

Algoritmo de evaluación de conflictos de interés (COI)

1. Objetivo del algoritmo

El algoritmo evalúa un artículo científico (PDF o HTML) y produce una estimación estructurada del **riesgo de conflictos de interés (COI)** y de la **credibilidad editorial** del trabajo.

Puntos clave:

- Usa el texto del artículo como única fuente de información.
 - Combina:
 - Extracción semántica mediante un modelo de lenguaje.
 - Un conjunto fijo de **reglas y umbrales**.
 - Siempre devuelve:
 - Una puntuación global `overall_score` (0–100).
 - Un nivel global de riesgo `overall_risk` (low | medium | high).
 - Puntuaciones por dimensión.
 - Un informe explicativo con estructura estable.
-

2. Flujo general del análisis

1. **Ingesta del documento**
 - El usuario sube un PDF o proporciona una URL.
 - El sistema extrae:
 - Texto plano (título, autores, afiliaciones, secciones).
 - Estructura básica del documento (encabezados si existen).
2. **Preprocesamiento**
 - Limpieza de texto (eliminación de artefactos, espacios duplicados, cabeceras de páginas repetidas).
 - Detección aproximada de secciones clásicas:
 - Título, autores, afiliaciones.
 - Abstract, Introducción, Métodos, Resultados, Discusión, Conclusiones.
 - Secciones clave: Conflict of Interest, Competing Interests, Funding, Acknowledgements.

3. Extracción semántica con IA

- El modelo de lenguaje se usa para:
 - Identificar nombres de autores, instituciones y posibles financiadores.
 - Detectar si hay declaraciones de COI explícitas o implícitas.
 - Localizar fragmentos donde se hable de financiación, patrocinadores, empresas, fundaciones, etc.
 - Detectar patrones de lenguaje (promocional vs crítico) y presencia/ausencia de limitaciones.
- Salida de esta fase: un conjunto de **hechos textuales estructurados** que alimentan las reglas del algoritmo.

4. Aplicación de reglas por dimensiones

- A partir de los hechos extraídos, se calculan puntuaciones 0–100 para cada una de las 5 dimensiones definidas.
- Estas puntuaciones no dependen de “intuiciones” del modelo, sino de reglas predefinidas (heurísticas) basadas en buenas prácticas (ICMJE, COPE, WAME, CONSORT, PRISMA, DOAJ).

5. Cálculo de puntuación global

- Se calcula la media de las 5 dimensiones:
 - $\text{overall_score} = \text{media_dimensiones} (0\text{--}100)$.
- Se asigna:
 - $\text{overall_risk} = \text{"low"}$ si 0–33
 - $\text{overall_risk} = \text{"medium"}$ si 34–66
 - $\text{overall_risk} = \text{"high"}$ si 67–100

6. Generación del informe

- El modelo redacta un informe en lenguaje natural, pero:
 - Con estructura fija.
 - Con etiquetas y niveles estables.
 - Con referencias directas a las reglas que han motivado cada puntuación.

3. Las 5 dimensiones del algoritmo

El algoritmo utiliza una **metodología única** para todo tipo de artículos, organizada en 5 dimensiones:

3.1. Disclosure & Funding Transparency

(Transparencia de conflictos declarados y financiación)

Evalúa:

- Presencia o ausencia de:
 - Sección de Conflict of interest / Competing interests / Disclosure.
 - Sección de Funding / Acknowledgements relacionada con financiación.
- Claridad de las declaraciones:
 - Fórmulas como “The authors declare no conflicts of interest”.
 - Descripción concreta de quién financia qué.
- Penalizaciones:
 - Estudios de alto riesgo potencial (ensayos clínicos, fármacos, dispositivos médicos, evaluaciones de productos, etc.) **sin ninguna declaración de COI.**
 - Declaraciones vagas o ambiguas.

Reglas básicas (ejemplos):

- No hay ninguna mención a COI ni financiación en un estudio sensible → score 75–90, level "high".
 - Hay sección de COI con “no conflicts of interest” + financiación claramente descrita → score 20–35, level "low" o "medium" bajo.
 - Hay financiación, pero no se menciona COI en absoluto → score 40–60, level "medium".
-

3.2. Funding–Outcome Alignment

(Relación entre financiación y resultados)

Evalúa:

- Identificación del sponsor:
 - Empresa, fundación, organismo público, universidad, etc.
- Alineamiento entre:
 - Quién financia.
 - Lo favorables que son resultados y conclusiones hacia sus intereses.
 - El grado de discusión crítica (limitaciones, riesgos, alternativas).

Reglas básicas (ejemplos):

- Sponsor comercial claro + resultados muy positivos + lenguaje triunfalista + escasa discusión crítica → score 60–85, level "high".

- Sponsor público/académico + resultados matizados + discusión amplia de limitaciones → score 20–40, level "low" o "medium".
 - No se identifica sponsor y el texto no permite inferirlo → score 40–55, level "medium" por incertidumbre.
-

3.3. Author–Institution–Sponsor Network

(Red autores–instituciones–financiadores)

Evalúa:

- Autores y sus afiliaciones:
 - Universidades, hospitales, centros de investigación, empresas, institutos, etc.
- Coincidencias entre:
 - Instituciones de los autores.
 - Entidades que actúan como sponsors o que se benefician directamente del producto/resultado.
- Concentración de poder:
 - Muchos autores de la misma empresa u organización comercial directamente interesada.

Reglas básicas (ejemplos):

- Varios autores empleados de la empresa que financia el estudio o propietaria del producto analizado → score 70–90, level "high".
 - Afiliaciones académicas o institucionales diversas, sin sponsor comercial directo → score 20–40, level "low"/"medium".
 - Afiliaciones ausentes, incompletas o genéricas ("Research Center" sin más contexto) → score 60–80, level "high".
-

3.4. Journal / Editorial Integrity

(Integridad editorial y riesgo de predatory journal)

Evalúa, a partir de la información disponible:

- Nombre y características de la revista.
- Indicios de:
 - Revisión por pares real.
 - Políticas éticas.
 - Registro de ensayos (clinical trial registration).
- Señales de **predatory journal**:
 - Nombre genérico o sospechosamente rimbombante.
 - Mención a “rapid acceptance”, “fast publication” sin explicación del proceso de revisión.
 - Foco desproporcionado en tasas de publicación (APCs) frente a calidad científica.

Reglas básicas (ejemplos):

- Múltiples señales de revista depredadora → score 70–90, level "high".
 - Indicios de buenas prácticas editoriales (mención a peer review, políticas de ética, etc.) → score 20–40, level "low"/"medium".
 - Información insuficiente sobre la revista → score 40–60, level "medium".
-

3.5. Textual Bias & Reporting Quality

(Sesgos de lenguaje y calidad de reporte)

Evalúa:

- Lenguaje utilizado:
 - Promocional vs sobrio.
 - Uso de adjetivos exagerados (“breakthrough”, “revolutionary”, “game-changer”) sin matices.
- Tratamiento de limitaciones y riesgos:
 - Presencia de sección de limitaciones.
 - Autocrítica sobre sesgos, tamaño muestral, generalización.
- Autocitación:
 - Porcentaje aproximado de citas al propio grupo de autores.
- Grado de transparencia metodológica respecto a estándares esperables:
 - CONSORT (ensayos clínicos).
 - PRISMA (revisiones sistemáticas), etc., de forma aproximada.

Reglas básicas (ejemplos):

- Lenguaje publicitario + ausencia de limitaciones + muchas autocitas → score 60–80, level "high".
 - Lenguaje técnico y sobrio + limitaciones descritas + discusión equilibrada → score 20–40, level "low"/"medium".
-

4. Sistema de puntuación y categorización

Para cada dimensión:

- Se asigna un score de **0 a 100**.
- Se asigna un level categórico:
 - 0–33 → "low"
 - 34–66 → "medium"
 - 67–100 → "high"

La conversión score → level es fija y **no se modifica dinámicamente**.

4.1. Cálculo del riesgo global

- overall_score = media simple de los 5 score de dimensión.
- overall_risk:
 - 0–33 → "low"
 - 34–66 → "medium"
 - 67–100 → "high"

Si falta información en una dimensión, el algoritmo:

- Marca esa dimensión en rango medio (p. ej. 45–55).
 - Explicita en el informe que la falta de datos genera incertidumbre.
-

5. Rol del modelo de IA dentro del algoritmo

La IA no decide “a ojo” los niveles de riesgo. Su función es:

1. **Extraer información** del texto:

- Secciones, fragmentos, menciones a empresas, agencias, universidades.
- Declaraciones de COI y financiación.
- Patrones de lenguaje (promocional vs crítico).
- 2. **Mapear hallazgos a reglas:**
 - A partir de lo que ha detectado, clasifica la situación de cada dimensión en las categorías definidas por las reglas.
 - Aplica los umbrales predefinidos para convertir la situación en un score.
- 3. **Redactar el informe** con estructura fija:
 - Explica, para cada dimensión:
 - Evidencias encontradas (citas o descripciones breves).
 - Regla aplicada (“Sponsor comercial + resultados muy positivos + poca crítica ⇒ riesgo alto”, etc.).

El diseño está orientado a maximizar:

- **Estabilidad y reproducibilidad:** mismo paper → resultados prácticamente idénticos.
- **Transparencia:** el informe explica qué ha disparado cada puntuación.

6. Estructura del informe generado

El informe que ve el usuario sigue siempre la misma plantilla:

1. **Paper analizado**
 - Título (si se puede extraer).
 - Autores principales.
 - Revista / fuente (si se puede inferir).
2. **Resumen global de conflictos de interés**
 - Overall COI risk: low/medium/high
 - Overall COI score: X/100
 - Párrafo breve justificando el nivel.
3. **Desglose por dimensiones (1–5)**

Para cada dimensión:

- Nombre de la dimensión.
- Score: X/100
- Risk level: low/medium/high
- Evidencia principal:
 - Lista de puntos con hallazgos clave.
- Reglas aplicadas:
 - Descripción breve de las reglas que han llevado a esa puntuación.
- 4. **Limitaciones del análisis**
 - Aclara que:

- Solo se analiza el texto disponible.
- No se accede a formularios de COI externos, registros de ensayos, ni bases de datos externas.
- El resultado es una ayuda para la lectura crítica, no un veredicto legal ni definitivo.

5. Recomendaciones para el lector

- Lista de acciones sugeridas:
 - Revisar formularios originales de COI de la revista.
 - Contrastar con otros estudios independientes.
 - Verificar si hay revisiones sistemáticas no financiadas por las mismas entidades.

7. Ejemplo simplificado de salida estructurada (esquema lógico)

A nivel interno, la información se puede representar de forma equivalente a:

```
{
  "overall_score": 65,
  "overall_risk": "medium",
  "dimensions": [
    {
      "name": "Disclosure & Funding Transparency",
      "score": 70,
      "level": "high"
    },
    {
      "name": "Funding–Outcome Alignment",
      "score": 60,
      "level": "medium"
    },
    {
      "name": "Author–Institution–Sponsor Network",
      "score": 55,
      "level": "medium"
    },
    {
      "name": "Journal / Editorial Integrity",
      "score": 45,
      "level": "medium"
    },
    {
      "name": "Textual Bias & Reporting Quality",
      "score": 55,
      "level": "medium"
    }
  ]
}
```

Encima de esta estructura se genera el informe en texto que ve el usuario.

8. Limitaciones actuales del algoritmo

- Se basa únicamente en el contenido textual del paper:

- No accede de forma automática a DOAJ, COPE, ICMJE, ni a bases de datos externas.
- La detección de “predatory journals” es aproximada y basada en señales textuales internas.
- El algoritmo no “demuestra” la existencia de un conflicto de interés; indica **riesgo** de COI en función de:
 - Transparencia.
 - Coherencia entre financiación y resultados.
 - Red de autores e instituciones.
 - Integridad editorial percibida.
 - Calidad y sesgos del reporte.

El propósito es servir como **herramienta de lectura crítica y activismo contra la desinformación científica**, no como tribunal de verdad científica.