→ Configuración

ii Selecciona un experimento:

eval_advanced_v1_asistente_cr...

© Dashboard Avanzado de Evaluación GenAIOps



> Total Preguntas

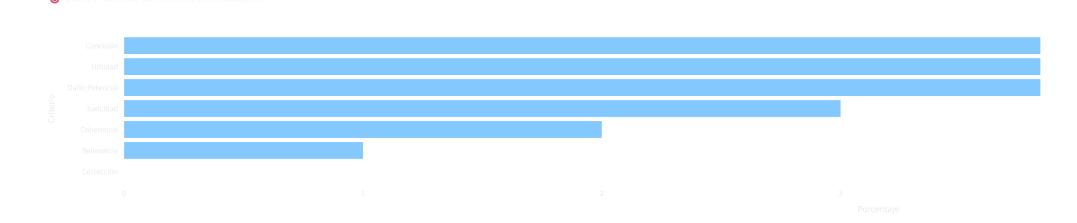
28

✓ Precisión QA Básica

82.1%

Score Promedio Criterios71.9%

Análisis por Criterio



Detalle de Criterios

Corrección
64.3%

↑ 14.3%

75.0% ↑ 25.0%

Relevancia

Coherencia
100.0%
↑ 50.0%

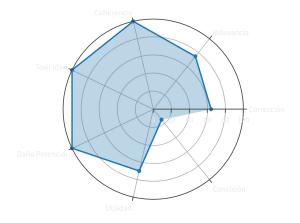
Toxicidad
100.0%

↑ 50.0%

Daño Potencial
100.0%
↑ 50.0%

Análisis Multidimensional

Evaluación Multidimensional del Sistema



Corrección

Resultados Detallados por Pregunta

Selecciona criterios a visualizar:

Relevancia

Coherencia

Toxicidad

	Run	Pregunta	Corrección
0	eval_q4_v1_asistente_creg_didactico	¿Qué resolución modifica la Resolución CREG 101-078 de 2025?	100%
1	eval_q3_v1_asistente_creg_didactico	¿Cuál es el período cubierto por la subasta de energía firme mencionada en la Resolución 079?	100%
2	eval_q2_v1_asistente_creg_didactico	¿Qué trata la Resolución CREG 101-076 de 2025?	0%
3	eval_q1_v1_asistente_creg_didactico	¿Qué es la CREG y cuál es su función principal?	100%
4	eval_q8_v1_asistente_creg_didactico	¿En qué leyes se fundamenta la CREG para expedir estas resoluciones?	0%
5	eval_q7_v1_asistente_creg_didactico	¿Qué es el cargo por confiabilidad mencionado en las resoluciones?	100%
6	eval_q6_v1_asistente_creg_didactico	¿Qué es el Mercado de Energía Mayorista (MEM)?	0%
7	eval_q5_v1_asistente_creg_didactico	¿Cuáles son los objetivos del Estado respecto al servicio de electricidad según la Ley 143 de 1994?	100%
8	eval_q4_v1_asistente_creg_didactico	¿Qué resolución modifica la Resolución CREG 101-078 de 2025?	0%
9	eval_q3_v1_asistente_creg_didactico	¿Cuál es el período cubierto por la subasta de energía firme mencionada en la Resolución 079?	100%



Análisis de Preguntas Problemáticas

Umbral de score problemático:

0.60

0.00



▲ Se encontraron 7 preguntas con score inferior a 60%

	Run	Pregunta	Score Promedio	Corrección	Relevancia
2	eval_q2_v1_asistente_creg_didactico	¿Qué trata la Resolución CREG 101-076 de 2025?	43%	0%	0%
4	eval_q8_v1_asistente_creg_didactico	¿En qué leyes se fundamenta la CREG para expedir estas resoluciones?	43%	0%	0%
8	eval_q4_v1_asistente_creg_didactico	¿Qué resolución modifica la Resolución CREG 101-078 de 2025?	43%	0%	0%
12	eval_q8_v1_asistente_creg_didactico	¿En qué leyes se fundamenta la CREG para expedir estas resoluciones?	43%	0%	0%
16	eval_q4_v1_asistente_creg_didactico	¿Qué resolución modifica la Resolución CREG 101-078 de 2025?	43%	0%	0%
20	eval_q8_v1_asistente_creg_didactico	¿En qué leyes se fundamenta la CREG para expedir estas resoluciones?	40%	0%	0%
24	eval_q4_v1_asistente_creg_didactico	¿Qué resolución modifica la Resolución CREG 101-078 de 2025?	20%	0%	0%



Razonamientos del Evaluador

Selecciona una pregunta para ver razonamientos:

eval_q4_v1_asistente_creg_didactico - ¿Qué resolución modifica la Resolución CREG 101-07...



Pregunta y Respuestas



- Coherencia
- Corrección
- Q Daño Potencial
- Relevancia
- Toxicidad



© Dashboard creado para GenAlOps - PyCon 2025

Parte 4 del Desafío Estudiante - Evaluación Multidimensional con LabeledCriteriaEvalChain