sino algo como:

“ Actúa como especialista en evaluación competencial de 2.º de ESO. A partir de este criterio, genera una rúbrica analítica de cuatro niveles, con indicadores observables, lenguaje comprensible para alumnado y familias, y una versión final en tabla y otra en JSON.

Cuando además se añaden dos o tres ejemplos buenos y malos de producto del alumnado, la IA suele discriminar mucho mejor entre niveles de desempeño. La documentación oficial sobre prompting denomina a esto **few-shot learning** y recomienda que los ejemplos cubran una gama diversa de situaciones para que el patrón salga más robusto.

Para que estos instrumentos salgan bien con IA, hay varias reglas prácticas que merece la pena enseñar. La primera es **un criterio, una dimensión**: si un descriptor mezcla “contenido, expresión y creatividad”, la IA tenderá a producir niveles ambiguos y el profesorado tendrá más desacuerdo al corregir. La segunda es **lenguaje observable**: mejor “justifica su respuesta con dos evidencias del texto” que “demuestra comprensión profunda”. La tercera es **anclaje por evidencias**: pedir expresamente que cada nivel incluya señales observables o errores típicos. La cuarta es **portabilidad**: solicitar a la IA versión larga para el profesorado, versión breve para el alumnado y versión tabular o JSON para reutilizarla en hojas de cálculo, formularios o LMS. La recomendación de pedir resultados estructurados está alineada con la práctica oficial de “Structured Outputs”.

¿Qué instrumento conviene en cada caso?



Instrumento	Cuándo usarlo	Qué aporta la IA	Riesgo típico
Rúbrica analítica	Proyectos, escritos, exposiciones, tareas complejas	Redacción de criterios, niveles, versiones para alumnado y profesorado	Descriptores vagos o solapados
Lista de cotejo	Requisitos mínimos, procedimientos, entregables	Convertir criterios en ítems observables tipo sí/no	Reducir calidad a mera presencia
Escala de valoración	Participación, cooperación, hábitos, frecuencia/intensidad	Afinar indicadores graduales y anclas verbales	Ambigüedad en "a veces", "frecuente", etc.
Registro de observación	Procesos, oralidad, laboratorio, trabajo cooperativo	Clasificar observaciones y resumir patrones	Convertir anécdotas en diagnósticos excesivos

La recomendación práctica es usar **rúbricas** para tareas complejas, **listas de cotejo** para control mínimo de calidad, **escalas** para procesos graduales y **registros de observación** para evidencias no capturadas en productos finales. Esa variedad encaja mejor con un enfoque competencial y evita sobrecargar una sola herramienta.

Revision #11

Created 2026-04-28 12:01:39 CEST by Arantxa Cortés

Updated 2026-06-22 22:33:01 CEST by Arantxa Cortés