

Diccionario de Datos Sistema de Preguntas ICFES

Sistema de Evaluación Educativa

17 de agosto de 2025

Índice

1. Introducción	6
2. Campos de Identificación Básica	6
2.1. Campo 1: ID_Pregunta	6
2.2. Campo 2: Área_Evaluada	6
2.3. Campo 3: Competencia	7
2.4. Campo 4: Componente	8
2.5. Campo 5: Proceso_Cognitivo	8
2.6. Campo 6: Tipo_Conocimiento	9
3. Características Psicométricas	9
3.1. Campo 7: Nivel_Dificultad	9
3.2. Campo 41: Índice_Discriminación	9
3.3. Campo 42: Parámetro_IRT_A	9
3.4. Campo 43: Parámetro_IRT_B	9
3.5. Campo 44: Parámetro_IRT_C	10
4. Contenido Pedagógico	10
4.1. Campo 8: Afirmación	10
4.2. Campo 9: Evidencia	10
4.3. Campo 10: Pregunta	10
4.4. Campo 11: Respuesta_Correcta	10
4.5. Campo 12: Nivel_Desempeño_Esperado	10
4.6. Campo 13: Tiempo_Estimado	11
4.7. Campo 14: Tema_Específico	11
4.8. Campo 45: Subtema	11
4.9. Campo 15: Grado_Escolar	12
4.10. Campo 16: Periodo_Aplicación	12
5. Opciones de Respuesta	12
5.1. Campo 17: Opcion_A	12
5.2. Campo 18: Opcion_B	12
5.3. Campo 19: Opcion_C	12
5.4. Campo 20: Opcion_D	12
6. Recursos Visuales	13
6.1. Campo 21: Requiere_Imagen	13
6.2. Campo 22: Imagen_Pregunta_URL	13
6.3. Campo 23: Imagen_Opcion_A_URL	13
6.4. Campo 24: Imagen_Opcion_B_URL	13
6.5. Campo 25: Imagen_Opcion_C_URL	13
6.6. Campo 26: Imagen_Opcion_D_URL	13

7. Gamificación y Ayuda	13
7.1. Campo 27: Puntos_XP	13
7.2. Campo 28: Pista_1	14
7.3. Campo 29: Pista_2	14
7.4. Campo 30: Pista_3	14
7.5. Campo 31: Explicación_Respuesta	14
7.6. Campo 32: Error_Común	14
8. Contexto Compartido	14
8.1. Campo 33: Pregunta_con_contexto	14
8.2. Campo 34: Pregunta_Libro	14
8.3. Campo 35: Orden_En_Contexto	15
8.4. Campo 36: Contexto_Requerido	15
8.5. Campo 37: Imagen_Contexto_Comp	15
8.6. Campo 38: ID_Contexto_Compartido	15
9. Gestión de Archivos	15
9.1. Campo 39: Ruta_absoluta_Archivo	15
9.2. Campo 40: Nombre_del_Archivo	15
10. Campos Transversales de Análisis	15
10.1. Campo 46: Estrategia_Discursiva	15
10.2. Campo 47: Tipo_Razonamiento	16
10.3. Campo 48: Complejidad_Cognitiva	16
10.4. Campo 49: Contexto_Aplicación	16
11. Campos Específicos por Materia	16
11.1. Lectura Crítica	16
11.1.1. Campo 50: Tipo_Texto	16
11.1.2. Campo 51: Género_Textual	16
11.1.3. Campo 52: Función_Comunicativa	16
11.2. Matemáticas	17
11.2.1. Campo 53: Pensamiento_Matemático	17
11.3. Ciencias Naturales	17
11.3.1. Campo 54: Disciplina_Predominante	17
11.3.2. Campo 55: Concepto_Científico	17
11.3.3. Campo 56: Proceso_Científico	17
11.3.4. Campo 57: Nivel_Representación	17
11.4. Ciencias Sociales	17
11.4.1. Campo 58: Periodo_Histórico	17
11.4.2. Campo 59: Ámbito_Análisis	17
11.4.3. Campo 60: Escala_Espacial	18
11.4.4. Campo 61: Tipo_Fuente	18
11.5. Inglés	18
11.5.1. Campo 62: Nivel_MCER	18
11.5.2. Campo 63: Parte_Prueba	18
11.5.3. Campo 64: Competencia_Comunicativa	18

11.5.4. Campo 65: Habilidad_Específica	19
12. Análisis de Distractores	19
12.1. Campo 66: Distractor_A_Concepto	19
12.2. Campo 67: Distractor_B_Concepto	19
12.3. Campo 68: Distractor_C_Concepto	19
12.4. Campo 69: Frecuencia_Error_A	19
12.5. Campo 70: Frecuencia_Error_B	19
12.6. Campo 71: Frecuencia_Error_C	20
13. Rutas de Aprendizaje	20
13.1. Campo 72: Concepto_Refuerzo_1	20
13.2. Campo 73: Concepto_Refuerzo_2	20
13.3. Campo 74: Recurso_Sugerido	20
13.4. Campo 75: Tiempo_Estudio_Estimado	20
14. Campos de Auditoría	20
14.1. Campo 76: Fecha_Creación	20
14.2. Campo 77: Fecha_Actualización	21
14.3. Campo 78: Versión_Item	21
14.4. Campo 79: Banco_Origen	21
14.5. Campo 80: Fecha_Calibración	21
15. Reglas de Aplicación por Área	21
15.1. Lectura Crítica	21
15.2. Matemáticas	21
15.3. Ciencias Naturales	22
15.4. Ciencias Sociales	22
15.5. Inglés	22
16. Ejemplos de Llenado por Área	22
16.1. Ejemplo: Lectura Crítica	22
16.2. Ejemplo: Matemáticas	23
16.3. Ejemplo: Ciencias Naturales	23
16.4. Ejemplo: Inglés	23
17. Instructivo de Prompts para Automatización	23
17.1. Uso de NotebookLM para Extracción de Preguntas	23
17.1.1. Objetivo	23
17.1.2. Proceso Recomendado	23
17.1.3. Prompt para NotebookLM	24
17.2. Uso de Claude para Generación de Formato de 80 Columnas	25
17.2.1. Objetivo	25
17.2.2. Proceso Recomendado	25
17.2.3. Prompt para Claude	25

18. Flujo de Trabajo Automatizado	28
18.1. Proceso Completo NotebookLM → Claude	28
18.1.1. Paso 1: Preparación en NotebookLM	28
18.1.2. Paso 2: Conversión en Claude	28
18.1.3. Paso 3: Validación Final	28
18.2. Beneficios del Flujo Automatizado	28
19. Notas Importantes	29

1 Introducción

Este diccionario de datos define la estructura, tipos de datos, valores válidos y reglas de llenado para la base de datos del sistema de preguntas ICFES. Contiene 80 campos organizados por categorías funcionales y áreas de evaluación específicas.

Áreas de Evaluación Contempladas:

- Lectura Crítica
- Matemáticas
- Ciencias Naturales
- Ciencias Sociales
- Inglés

2 Campos de Identificación Básica

2.1 Campo 1: ID_Pregunta

- **Tipo de Dato:** VARCHAR(50) - CLAVE PRIMARIA
- **Obligatorio:** SÍ
- **Formato:** Código único alfanumérico
- **Patrón sugerido:** [ÁREA]-[AÑO]-[NÚMERO_SECUENCIAL]
- **Ejemplos:** LC_2024_001, MAT_2024_045, CN_2024_123
- **Descripción:** Identificador único e inequívoco para cada pregunta del sistema

2.2 Campo 2: Área_Evaluada

- **Tipo de Dato:** VARCHAR(30)
- **Obligatorio:** SÍ
- **Valores válidos:**
 - Lectura Crítica
 - Matemáticas
 - Ciencias Naturales
 - Ciencias Sociales
 - Inglés
- **Descripción:** Área de conocimiento principal que evalúa la pregunta

2.3 Campo 3: Competencia

- **Tipo de Dato:** VARCHAR(100)
- **Obligatorio:** Sí
- **Descripción:** Competencia específica evaluada según el área

Valores por área:

Para Lectura Crítica (exactamente estos 3 valores):

1. Identificar y entender los contenidos locales que conforman un texto
2. Comprender cómo se articulan las partes de un texto para darle un sentido global
3. Reflexionar a partir de un texto y evaluar su contenido

Para Matemáticas (con distribución porcentual):

- Interpretación y representación (34 %)
- Formulación y ejecución (43 %)
- Argumentación (23 %)

Para Ciencias Naturales (con distribución aproximada):

- Indagación (40 %)
- Uso comprensivo del conocimiento científico (30 %)
- Explicación de fenómenos (30 %)

Para Ciencias Sociales:

- Pensamiento social
- Interpretación y análisis de perspectivas
- Pensamiento reflexivo y sistémico

Para Inglés:

- Lingüística
- Pragmática
- Sociolingüística

2.4 Campo 4: Componente

- **Tipo de Dato:** VARCHAR(50)
- **Obligatorio:** Según área
- **Descripción:** Componente específico dentro del área evaluada

Para Matemáticas:

- Álgebra_y_cálculo
- Geometría
- Estadística_y_probabilidad

Para Ciencias Naturales (33.3 % cada uno):

- Entorno_vivo
- Entorno_físico
- Ciencia_tecnología_sociedad

Para Ciencias Sociales:

- Historia (40 % de la evaluación de Sociales)
- Geografía
- Constitución_política_democracia
- Competencias_ciudadanas (60 % del total de la evaluación)

2.5 Campo 5: Proceso_Cognitivo

- **Tipo de Dato:** VARCHAR(30)
- **Obligatorio:** NO
- **Descripción:** Proceso mental requerido para resolver la pregunta

Para Lectura Crítica:

- Comprensión_literal
- Comprensión_inferencial
- Comprensión_crítica
- Análisis_estructural
- Evaluación_argumentativa

Para Ciencias Naturales:

- Observación
- Formulación hipótesis
- Experimentación
- Análisis_datos
- Comunicación_resultados

2.6 Campo 6: Tipo_Conocimiento

- **Tipo de Dato:** VARCHAR(30)
- **Obligatorio:** Para Matemáticas
- **Valores válidos:** Genérico, No_genérico, Procedimental, Conceptual
- **Descripción:** Clasificación del tipo de conocimiento evaluado

3 Características Psicométricas

3.1 Campo 7: Nivel_Dificultad

- **Tipo de Dato:** DECIMAL(4,3)
- **Rango válido:** 0.000 - 1.000
- **Descripción:** Índice de dificultad psicométrica del ítem
- **Ejemplo:** 0.750 significa 75 % de dificultad

3.2 Campo 41: Índice_Discriminación

- **Tipo de Dato:** DECIMAL(4,3)
- **Rango válido:** -1.000 - 1.000
- **Descripción:** Capacidad del ítem para discriminar entre estudiantes de alto y bajo rendimiento

3.3 Campo 42: Parámetro_IRT_A

- **Tipo de Dato:** DECIMAL(4,3)
- **Descripción:** Parámetro de discriminación en Teoría de Respuesta al Ítem

3.4 Campo 43: Parámetro_IRT_B

- **Tipo de Dato:** DECIMAL(4,3)
- **Descripción:** Parámetro de dificultad en Teoría de Respuesta al Ítem

3.5 Campo 44: Parámetro_IRT_C

- **Tipo de Dato:** DECIMAL(4,3)
- **Descripción:** Parámetro de pseudoadivinanza en Teoría de Respuesta al Ítem

4 Contenido Pedagógico

4.1 Campo 8: Afirmación

- **Tipo de Dato:** TEXT
- **Descripción:** Enunciado general que contextualiza la competencia evaluada
- **Recomendación:** Máximo 15 palabras para síntesis efectiva

4.2 Campo 9: Evidencia

- **Tipo de Dato:** TEXT
- **Descripción:** Indicador específico de lo que debe demostrar el estudiante
- **Recomendación:** Máximo 15 palabras para claridad conceptual

4.3 Campo 10: Pregunta

- **Tipo de Dato:** TEXT
- **Obligatorio:** Sí
- **Descripción:** Enunciado completo de la pregunta tal como aparece al estudiante

4.4 Campo 11: Respuesta_Correcta

- **Tipo de Dato:** CHAR(1)
- **Obligatorio:** Sí
- **Valores válidos:** A, B, C, D
- **Descripción:** Letra que identifica la opción correcta

4.5 Campo 12: Nivel_Desempeño_Esperado

- **Tipo de Dato:** VARCHAR(20)
- **Descripción:** Nivel de desempeño esperado según el área

Niveles estándar ICFES (todas las áreas):

- Insuficiente: 0-35 puntos

- Mínimo: 36-50 puntos
- Satisfactorio: 51-70 puntos
- Avanzado: 71-100 puntos

Niveles específicos para Matemáticas:

- Nivel_1: 0-35 puntos
- Nivel_2: 36-50 puntos
- Nivel_3: 51-70 puntos
- Nivel_4: 71-100 puntos

Niveles específicos para Inglés (MCER):

- A-: 0-47 puntos
- A1: 48-57 puntos
- A2: 58-67 puntos
- B1: 68-78 puntos
- B+: 79-100 puntos

4.6 Campo 13: Tiempo_Estimado

- **Tipo de Dato:** INTEGER
- **Unidad:** Segundos
- **Rango típico:** 60-300 segundos (1-5 minutos)
- **Descripción:** Tiempo estimado de resolución para un estudiante promedio

4.7 Campo 14: Tema_Específico

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Tema principal que evalúa la pregunta
- **Ejemplos:** "Funciones lineales", "Comprensión lectora", "Genética"

4.8 Campo 45: Subtema

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Subtema más específico dentro del tema principal

4.9 Campo 15: Grado_Escolar

- **Tipo de Dato:** VARCHAR(10)
- **Valores típicos:** 11°, 10°, 9°
- **Descripción:** Grado académico objetivo de la pregunta

4.10 Campo 16: Periodo_Aplicación

- **Tipo de Dato:** VARCHAR(20)
- **Formato:** YYYY-N (donde N es el semestre)
- **Ejemplos:** 2024-1, 2024-2, 2025-1
- **Descripción:** Período de aplicación de la prueba

5 Opciones de Respuesta

5.1 Campo 17: Opcion_A

- **Tipo de Dato:** TEXT
- **Descripción:** Texto completo de la primera opción de respuesta

5.2 Campo 18: Opcion_B

- **Tipo de Dato:** TEXT
- **Descripción:** Texto completo de la segunda opción de respuesta

5.3 Campo 19: Opcion_C

- **Tipo de Dato:** TEXT
- **Descripción:** Texto completo de la tercera opción de respuesta

5.4 Campo 20: Opcion_D

- **Tipo de Dato:** TEXT
- **Descripción:** Texto completo de la cuarta opción de respuesta

6 Recursos Visuales

6.1 Campo 21: Requiere_Imagen

- **Tipo de Dato:** BOOLEAN
- **Valores:** TRUE/FALSE
- **Default:** FALSE
- **Descripción:** Indica si la pregunta requiere imagen para su resolución

6.2 Campo 22: Imagen_Pregunta_URL

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL completa o ruta del archivo de imagen asociada a la pregunta

6.3 Campo 23: Imagen_Opcion_A_URL

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL completa o ruta del archivo de imagen de la opción A

6.4 Campo 24: Imagen_Opcion_B_URL

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL completa o ruta del archivo de imagen de la opción B

6.5 Campo 25: Imagen_Opcion_C_URL

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL completa o ruta del archivo de imagen de la opción C

6.6 Campo 26: Imagen_Opcion_D_URL

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL completa o ruta del archivo de imagen de la opción D

7 Gamificación y Ayuda

7.1 Campo 27: Puntos_XP

- **Tipo de Dato:** INTEGER
- **Default:** 10

- **Rango típico:** 5-50 puntos
- **Descripción:** Puntos de experiencia otorgados al resolver correctamente

7.2 Campo 28: Pista_1

- **Tipo de Dato:** TEXT
- **Descripción:** Primera ayuda progresiva para orientar al estudiante

7.3 Campo 29: Pista_2

- **Tipo de Dato:** TEXT
- **Descripción:** Segunda ayuda progresiva, más específica que la primera

7.4 Campo 30: Pista_3

- **Tipo de Dato:** TEXT
- **Descripción:** Tercera ayuda progresiva, casi revelando la respuesta

7.5 Campo 31: Explicación_Respuesta

- **Tipo de Dato:** TEXT
- **Descripción:** Explicación pedagógica detallada de por qué la respuesta es correcta

7.6 Campo 32: Error_Común

- **Tipo de Dato:** TEXT
- **Descripción:** Descripción del error más frecuente que cometen los estudiantes

8 Contexto Compartido

8.1 Campo 33: Pregunta_con_contexto

- **Tipo de Dato:** BOOLEAN
- **Descripción:** Indica si la pregunta requiere un contexto previo para ser resuelta

8.2 Campo 34: Pregunta_Libro

- **Tipo de Dato:** BOOLEAN
- **Descripción:** Indica si forma parte de un conjunto de preguntas agrupadas

8.3 Campo 35: Orden_En_Contexto

- **Tipo de Dato:** INTEGER
- **Descripción:** Orden secuencial de la pregunta dentro del contexto compartido

8.4 Campo 36: Contexto_Requerido

- **Tipo de Dato:** TEXT
- **Descripción:** Texto completo del contexto necesario para responder la pregunta

8.5 Campo 37: Imagen_Contexto_Comp

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** URL o ruta de imagen del contexto compartido entre preguntas

8.6 Campo 38: ID_Contexto_Compartido

- **Tipo de Dato:** VARCHAR(50)
- **Descripción:** Identificador único del grupo de preguntas con contexto común

9 Gestión de Archivos

9.1 Campo 39: Ruta_absoluta_Archivo

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** Ruta completa del archivo en el sistema de almacenamiento

9.2 Campo 40: Nombre_del_Archivo

- **Tipo de Dato:** VARCHAR(255)
- **Descripción:** Nombre original del archivo fuente, incluyendo extensión

10 Campos Transversales de Análisis

10.1 Campo 46: Estrategia_Discursiva

- **Tipo de Dato:** VARCHAR(30)
- **Valores válidos:** Deductiva, Inductiva, Analítica, Sintética, Comparativa
- **Descripción:** Estrategia de razonamiento predominante en la pregunta

10.2 Campo 47: Tipo_Razonamiento

- **Tipo de Dato:** VARCHAR(30)
- **Valores válidos:** Lógico_deductivo, Espacial, Abstracto, Concreto, Analógico
- **Descripción:** Tipo de razonamiento requerido para resolver la pregunta

10.3 Campo 48: Complejidad_Cognitiva

- **Tipo de Dato:** VARCHAR(20)
- **Valores válidos:** Recordar, Comprender, Aplicar, Analizar, Evaluar, Crear
- **Base teórica:** Taxonomía de Bloom revisada
- **Descripción:** Nivel de complejidad cognitiva según la taxonomía educativa

10.4 Campo 49: Contexto_Aplicación

- **Tipo de Dato:** VARCHAR(30)
- **Valores válidos:** Familiar, Laboral, Comunitario, Académico, Científico, Situaciones_cotidianas, Contextos_académicos, Formatos_no_convencionales, Aplicaciones_prácticas
- **Descripción:** Contexto de aplicación del conocimiento evaluado

11 Campos Específicos por Materia

11.1 Lectura Crítica

11.1.1 Campo 50: Tipo_Texto

- **Valores válidos:** Continuo_literario, Continuo_informativo, Continuo_argumentativo, Discontinuo_infografía, Discontinuo_tabla, Discontinuo_gráfico, Mixto
- **Descripción:** Clasificación del tipo de texto presentado

11.1.2 Campo 51: Género_Textual

- **Valores válidos:** Narrativo, Expositivo, Argumentativo, Descriptivo, Instructivo
- **Descripción:** Género literario o discursivo del texto

11.1.3 Campo 52: Función_Comunicativa

- **Valores válidos:** Informar, Persuadir, Explicar, Narrar, Describir
- **Descripción:** Función comunicativa principal del texto

11.2 Matemáticas

11.2.1 Campo 53: Pensamiento_Matemático

- **Valores válidos:** Numérico_variacional, Geométrico_métrico, Aleatorio
- **Descripción:** Tipo de pensamiento matemático predominante

11.3 Ciencias Naturales

11.3.1 Campo 54: Disciplina_Predominante

- **Valores válidos:** Biología, Física, Química, Interdisciplinar_CTS
- **Descripción:** Disciplina científica principal de la pregunta

11.3.2 Campo 55: Concepto_Científico

- **Valores válidos:** Estructura_función_biológica, Homeostasis, Herencia_reproducción, Ecología_ecosistemas, Evolución_diversidad, Cinemática_dinámica, Energía_conservación, Ondas_electromagnetismo, Enlaces_químicos, Reacciones_químicas, Estequiometría
- **Descripción:** Concepto científico principal evaluado

11.3.3 Campo 56: Proceso_Científico

- **Valores válidos:** Observación, Formulación_hipótesis, Experimentación, Análisis_datos, Comunicación_resultados
- **Descripción:** Proceso del método científico evaluado

11.3.4 Campo 57: Nivel_Representación

- **Valores válidos:** Macroscópico, Microscópico, Simbólico
- **Descripción:** Nivel de representación del fenómeno científico

11.4 Ciencias Sociales

11.4.1 Campo 58: Periodo_Histórico

- **Valores válidos:** Precolombino, Colonial, Independencia, República_siglo_XIX, Siglo_XX, Contemporáneo
- **Descripción:** Período histórico al que se refiere la pregunta

11.4.2 Campo 59: Ámbito_Análisis

- **Valores válidos:** Político, Económico, Social, Cultural, Ambiental
- **Descripción:** Ámbito principal de análisis social

11.4.3 Campo 60: Escala_Espacial

- **Valores válidos:** Local, Regional, Nacional, Global
- **Descripción:** Escala geográfica de análisis del fenómeno

11.4.4 Campo 61: Tipo_Fuente

- **Valores válidos:** Primaria_documento, Primaria_testimonio, Secundaria_académica, Iconográfica, Estadística
- **Descripción:** Tipo de fuente histórica o social utilizada

11.5 Inglés

11.5.1 Campo 62: Nivel_MCER

- **Valores válidos y rangos:**
 - A-: 0-47 puntos (Nivel pre-básico)
 - A1: 48-57 puntos (Nivel básico inicial)
 - A2: 58-67 puntos (Nivel básico)
 - B1: 68-78 puntos (Nivel intermedio)
 - B+: 79-100 puntos (Nivel intermedio alto)
- **Descripción:** Nivel según Marco Común Europeo de Referencia

11.5.2 Campo 63: Parte_Prueba

- **Valores válidos y distribución:**
 - Parte_1_avisos (11 %) - Comprensión de avisos y señales
 - Parte_2_vocabulario (11 %) - Conocimiento de vocabulario
 - Parte_3_conversaciones (11 %) - Comprensión de conversaciones
 - Parte_4_gramática_básica (18 %) - Gramática fundamental
 - Parte_5_lectura_literal (16 %) - Comprensión literal
 - Parte_6_lectura_inferencial (11 %) - Comprensión inferencial
 - Parte_7_gramática_avanzada (22 %) - Gramática compleja
- **Descripción:** Parte específica de la prueba de inglés

11.5.3 Campo 64: Competencia_Comunicativa

- **Valores válidos:** Lingüística, Pragmática, Sociolingüística
- **Descripción:** Competencia comunicativa principal evaluada

11.5.4 Campo 65: Habilidad_Específica

- **Valores válidos:** Reading_comprehension, Vocabulary_knowledge, Grammar_usage, Language_use
- **Descripción:** Habilidad específica del idioma evaluada

Función Comunicativa adicional para Inglés: Información_personal, Invitación_sugerencia, Descripción, Instrucción, Argumentación, Expresión_opinión

12 Análisis de Distractores

12.1 Campo 66: Distractor_A_Concepto

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Error conceptual específico asociado al distractor A

12.2 Campo 67: Distractor_B_Concepto

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Error conceptual específico asociado al distractor B

12.3 Campo 68: Distractor_C_Concepto

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Error conceptual específico asociado al distractor C

12.4 Campo 69: Frecuencia_Error_A

- **Tipo de Dato:** DECIMAL(5,2)
- **Rango válido:** 0.00-100.00
- **Descripción:** Porcentaje histórico de selección del distractor A

12.5 Campo 70: Frecuencia_Error_B

- **Tipo de Dato:** DECIMAL(5,2)
- **Rango válido:** 0.00-100.00
- **Descripción:** Porcentaje histórico de selección del distractor B

12.6 Campo 71: Frecuencia_Error_C

- **Tipo de Dato:** DECIMAL(5,2)
- **Rango válido:** 0.00-100.00
- **Descripción:** Porcentaje histórico de selección del distractor C

13 Rutas de Aprendizaje

13.1 Campo 72: Concepto_Refuerzo_1

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Primer concepto que el estudiante debe reforzar

13.2 Campo 73: Concepto_Refuerzo_2

- **Tipo de Dato:** VARCHAR(100)
- **Descripción:** Segundo concepto que el estudiante debe reforzar

13.3 Campo 74: Recurso_Sugerido

- **Tipo de Dato:** VARCHAR(500)
- **Descripción:** Enlaces, libros o referencias de estudio recomendadas

13.4 Campo 75: Tiempo_Estudio_Estimado

- **Tipo de Dato:** INTEGER
- **Unidad:** Horas
- **Descripción:** Tiempo recomendado de práctica para dominar el concepto

14 Campos de Auditoría

14.1 Campo 76: Fecha_Creación

- **Tipo de Dato:** TIMESTAMP
- **Default:** CURRENT_TIMESTAMP
- **Descripción:** Fecha y hora de creación del registro

14.2 Campo 77: Fecha_Actualización

- **Tipo de Dato:** TIMESTAMP
- **Default:** CURRENT_TIMESTAMP
- **Descripción:** Fecha y hora de la última modificación

14.3 Campo 78: Versión_Item

- **Tipo de Dato:** INTEGER
- **Default:** 1
- **Descripción:** Número de versión del ítem para control de cambios

14.4 Campo 79: Banco_Origen

- **Tipo de Dato:** VARCHAR(50)
- **Ejemplos:** ICFES_2024, SIMULACRO_INSTITUCIONAL, BANCO_PRUEBAS
- **Descripción:** Identificador del banco de preguntas de origen

14.5 Campo 80: Fecha_Calibración

- **Tipo de Dato:** DATE
- **Descripción:** Fecha de calibración de los parámetros psicométricos

15 Reglas de Aplicación por Área

15.1 Lectura Crítica

- **Campos obligatorios específicos:** Tipo_Texto, Género_Textual, Función_Comunicativa
- **Campos opcionales:** Proceso_Cognitivo
- **Nivel de desempeño:** Usar estándar ICFES (Insuficiente, Mínimo, Satisfactorio, Avanzado)
- **Competencias:** Usar exactamente los 3 valores especificados

15.2 Matemáticas

- **Campos obligatorios específicos:** Componente, Pensamiento_Matemático, Tipo_Conocimiento
- **Nivel de desempeño:** Usar Nivel_1 a Nivel_4 en lugar de estándar
- **Distribución de competencias:** Respetar porcentajes (34 %, 43 %, 23 %)
- **Contexto de aplicación:** Priorizar valores específicos de Matemáticas

15.3 Ciencias Naturales

- **Campos obligatorios específicos:** Componente, Disciplina_Predominante, Proceso_Científico
- **Campos opcionales:** Concepto_Científico, Nivel_Representación
- **Distribución de componentes:** 33.3 % cada uno
- **Distribución de competencias:** 40 %, 30 %, 30 %

15.4 Ciencias Sociales

- **Campos obligatorios específicos:** Componente, Ámbito_Análisis, Escala_Espacial
- **Campos opcionales:** Periodo_Histórico, Tipo_Fuente
- **Distribución:** Historia 40 %, Competencias ciudadanas 60 %
- **Competencias:** Usar las 3 específicas de Ciencias Sociales

15.5 Inglés

- **Campos obligatorios específicos:** Nivel_MCER, Parte_Prueba, Competencia_Comunicativa, Habilidad_Específica
- **Nivel de desempeño:** Usar niveles MCER (A-, A1, A2, B1, B+) en lugar de estándar
- **Distribución de partes:** Respetar porcentajes específicos de cada parte
- **Función comunicativa:** Usar valores adicionales específicos para Inglés

16 Ejemplos de Llenado por Área

16.1 Ejemplo: Lectura Crítica

ID_Pregunta: "LC_2024_001"

Área_Evaluada: "Lectura Crítica"

Competencia: "Reflexionar a partir de un texto y evaluar su contenido"

Tipo_Texto: "Continuo_argumentativo"

Género_Textual: "Argumentativo"

Función_Comunicativa: "Persuadir"

Proceso_Cognitivo: "Comprensión_crítica"

Nivel_Desempeño_Esperado: "Satisfactorio"

16.2 Ejemplo: Matemáticas

ID_Pregunta: "MAT_2024_045"
Área_Evaluada: "Matemáticas"
Competencia: "Formulación y ejecución"
Componente: "Álgebra_y_cálculo"
Pensamiento_Matemático: "Numérico_variacional"
Tipo_Conocimiento: "Procedimental"
Nivel_Desempeño_Esperado: "Nivel_3"

16.3 Ejemplo: Ciencias Naturales

ID_Pregunta: "CN_2024_123"
Área_Evaluada: "Ciencias Naturales"
Competencia: "Explicación de fenómenos"
Componente: "Entorno_vivo"
Disciplina_Predominante: "Biología"
Concepto_Científico: "Herencia_reproducción"
Proceso_Científico: "Análisis_datos"
Nivel_Representación: "Microscópico"

16.4 Ejemplo: Inglés

ID_Pregunta: "ING_2024_089"
Área_Evaluada: "Inglés"
Competencia: "Pragmática"
Nivel_MCER: "B1"
Parte_Prueba: "Parte_5_lectura_literal"
Competencia_Comunicativa: "Pragmática"
Habilidad_Específica: "Reading_comprehension"
Función_Comunicativa: "Información_personal"

17 Instructivo de Prompts para Automatización

17.1 Uso de NotebookLM para Extracción de Preguntas

17.1.1 Objetivo

Facilitar la extracción y estructuración inicial de preguntas desde documentos fuente utilizando NotebookLM para el procesamiento automático de contenido educativo.

17.1.2 Proceso Recomendado

1. Cargar documentos fuente en NotebookLM
2. Aplicar prompt de extracción específico
3. Obtener estructura preliminar de preguntas

4. Transferir a sistema principal para completar metadatos

17.1.3 Prompt para NotebookLM

****EXTRACTOR AUTOMÁTICO DE PREGUNTAS ICFES - NOTEBOOKLM****

****INSTRUCCIÓN PRINCIPAL:****

Analiza todos los documentos cargados y extrae TODAS las preguntas de evaluación encontradas. Organiza la información en bloques de máximo 25 preguntas por respuesta, siguiendo el formato exacto especificado.

****METODOLOGÍA DE EXTRACCIÓN:****

1. ****IDENTIFICACIÓN AUTOMÁTICA:****

- Localiza todas las preguntas numeradas o con formato de evaluación
- Detecta automáticamente contextos compartidos (ej: "Responde las preguntas 1 a 3 de acuerdo con...")
- Identifica elementos visuales asociados (tablas, gráficos, figuras, mapas)
- Extrae todas las opciones de respuesta (A, B, C, D)
- Busca respuestas correctas y argumentos del autor

2. ****DETECCIÓN DE CONTEXTOS COMPARTIDOS:****

- Frases clave: "Responde las preguntas X a Y", "De acuerdo con la siguiente información", "Basándote en el texto anterior"
- Asigna numeración secuencial dentro del contexto (1, 2, 3...)
- Preguntas independientes: siempre orden 1

3. ****DETECCIÓN DE DEPENDENCIA VISUAL:****

- Busca referencias a: "según la gráfica", "en la tabla", "en la figura", "observa el mapa", "según la imagen"
- Marca como "Verdadero" si hay dependencia visual, "Falso" si no

****FORMATO DE SALIDA EXACTO:****

****Fuente:**** [Nombre exacto del archivo PDF]

● ****Pregunta X****

- ****Pregunta:**** [Texto completo de la pregunta SIN modificar]
- ****Opciones de Respuesta:**** A. [opción A] B. [opción B] C. [opción C] D. [opción D]
- ****Respuesta Correcta:**** [Letra correcta].
- ****Argumento del Autor:**** [Si está disponible, caso contrario: "No disponible"]
- ****Área del Conocimiento:**** [Detectar automáticamente: Lectura Crítica, Matemáticas, Ciencias Naturales, Ciencias Sociales, Inglés, Competencias Ciudadanas]
- ****Depende de Imagen:**** [Verdadero/Falso]

- ****Contexto Compartido:**** [Verdadero/Falso]
- ****Orden de Contexto Compartido:**** [1,2,3... o solo 1 si no hay contexto]

****REGLAS CRÍTICAS:****

- NO modifiques el texto de las preguntas
- NO alteres las opciones de respuesta
- Extrae TODO sin omitir preguntas
- Máximo 25 preguntas por bloque de respuesta
- Si hay más de 25, continúa en siguiente bloque numerando secuencialmente
- Preserva numeración original de las preguntas del documento

****PROCESAMIENTO INTELIGENTE:****

- Identifica patrones de formato consistentes en el documento
- Si una pregunta hace referencia a contenido anterior, marca contexto compartido
- Extrae preguntas en el orden que aparecen, sin asumir numeración consecutiva

¿Estás listo para extraer todas las preguntas del documento cargado?

17.2 Uso de Claude para Generación de Formato de 80 Columnas

17.2.1 Objetivo

Automatizar la generación del formato completo de 80 columnas para preguntas IC-FES, garantizando consistencia y completitud de metadatos según el área de evaluación.

17.2.2 Proceso Recomendado

1. Proporcionar pregunta base y área de evaluación
2. Aplicar prompt de estructuración específico
3. Obtener formato completo con todos los campos
4. Revisar y ajustar según necesidades específicas

17.2.3 Prompt para Claude

****CONVERSION AUTOMÁTICO A FORMATO 80 COLUMNAS ICFES - CLAUDE AI****

****INSTRUCCIÓN PRINCIPAL:****

Convierte el bloque de preguntas extraído por NotebookLM al formato estándar de 80 columnas ICFES. Procesa cada pregunta individualmente con análisis psicométrico y pedagógico inteligente.

****ENTRADA ESPERADA:****

Recibirás bloques de preguntas en formato NotebookLM con estructura:

- Fuente, Pregunta, Opciones, Respuesta Correcta, Argumento del Autor, etc.

****SALIDA REQUERIDA:****

Tabla con exactamente 80 columnas separadas por tabulaciones, una fila por pregunta.

****ENCABEZADOS DE COLUMNAS (orden exacto):****

ID_Pregunta Área_Evaluada Competencia Componente Proceso_Cognitivo
 Tipo_Conocimiento Nivel_Dificultad Afirmación Evidencia Pregunta
 Respuesta_Correcta Nivel_Desempeño_Esperado Tiempo_Estimado
 Tema_Específico Grado_Escolar Periodo_Aplicación Opcion_A Opcion_B
 Opcion_C Opcion_D Requiere_Imagen Imagen_Pregunta_URL
 Imagen_Opcion_A_URL Imagen_Opcion_B_URL Imagen_Opcion_C_URL
 Imagen_Opcion_D_URL Puntos_XP Pista_1 Pista_2 Pista_3
 Explicación_Respuesta Error_Común Pregunta_con_contexto Pregunta_Libro
 Orden_En_Contexto Contexto_Requerido Imagen_Contexto_Comp
 ID_Contexto_Compartido Ruta_absoluta_Archivo Nombre_del_Archivo
 Índice_Discriminación Parámetro_IRT_A Parámetro_IRT_B Parámetro_IRT_C
 Subtema Estrategia_Discursiva Tipo_Razonamiento Complejidad_Cognitiva
 Contexto_Aplicación Tipo_Texto Género_Textual Función_Comunicativa
 Pensamiento_Matemático Disciplina_Predominante Concepto_Científico
 Proceso_Científico Nivel_Representación Periodo_Histórico
 Ámbito_Análisis Escala_Espacial Tipo_Fuente Nivel_MCER Parte_Prueba
 Competencia_Comunicativa Habilidad_Específica Distractor_A_Concepto
 Distractor_B_Concepto Distractor_C_Concepto Frecuencia_Error_A
 Frecuencia_Error_B Frecuencia_Error_C Concepto_Refuerzo_1
 Concepto_Refuerzo_2 Recurso_Sugerido Tiempo_Estudio_Estimado
 Fecha_Creación Fecha_Actualización Versión_Item Banco_Origen
 Fecha_Calibración

****LÓGICA DE CONVERSIÓN AUTOMÁTICA:****1. ****PRESERVACIÓN DE DATOS ORIGINALES:****

- Pregunta: Conservar texto exacto de NotebookLM
- Opciones A,B,C,D: Mantener opciones exactas
- Respuesta_Correcta: Usar letra proporcionada
- Área_Evaluada: Usar "Área del Conocimiento" de NotebookLM
- Requiere_Imagen: Convertir "Depende de Imagen" a TRUE/FALSE
- Pregunta_con_contexto: Convertir "Contexto Compartido" a TRUE/FALSE
- Orden_En_Contexto: Usar "Orden de Contexto Compartido"
- Nombre_del_Archivo: Usar "Fuente" de NotebookLM

2. ****GENERACIÓN AUTOMÁTICA ID_Pregunta:****

- Formato: [ÁREA_ABREV]_2025_[NUM_SECUENCIAL]
- LC_2025_001 (Lectura Crítica)
- MAT_2025_001 (Matemáticas)
- CN_2025_001 (Ciencias Naturales)

- CS_2025_001 (Ciencias Sociales)
- ING_2025_001 (Inglés)

3. **ANÁLISIS INTELIGENTE POR ÁREA:**

CIENCIAS SOCIALES:

- Competencia: Analizar contenido → "Pensamiento social" | "Interpretación y análisis de perspectivas" | "Pensamiento reflexivo y sistémico"
- Componente: Detectar tema → Historia | Geografía | Constitución_política_democracia Competencias_ciudadanas
- Ámbito_Análisis: Político, Económico, Social, Cultural, Ambiental
- Escala_Espacial: Local, Regional, Nacional, Global

LECTURA CRÍTICA:

- Competencia: "Identificar y entender los contenidos locales" | "Comprender articulación" | "Reflexionar y evaluar"
- Tipo_Texto: Continuo_literario, Continuo_informativo, etc.
- Género_Textual: Narrativo, Expositivo, Argumentativo, etc.

4. **CAMPOS CALCULADOS INTELIGENTEMENTE:**

- Nivel_Dificultad: 0.300-0.900 según complejidad verbal y conceptual
- Tiempo_Estimado: 60-180 segundos según longitud y complejidad
- Índice_Discriminación: 0.200-0.800 según calidad de distractores
- Parámetros IRT: Coherentes con dificultad estimada

5. **GENERACIÓN PEDAGÓGICA:**

- Afirmación: Resumen de respuesta correcta (máx 15 palabras)
- Evidencia: Competencia evaluada específica (máx 15 palabras)
- Explicación_Respuesta: Usar "Argumento del Autor" o generar explicación
- Pista_1/2/3: Pistas progresivas específicas para esa pregunta
- Error_Común: Error conceptual más probable
- Distractor_X_Concepto: Error específico de cada opción incorrecta

6. **CAMPOS POR DEFECTO:**

- Grado_Escolar: 11°
- Periodo_Aplicación: 2025-1
- Puntos_XP: 10
- Fecha_Creación: 2025-01-20
- Fecha_Actualización: 2025-01-20
- Versión_Item: 1
- Banco_Origen: ICFES_OFICIAL

REGLAS CRÍTICAS:

- Una fila = Una pregunta = Exactamente 80 columnas
- Usar valores categóricos exactos del diccionario ICFES
- Campos no aplicables: "No Aplica"

- Campos específicos de área: Solo llenar los relevantes
- Análisis individual por pregunta, no usar plantillas genéricas

¿Listo para convertir el bloque de preguntas de NotebookLM?

18 Flujo de Trabajo Automatizado

18.1 Proceso Completo NotebookLM → Claude

18.1.1 Paso 1: Preparación en NotebookLM

1. Cargar documentos PDF con preguntas ICFES
2. Aplicar el prompt de extracción específico
3. Obtener bloques de máximo 25 preguntas formateadas
4. Verificar completitud de la extracción

18.1.2 Paso 2: Conversión en Claude

1. Copiar el output de NotebookLM
2. Aplicar el prompt de conversión a 80 columnas
3. Obtener tabla estructurada con análisis psicométrico
4. Verificar formato y completitud de campos

18.1.3 Paso 3: Validación Final

1. Verificar que cada fila tenga exactamente 80 columnas
2. Confirmar valores categóricos según diccionario
3. Revisar coherencia de campos específicos por área
4. Exportar a formato deseado (CSV, Excel, SQL)

18.2 Beneficios del Flujo Automatizado

- **Eficiencia:** Procesamiento masivo de preguntas en minutos
- **Consistencia:** Formato estándar garantizado en todas las preguntas
- **Calidad:** Análisis psicométrico automático por IA
- **Trazabilidad:** Preservación de fuente original y metadatos
- **Escalabilidad:** Capacidad de procesar múltiples documentos
- **Automatización:** Flujo completo sin intervención manual

19 Notas Importantes

- **Consistencia:** Este diccionario debe consultarse antes de crear cualquier registro
- **Validación:** Todos los valores categóricos deben coincidir exactamente con los especificados
- **Obligatoriedad:** Los campos marcados como obligatorios no pueden quedar vacíos
- **Área específica:** Usar solo los campos correspondientes a cada área de evaluación
- **Automatización:** Utilizar el flujo NotebookLM → Claude para procesamiento eficiente
- **Formato plano:** Este formato evita problemas de tablas complejas en LaTeX
- **Actualización:** Mantener sincronizada la documentación con cambios en la estructura

Versión del documento: 2.1

Fecha de creación: 17 de agosto de 2025

Última actualización: 17 de agosto de 2025

Estado: Formato plano optimizado - Prompts implementados

Formato: Libre de problemas de tablas LaTeX