

## INNOVACIÓN EN LA GESTIÓN EDITORIAL OPEN JOURNAL SYSTEM: EL PLUGIN *ISSUESPOTLIGHT IA*

*Francisco Máñez*

*Ingeniero informático. Responsable plataforma Revistas UPC (OJS)*

### RESUMEN

El presente artículo describe el desarrollo e implementación técnica de *IssueSpotlight IA*, un plugin innovador para la plataforma de gestión editorial Open Journal Systems (OJS) que integra modelos de lenguaje de gran tamaño (LLM), específicamente Google Gemini en su versión "Tier Free" Gemini 2.5 flash, para la extracción automatizada de conocimiento y análisis inteligente de cada número publicado de la revista. IssueSpotlight IA ofrece cuatro dimensiones de análisis avanzado: 1. Una síntesis editorial temática inteligente. 2. Un radar de innovación que identifica y clasifica tendencias tecnológicas y metodológicas. 3. Una evaluación de impacto alineada con los Objetivos de Desarrollo Sostenible (ODS). 4. Un sistema de geolocalización de filiaciones para el mapeo global de autores.

El trabajo detalla una arquitectura fundamentada en el desarrollo de "plugin OJS" (fundamentado en la "modificación cero del núcleo" y basada en el sistema de hooks de PKP), garantizando la sostenibilidad tecnológica y compatibilidad del sistema. Se analizan los desafíos de la ingeniería de prompts multilingüe y la persistencia de datos bajo demanda. Los resultados demuestran que la IA no solo optimiza los tiempos de gestión editorial, sino que actúa como un catalizador para el descubrimiento científico y la visibilidad institucional, convirtiendo los datos de los diferentes números de cada revista, en una herramienta de análisis de datos y de transferencia de conocimiento en el marco de la Ciencia Abierta. URL para descargar el plugin: <https://github.com/franmanez/issueSpotlight>

### 1. INTRODUCCIÓN

La gestión de revistas científicas ha experimentado una transformación digital radical en las últimas dos décadas, con *Open Journal Systems (OJS)* consolidándose como el estándar de facto para la publicación en acceso abierto. Sin embargo, a pesar de su robustez, los ecosistemas editoriales tradicionales presentan la información de cada número de forma estática y fragmentada. Esto limita la capacidad de los lectores para identificar, de forma inmediata, las tendencias emergentes, el impacto social de las investigaciones o la distribución geográfica de las contribuciones científicas.

Con la eclosión de la Inteligencia Artificial Generativa (GenAI), se abre una oportunidad sin precedentes para redefinir el valor de los metadatos académicos. No obstante, la integración de la IA en OJS ha sido, hasta la fecha, escasa y limitada a tareas administrativas internas. El plugin *IssueSpotlight IA* nace con el objetivo de cerrar esta brecha, proporcionando una capa de análisis inteligente que "cobra vida" cada vez que se publica un nuevo número, transformando títulos y resúmenes en conocimiento estructurado y visual.

## 2. FUNDAMENTACIÓN TEÓRICA Y ESTADO DEL ARTE

### 2.1. La Inteligencia Artificial en el Flujo Editorial: De la Automatización a la Generación

Históricamente, la aplicación de la informática en OJS se ha centrado en sistemas de gestión de flujo de trabajo (workflow management systems) que automatizan la recepción y publicación. Sin embargo, la irrupción de los Modelos de Lenguaje de Gran Tamaño (LLMs) marca un cambio de paradigma: pasamos de sistemas que simplemente "mueven documentos" a sistemas que "entienden y sintetizan conocimiento".

Tras una revisión exhaustiva de la PKP Plugin Gallery (el repositorio oficial de extensiones de OJS), se observa que el estado del arte actual en la integración de IA se limita a:

- *Detección de Plagio*: Herramientas como Similarity Check que utilizan comparación de patrones, aunque recientemente han empezado a integrar capas de detección de texto generado por IA.
- *Gestión de Revisores*: Algoritmos de búsqueda de expertos basados en palabras clave, no en análisis semántico de la trayectoria del investigador.
- *Traducción Estática*: El soporte multilingüe actual sigue dependiendo de la traducción manual de archivos de localización, sin aprovechar el potencial de la traducción automática neuronal en tiempo real para el contenido científico.

IssueSpotlight IA se desmarca de este ecosistema al ser, según nuestro conocimiento, uno de los primeros plugins en integrar de forma nativa la API de un LLM (Google Gemini) para realizar un proceso de "análisis inteligente automatizado", centrándonos en el descubrimiento de conocimiento post-publicación. Esto permite que el sistema no solo gestione el artículo como un objeto digital, si no que extraiga su valor semántico para generar una narrativa coherente de todo el número.

### 2.2. Bibliometría de Tercera Generación y Cienciometría Visual

La bibliometría tradicional se basa en el recuento de citas e índices de impacto (como el factor de impacto de JCR o el índice h). La "segunda generación" introdujo el análisis de redes de citas. No obstante, estamos entrando en lo que algunos autores denominan *Bibliometría Semántica o de Tercera Generación*, donde la unidad de medida no es la cita, sino el concepto y su contexto.

El desafío de la Ciencia Abierta (Open Science) no es solo el acceso gratuito a los PDFs, sino la *"inteligibilidad abierta"*. Con miles de artículos publicados diariamente, el acceso no garantiza el descubrimiento. Las herramientas de visualización cuantitativa, como los mapas de densidad y los radares de tendencias, actúan como filtros cognitivos necesarios para que la comunidad científica pueda navegar la infoxicación académica.

Bajo este marco de bibliometría semántica, las funcionalidades de IssueSpotlight IA se integran como herramientas de análisis de valor añadido que trascienden el recuento de metadatos tradicionales:

- *Síntesis Editorial Inteligente*: Responde a la necesidad de síntesis narrativa, donde la IA actúa como un puente semántico que identifica los hilos conductores del número, facilitando la comprensión transdisciplinar.
- *Radar de Innovación*: Ejemplifica la detección de tendencias por densidad conceptual, permitiendo que términos emergentes sean ponderados por su novedad metodológica y no solo por su frecuencia de aparición.

- *Evaluación de Impacto ODS*: Se alinea con la bibliometría de impacto social (Altmetrics de contexto), utilizando el razonamiento del LLM para mapear la investigación técnica con metas globales de sostenibilidad, una tarea que sería imposible mediante simples algoritmos de búsqueda por palabras clave.
- *Mapeo Geográfico Normalizado*: Aborda el desafío de la limpieza de entidades institucionales, donde la IA aplica un conocimiento enciclopédico para corregir la fragmentación de afiliaciones, convirtiendo datos ruidosos en un mapa de procedencia científica fiable y visual.

### 3. DESCRIPCIÓN FUNCIONAL DEL PLUGIN

*IssueSpotlight IA* se despliega como un panel de control interactivo que actúa como una capa de inteligencia sobre los metadatos de OJS. Su arquitectura se divide en cuatro pilares funcionales, cada uno gobernado por una lógica de prompting especializada:

#### 3.1. Síntesis Editorial Inteligente

A diferencia de un resumen automático convencional, este componente emula el rol de un Editor Jefe. El sistema procesa la totalidad de los título y abstracts del número (hasta 100.000 caracteres de contexto) para generar una narrativa fluida de máximo 250 palabras.

- *Lógica de Agrupación*: La IA identifica hilos conductores transversales, agrupando los artículos por afinidad temática o metodológica en lugar de seguir un orden secuencial
- *Salida Estructurada*: El resultado se entrega en formato HTML semántico, listo para ser revisado por el equipo editorial o publicado directamente como presentación del número.

#### 3.2. Radar de Innovación (Innovation Radar)

Es el núcleo visual del plugin, basado en un gráfico de burbujas (Packed Bubble Chart) que prioriza el valor terminológico sobre la simple frecuencia de palabras.

- *Reglas de Especificidad Estrictas*: Para garantizar un análisis de alto nivel, el sistema exige el uso de *bigramas* y *trigramas* (ej. "Deep Learning" en lugar de "Learning"). Además, cuenta con una "lista negra" de términos genéricos prohibidos (ej. "Investigación", "Análisis", "Sistema") que suelen contaminar las nubes de palabras tradicionales.
- *Categorización de Tendencias*: La IA clasifica dinámicamente cada concepto en tres estadios:
  - Stable: Métodos base o estándares consolidados.
  - Rising: Conceptos con tracción creciente en la disciplina.
  - New: Tecnologías disruptivas o enfoques experimentales detectados.
- *Estandarización Científica*: El radar se genera obligatoriamente en inglés para asegurar la interoperabilidad con tesauros científicos internacionales.

#### 3.3. Impacto ODS (Objetivos de Desarrollo Sostenible)

Este módulo traduce la investigación técnica a un lenguaje de impacto global. La IA evalúa la contribución del número a la Agenda 2030 de la ONU.

- *Ponderación Relativa*: El sistema distribuye un "presupuesto" de relevancia del 100% entre un mínimo de 3 y un máximo de 7 Objetivos de Desarrollo Sostenible.
- *Justificación Cualitativa*: Cada ODS detectado incluye un razonamiento (reasoning) generado

por la IA que explica por qué ese número de la revista es relevante para esa meta específica, permitiendo a la revista generar informes de impacto social instantáneos.

### 3.4. Mapa Global Institucional

El mapa interactivo (basado en Leaflet.js) no solo muestra puntos, sino que realiza una curación de datos socioprofesionales.

- *Normalización de Afiliaciones:* La IA actúa como un experto bibliométrico que resuelve la fragmentación de nombres (ej. unificando "UPC", "Politécnica de BCN" y "Univ. Politècnica de Catalunya" bajo una única entidad normalizada).
- *Enriquecimiento Geográfico:* A partir del nombre de la institución, la IA infiere la ciudad, el país y las coordenadas geográficas exactas, permitiendo visualizar la densidad investigadora y la procedencia global de los autores sin necesidad de bases de datos externas como Scopus o WoS.

## 4. IMPLEMENTACIÓN TÉCNICA Y ARQUITECTURA

### 4.1. Filosofía de "Modificación Cero"

Uno de los mayores desafíos técnicos en el desarrollo de plugins para OJS es garantizar la compatibilidad con futuras actualizaciones del sistema. IssueSpotlight IA se ha implementado siguiendo estrictamente el sistema de Hooks de PKP. La inyección de la interfaz se realiza de forma no invasiva mediante el gancho `TemplateManager::display`, mientras que la lógica de negocio y las peticiones a la IA se gestionan a través de controladores propios (LoadHandler), y la persistencia de datos utiliza el motor ADODB integrado. Este enfoque garantiza que no se haya modificado ni una sola línea del núcleo de OJS, facilitando el mantenimiento a largo plazo.

### 4.2. Integración con la API de Google Gemini

Se ha seleccionado el modelo Gemini Flash vía API (*gemini-2.5-flash-lite*) por su balance óptimo entre latencia precisión en la generación de resultados y generosa ventana de contexto. La comunicación se realiza mediante llamadas cURL cifradas. Técnicamente, cada análisis completo de un número se desglosa en *cuatro llamadas secuenciales al LLM* (Síntesis, Radar, ODS y geolocalización). Esta segmentación permite:

- *Optimización de Prompts:* Cada tarea recibe instrucciones específicas, minimizando el riesgo de "alucinaciones".
- *Integridad Estructural:* Se garantiza que los objetos JSON devueltos sean válidos y no se vean truncados por límites de salida del modelo.

### 4.3. Estrategia Idiomática y Limitaciones de Contexto

Aunque OJS es una plataforma multilingüe, la integración de *IssueSpotlight IA* ha optado por ejecutar el análisis exclusivamente en el idioma principal de la revista (Primary Locale). Esta decisión de diseño responde a tres limitaciones técnicas críticas identificadas durante la fase de desarrollo:

1. *Restricción de la Ventana de Salida (Output Window):* Este es el factor más crítico. A pesar de que los modelos como Gemini Flash tienen una ventana de entrada inmensa, el límite de tokens de salida (generación) es mucho más estrecho. En números con un volumen alto de artículos, intentar generar los resultados en dos o más idiomas superaba con (algunas veces) este límite de generación. Esto provocaba que la respuesta de la IA se cortara a la mitad, invalidando el JSON y dejando el análisis incompleto.

2. *Gestión de Cuota en el Tier Gratuito*: El plan gratuito de Google Gemini 2.5 Flash Lite impone un límite de 20 llamadas diarias por clave API. Al segmentar el análisis en cuatro dimensiones para asegurar la máxima calidad, cada número consume ya un porcentaje importante de la cuota diaria (4 llamadas). Intentar replicar estas llamadas por cada idioma configurado en la revista (español, inglés, catalán, etc.) agotaría la capacidad operativa del plugin de forma inmediata, impidiendo analizar nuevos números en el mismo día.
3. *Consistencia y Estabilidad del Formato*: Forzar a la IA a trabajar con múltiples idiomas en una sola respuesta aumentaba significativamente la probabilidad de errores de code-switching (mezcla accidental de palabras de distintos idiomas) y fallos estructurales en los objetos JSON. Al centrar el esfuerzo en un solo idioma, se garantiza que la IA mantenga una "atención" constante en la precisión semántica y técnica del contenido.

#### 4.4. Modelo de Persistencia y Análisis Precalculado

La arquitectura del plugin rompe con el modelo de "IA en tiempo real para el lector" para adoptar un enfoque de análisis persistente en base de datos. El flujo de trabajo se basa en un proceso de generación por parte del editor y consulta estática por parte del usuario final.

- *Generación en "Batch"*: El análisis lo dispara de forma manual el editor desde el backend administrativo una única vez (o cuando hay cambios significativos). En ese momento, las cuatro dimensiones de análisis (Editorial, Radar, ODS y Geo) se procesan y se consolidan.
- *Almacenar respuesta en BBDD*: Una vez generado, el resultado se almacena en la tabla `issue_ai_analysis` utilizando el `issue_id` como clave única. Cuando un lector visita la revista, el sistema realiza una consulta SQL directa a la base de datos local.

*Tabla 1: Estructura simplificada de la base de datos de IssueSpotlight.*

<b>Campo</b>	<b>Tipo</b>	<b>Función</b>
<b>issue_id</b>	BIGINT	Clave primaria. Vincula el análisis al número de OJS.
<b>editorial_draft</b>	LONGTEXT	Almacena el código HTML de la síntesis temática.
<b>radar_analysis</b>	LONGTEXT	Bloque JSON con los conceptos, frecuencias y tendencias.
<b>ods_analysis</b>	LONGTEXT	JSON estructurado con los ODS y sus razonamientos.
<b>geo_analysis</b>	LONGTEXT	JSON con las coordenadas y nombres normalizados.
<b>date_generated</b>	DATETIME	Registro temporal para auditoría del análisis.

- **Ventajas Operativas:**

- *Coste Cero para el Lector*: El tráfico de usuarios finales no genera llamadas a la API de Google, eliminando costes inesperados o bloqueos por límite de cuota.
- *Velocidad Instantánea*: El lector visualiza las visualizaciones e informes de forma inmediata, sin tiempos de espera derivados del procesamiento de lenguaje natural.
- *Consistencia*: Se garantiza que todos los lectores vean exactamente el mismo análisis validado bibliométricamente.

## 5. INGENIERÍA DE PROMPTS (PROMPT ENGINEERING)

La calidad del análisis depende críticamente de cómo se formulan las instrucciones al LLM. IssueSpotlight IA utiliza un enfoque de prompts de sistema altamente estructurados, diseñados para garantizar la veracidad científica y la integridad del formato JSON.

### 5.1. El Prompt Editorial: Síntesis Narrativa y Role-Playing

Aquí se utiliza la técnica de Role-Playing para que la IA adopte el tono institucional adecuado.

*Prompt Original:*

- Actúa como Editor Jefe. Escribe una editorial corta (max 250 palabras) en HTML (usando <p>, <h3>, <ul>). Agrupa los artículos por temáticas comunes y destaca tendencias. Sé profesional y académico.

REGLA CRÍTICA: NO incluyas ninguna cabecera inicial como 'Editorial', 'Editorial del Editor Jefe' o similares. Empieza directamente con el análisis del contenido.

- Se implementa una constante negativa ("NO incluyas cabeceras") para asegurar que el texto sea inyectable directamente en el tema de OJS sin romper el diseño. El valor añadido aquí no es el resumen, sino la capacidad de agrupación, donde la IA debe encontrar patrones comunes entre artículos aparentemente inconexos.

### 5.2. El Prompt del Radar: Reglas de Especificidad y Ponderación

Este es el prompt más complejo, ya que debe filtrar el "ruido" de palabras comunes para encontrar la esencia tecnológica.

*Prompt Original:*

- Analiza la siguiente lista de Títulos y Resúmenes de artículos académicos de un número de revista.

Extrae los conceptos tecnológicos, metodológicos o teóricos más relevantes.

REGLAS DE ESPECIFICIDAD:

1. BIGRAMAS Y TRIGRAMAS: Evita términos de una sola palabra (ej. 'Diseño', 'Educación'). Prioriza conceptos compuestos de 2 o 3 palabras que definan el nicho (ej. 'Diseño Especulativo', 'Educación Híbrida'). Es decir, que se centre en el Problema-Solución o Método-Contexto.

2. TÉRMINOS PROHIBIDOS: No utilices conceptos genéricos como: Tecnología, Innovación, Análisis, Diseño, Desarrollo, Investigación, Estudio, Ciencia, Sistema, Método, Aplicación, Resultado, Datos.

3. NORMALIZACIÓN: Une sinónimos bajo el término más técnico.

Para cada concepto:

- Cuenta su frecuencia de aparición en los artículos (estimada).

- Determina su tendencia ('trend'): 'new', 'rising', 'stable' según su impacto en el texto.

- El éxito de este prompt reside en la *prohibición explícita*. Al eliminar términos vacíos de contenido técnico (como "Análisis" o "Datos"), obligamos a la IA a buscar en las capas más profundas del texto.

La instrucción sobre *bigramas* y *trigramas* es crucial para evitar visualizaciones genéricas, forzando al modelo a identificar intersecciones de conocimiento (ej. "Sostenibilidad Urbana" frente a "Sostenibilidad").

Sin embargo, el valor diferencial es la capacidad de la IA para realizar un *análisis de madurez tecnológica* mediante la clasificación de tendencias: las tecnologías base o estándares se etiquetan como *Stable*, aquellas con tracción y relevancia creciente como *Rising*, y los enfoques disruptivos o experimentales como *New*. Esta lógica heurística permite que el radar funcione como una herramienta de Horizon Scanning, diferenciando entre el "ruido" terminológico y las verdaderas señales de innovación dentro del número.

### 5.3. El Prompt de los ODS: Razonamiento Cualitativo

Este prompt obliga a la IA a realizar una tarea de "presupuestación" y justificación.

*Prompt Original:*

- Analiza el contenido de los artículos y determina su contribución a los Objetivos de Desarrollo Sostenible (ODS) de la ONU.

Distribuye un total de 100% entre los ODS más relevantes (mínimo 3, máximo 7).

Utiliza estrictamente esta tabla de colores oficiales hex para cada ODS: [Tabla de Colores...]

IMPORTANTE: Devuelve SOLAMENTE un array JSON válido de ODS con este formato exacto:

```
[{"ods": 4, "name": "Educación", "percentage": 30, "color": "#C5192D", "reasoning": "..."}, ...]
```

- Lo más destacable es el anclaje de datos externos (*colores oficiales* en hexadecimal de cada ODS) y la imposición de un *rango estricto de filtrado* (mínimo 3, máximo 7 ODS). Al limitar el número de objetivos, evitamos la "dilución de impacto" (donde la IA podría asignar porcentajes insignificantes a todos los ODS por compromiso semántico) y forzamos al modelo a realizar una jerarquización crítica. Esta restricción garantiza que solo se muestren aquellos objetivos donde la investigación tiene un peso real, proporcionando una visualización mucho más limpia y significativa para el lector. Además, la inclusión del campo *reasoning* añade una capa de transparencia necesaria para validar por qué se ha establecido dicha vinculación.

### 5.4. El Prompt de geolocalización: Curación de Entidades

Este prompt resuelve uno de los problemas históricos de OJS: el "ruido" en los nombres de las instituciones.

*Prompt Original:*

- Actúa como un experto en geología institucional y bibliometría.

Analiza la siguiente lista de afiliaciones de autores.

1. Normaliza las instituciones (ej: 'UPC' -> 'Universitat Politècnica de Catalunya').

2. Para cada institución única, encuentra su Ciudad, País y Coordenadas aproximadas (Latitud y Longitud).

3. Devuelve SOLAMENTE un JSON con este formato exacto:

```
{ "institutions": [{"name": "Nombre Real", "city": "Ciudad", "lat": 0.0, "lng": 0.0, ...}] }
```

- La IA aquí realiza una función de limpieza de base de datos. Al actuar como un "experto en geología institucional", el modelo utiliza su conocimiento interno sobre universidades de todo el mundo para corregir abreviaturas o errores tipográficos de los autores. La inferencia de coordenadas permite que el plugin dibuje el mapa sin depender de servicios de terceros como Google Maps API o Mapbox, que suelen requerir suscripciones de pago.

## 6. VISUALIZACIÓN Y EXPERIENCIA DEL USUARIO (UI/UX)

La interfaz de IssueSpotlight IA ha sido diseñada para transformar metadatos en narrativas visuales, utilizando bibliotecas JavaScript modernas. El objetivo es ofrecer una experiencia inmersiva y analítica.

### 6.1. Panel de Síntesis Editorial y Navegación

La primera capa de interacción presenta la Síntesis Editorial Generativa. A diferencia de los bloques de texto plano estándar, este componente utiliza HTML semántico inyectado dinámicamente para estructurar el contenido en secciones temáticas claras.

- Tecnología: Inyección DOM asíncrona mediante AJAX para no bloquear la carga inicial de la página.
- UX: Diseño limpio y tipografía jerarquizada que facilita la lectura rápida (skimming), permitiendo al usuario captar la esencia del número en segundos.

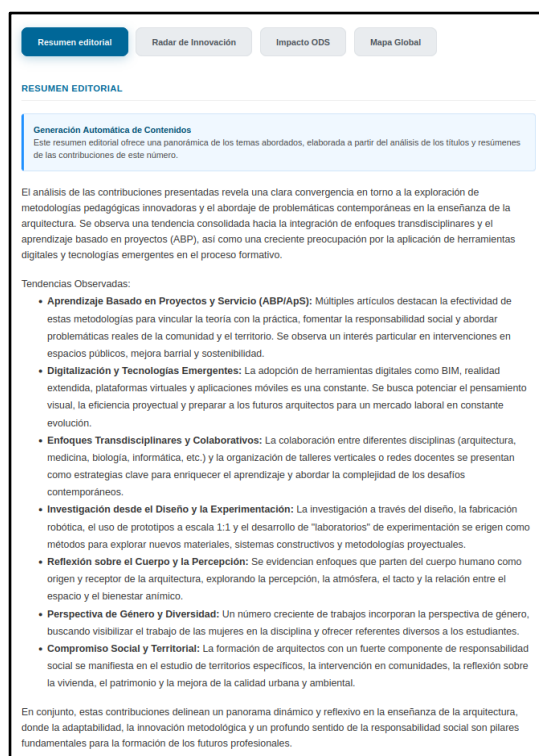


Imagen 1: Vista del bloque de síntesis editorial generado por IA, mostrando la agrupación temática de artículos y la identificación de hilos conductores.

## 6.2. Radar de Innovación (Packed Bubble Chart)

El componente central de descubrimiento visual. Se aleja de las nubes de palabras estáticas para ofrecer una visualización física dinámica donde el tamaño y color de las burbujas tienen significado semántico.

- *Tecnología:* Highcharts.js con el módulo *Packed Bubble*.
- *Interacción:*
  - Clustering: Las burbujas se agrupan gravitacionalmente según su tendencia (Stable, Rising, New).
  - Exploración: Al pasar el ratón, se destaca el concepto y su frecuencia estimada.

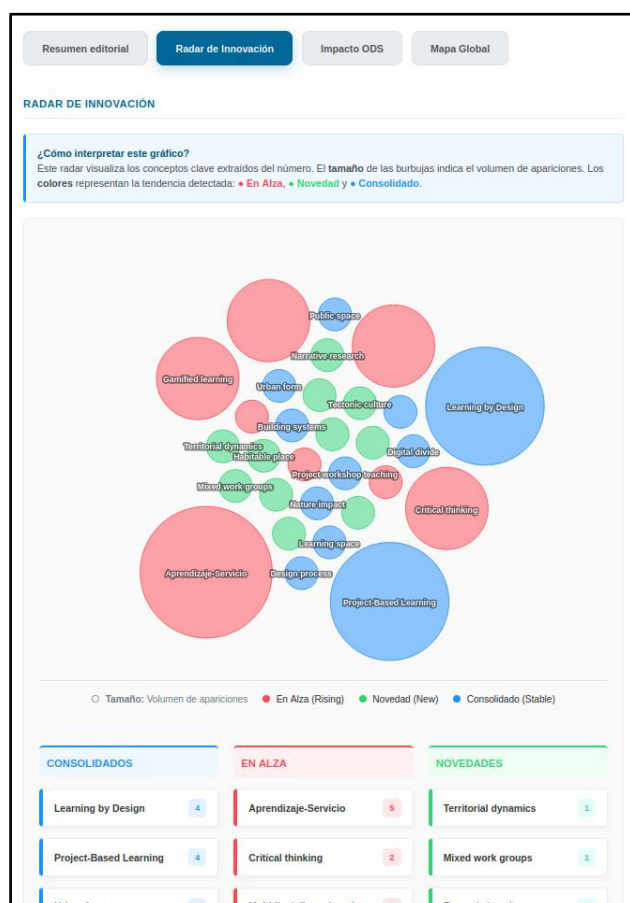


Imagen 2: Visualización del Radar de Innovación mostrando la distribución de conceptos tecnológicos clasificados por tendencias (Nuevos, En Alza, Consolidados).

## 6.3. Tarjetas de Impacto ODS

- Para visualizar la alineación con la Agenda 2030, se ha optado por un diseño de tarjetas (cards) modulares en lugar de listas o tablas.
- Tecnología: CSS Grid y Flexbox para un diseño responsivo que se adapta a móviles y tablets.
- Diseño Visual: Cada tarjeta utiliza el color oficial hexadecimal del ODS correspondiente (inyectado desde el JSON de la IA), garantizando el reconocimiento visual inmediato.
- UX: Incluye un despliegue progresivo de información; el razonamiento cualitativo de la IA se muestra bajo demanda para no saturar la vista principal.

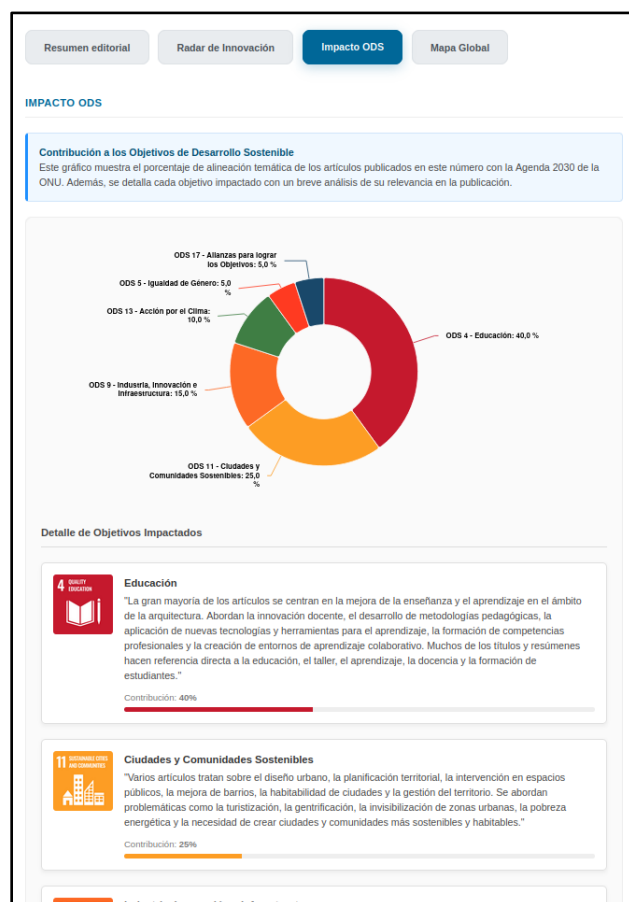


Imagen 3: Panel de impacto ODS, donde cada tarjeta muestra el objetivo, el porcentaje de relevancia y el código de color oficial de la ONU.

## 6.4. Mapa Global Institucional Interactivo

Sistema de Información Geográfica (GIS) para mostrar las filiaciones de los autores del número analizado.

- *Tecnología:* Leaflet.js sobre capas de mapa base ligeras (CartoDB Positron) para mantener la estética limpia.
- *Innovación UX:* Spiral Jittering. Para solucionar el problema de solapamiento, cuando múltiples instituciones pertenecen a la misma ciudad (ej. varias universidades en Barcelona), se aplica un algoritmo de dispersión en espiral. Esto permite que todos los puntos sean clicables individualmente sin ocultarse unos a otros.
- *Popups Enriquecidos:* Al hacer clic, se muestra el nombre normalizado de la institución y el conteo de autores.

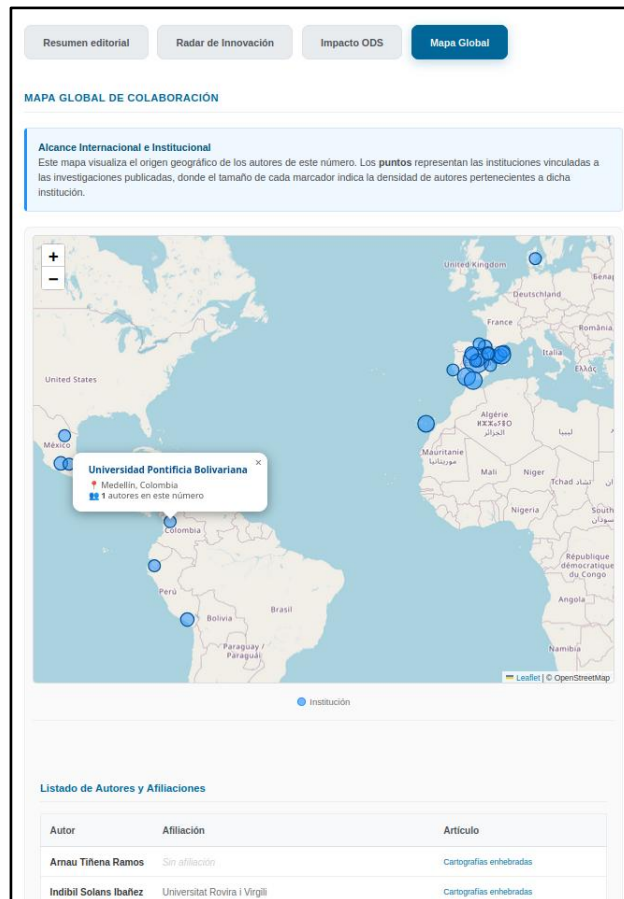


Imagen 3: Mapa interactivo de procedencia institucional. Se observa el agrupamiento (clusters) de de autores y la dispersión precisa de instituciones en zonas de alta densidad.

## 7. INSTALACIÓN Y CONFIGURACIÓN DEL PLUGIN

La implementación de IssueSpotlight IA se ha diseñado para ser accesible a cualquier administrador de OJS, independientemente de su nivel técnico. El proceso se divide en tres fases claras: instalación, obtención de credenciales y ejecución del análisis.

### 7.1. Instalación Estándar en OJS

El plugin se distribuye como un paquete estándar .tar.gz, compatible con la arquitectura de plugins de PKP.

- *Carga del Paquete:* El administrador del sitio (Site Admin) o el gestor de la revista (Journal Manager) accede al menú Ajustes > Sitio Web > Plugins > Cargar Plugin.
- *Despliegue Automático:* Al subir el archivo .tar.gz, OJS descomprime e instala automáticamente los archivos en el directorio plugins/generic/issueSpotlight.
- *Activación:* Una vez instalado, el plugin aparece en la lista de "Plugins Genéricos" y debe ser habilitado mediante el checkbox correspondiente. En este punto, se ejecutan los scripts de instalación (install.xml) que crean la tabla issue\_ai\_analysis en la base de datos si no existe.

## 7.2. Configuración de la API de Google Gemini

Para que el motor de IA funcione, es necesario vincular el plugin con el servicio de Google.

- *Acceso a Ajustes:* En la propia lista de plugins, al desplegar la flecha azul junto a IssueSpotlight IA, se selecciona la opción Ajustes.
- *Obtención de la API Key:* El formulario de configuración incluye un enlace directo a Google AI Studio (<https://aistudio.google.com/>), donde el usuario puede generar una clave API gratuita (Gemini API Key) con su cuenta de Google.
- *Vinculación:* Se introduce la clave en el campo "API Key" del formulario y, al guardar, el plugin valida la conexión. A partir de este momento, el sistema está listo para operar.

## 7.4. Ejecución del Análisis Inteligente (Workflow)

El análisis no es automático por defecto, sino que se activa bajo demanda para dar control total al editor sobre cuándo se "publican" los metadatos de IA.

- *Rol del Editor:* Un usuario con rol de Gestor o Editor accede a la gestión de Números (Issues) (ya sean Futuros o Anteriores).
- *Botón de Acción:* Junto a cada número en la rejilla (grid) de OJS, aparece un nuevo botón denominado "IssueSpotlight IA".
- *Generación y Persistencia:* Al pulsar el botón, se abre una modal que muestra el estado actual ("Pendiente" o "Analizado"). El editor pulsa "INICIAR ANÁLISIS CON IA", lo que desencadena el proceso en segundo plano. Una vez completado (aprox. 10-15 segundos), los resultados se guardan permanentemente en la base de datos y quedan visibles instantáneamente para los lectores en la portada del número.

## 8. DISCUSIÓN: APORTACIONES AL ECOSISTEMA DE CIENCIA ABIERTA

Actualmente, no existen en el repositorio oficial de plugins de PKP soluciones que integren IA generativa para el análisis bibliométrico a nivel de número. IssueSpotlight IA ofrece tres ventajas competitivas claras:

1. *Eficiencia editorial - Reducción del Tiempo de Síntesis:* Un editor puede tardar horas en redactar una editorial; la IA ofrece una base sólida en segundos.
2. *Visibilidad de Impacto Social:* La traducción automática a ODS permite a la revista alinearse con las políticas de ciencia abierta y sostenible de la Unión Europea. Esta funcionalidad es especialmente relevante en el marco del programa Horizonte Europa, que exige la monitorización del impacto social de la investigación financiada con fondos públicos.
3. *Normalización Geográfica:* Elimina el ruido en los datos de afiliación, permitiendo mapas precisos sin necesidad de bases de datos externas de pago (como Scopus o Web of Science).

Sin embargo, es necesario reconocer las limitaciones actuales. La dependencia de modelos externos como Gemini implica una conexión a Internet constante y la aceptación de las políticas de uso de Google. Aunque el plugin utiliza un enfoque de "zero-core-modification", la evolución de las APIs de IA es mucho más rápida que los ciclos de actualización de PKP, lo que requiere un mantenimiento proactivo del código de integración.

## 9. RESILIENCIA OPERACIONAL Y GESTIÓN DE FALLOS DEL LLM

La integración de servicios de IA externos introduce vectores de fallo inexistentes en los sistemas editoriales monolíticos. Durante la fase de desarrollo, se han identificado y mitigado los siguientes riesgos críticos:

### 9.1. Sobrecarga del Proveedor (Model Overload)

El error "**The model is overloaded. Please try again later**" se produce cuando los servidores de Google Gemini experimentan picos de demanda que exceden su capacidad de procesamiento instantáneo. Este fenómeno es especialmente frecuente durante las horas pico de uso global (14:00-18:00 UTC) y refleja una limitación inherente a los servicios de IA basados en la nube compartida.

El plugin no implementa un sistema de reintentos automáticos (retry logic) para evitar agravar la congestión del servicio. En su lugar, se ha optado por una estrategia de transparencia hacia el usuario: cuando se detecta este error, el sistema muestra un mensaje claro en la interfaz indicando que se trata de un problema temporal del proveedor y sugiere volver a ejecutar el análisis manualmente tras unos minutos. Esta decisión de diseño prioriza la honestidad operacional sobre la automatización opaca, permitiendo que los editores comprendan las limitaciones reales de la infraestructura de IA.

### 9.2. Error de Límite de Cuota (Quota Limit Exceeded)

El plan gratuito de Google Gemini impone un límite de *20 llamadas diarias* por clave API. Dado que cada análisis completo de un número consume *4 llamadas al LLM* (una por cada dimensión: Editorial, Radar, ODS y Geo), esto se traduce en una capacidad máxima de *5 análisis completos por día* en el tier gratuito.

Este límite tiene consecuencias prácticas para revistas con alta frecuencia de publicación. Una revista que publique semanalmente podría agotar su cuota en un solo día si regenera análisis múltiples veces. Para mitigar este problema, se han implementado dos medidas:

1. *Documentación Proactiva*: El sistema de configuración del plugin incluye advertencias claras sobre el consumo de cuota, calculando automáticamente el número máximo de análisis diarios disponibles.
2. *Persistencia de Resultados*: Los análisis se almacenan permanentemente en la base de datos, permitiendo su visualización ilimitada sin consumir llamadas adicionales. Solo la regeneración explícita consume cuota.

*Consideraciones de Escalabilidad*: Para revistas institucionales con volúmenes de publicación elevados, se recomienda la migración a planes de pago de Google Gemini o, alternativamente, la exploración de modelos de IA auto-hospedados (*self-hosted*) como Llama 3 o Mistral mediante infraestructuras como Ollama. La arquitectura modular del plugin facilitaría esta implementación sin necesidad de reescribir la lógica de negocio.

### 9.3. Dependencia Tecnológica y Riesgo de Vendor Lock-in

A pesar de las ventajas operativas de utilizar servicios gestionados como Google Gemini, es imperativo reconocer la dependencia estructural que esto genera respecto a las políticas comerciales de un único proveedor (Third-Party Dependency).

La viabilidad a largo plazo de IssueSpotlight IA en su configuración actual está supeditada a dos factores externos fuera del control de la comunidad de OJS:

- *Continuidad de los Modelos Gratuitos*: La disponibilidad de modelos de alto rendimiento (como Gemini Flash) en modalidades free-tier puede verse alterada o cancelada unilateralmente por Google en futuras actualizaciones de su estrategia de negocio.
- *Volatilidad de la API*: Los cambios disruptivos en la sintaxis de las llamadas (endpoints), la depreciación de versiones antiguas de modelos o la modificación de los límites de rate limiting requieren una vigilancia activa y un mantenimiento continuo del código del plugin.

Para mitigar este riesgo, la arquitectura de *IssueSpotlight IA* se ha diseñado de forma modular, desacoplando la lógica de negocio de la capa de conexión con el LLM. Esto facilitaría, en escenarios futuros de obsolescencia o cambio de condiciones, la migración hacia otros proveedores de IA (como OpenAI o Anthropic) o, idealmente, hacia infraestructuras de inferencia locales y soberanas basadas en modelos abiertos (Open Weights).

## 10. DISCUSIÓN ÉTICA, PRIVACIDAD Y SOBERANÍA DEL DATO

La integración de LLMs en el ecosistema científico no está exenta de controversia ética. Durante el desarrollo de IssueSpotlight, se han abordado tres dimensiones críticas:

### 10.1. La Alucinación en el Contexto Académico

Uno de los riesgos inherentes a los modelos generativos es la invención de datos ("alucinaciones"). Para mitigar este riesgo en un entorno donde la veracidad es innegociable, el plugin utiliza un sistema de **anclaje al texto fuente** (*Grounding*). Los prompts están diseñados para prohibir explícitamente a la IA inferir información que no esté presente en los metadatos proporcionados. En el caso del radar, si un concepto no aparece en los títulos o abstracts, la IA no puede inventarlo.

### 10.2. Privacidad y Datos de Afiliación

El envío de datos a servidores externos genera dudas razonables sobre la privacidad. Es importante destacar que IssueSpotlight solo envía a la API de Gemini información que ya es **pública y accesible** en OJS: títulos, resúmenes y nombres de instituciones. En ningún caso se envían correos electrónicos de autores, datos de revisiones por pares (ciegas) o metadatos de artículos no publicados.

### 10.3. Soberanía del Dato vs. Modelos Comerciales

El uso de Google Gemini plantea el debate sobre la soberanía tecnológica de las revistas académicas. Aunque hoy en día los modelos comerciales ofrecen una facilidad de integración superior, la arquitectura del plugin se ha diseñado de forma modular. Esto permitiría, en futuras iteraciones, sustituir la llamada a la API de Google por una conexión a modelos *open source* alojados en servidores locales (como Llama 3 o Mistral) mediante infraestructuras como Ollama o vLLM, garantizando una autonomía total para las instituciones que así lo requieran.

## 11. IMPACTO EN EL SEO ACADÉMICO

Más allá de la experiencia de usuario, IssueSpotlight IA actúa como un motor de enriquecimiento de metadatos que potencia la visibilidad de la revista en buscadores académicos (Google Scholar, Semantic Scholar).

Este impacto se materializa en dos mecanismos directos:

- *Densidad de Palabras Clave*: El radar de innovación extrae automáticamente bigramas y trigramas técnicos de alta especificidad que muchas veces no están presentes en las palabras clave originales de los autores. Al visualizarse como texto en la página del número, estos términos pasan a formar parte del índice rastreable de la revista.

- *Contextualización Automática*: La síntesis editorial generada por IA proporciona una capa adicional de texto estructurado y semánticamente rico que actúa como una "meta-descripción" avanzada del número. Esto favorece que los buscadores entiendan la temática global de la publicación, mejorando su posicionamiento en búsquedas por conceptos transversales y no solo por títulos de artículos individuales.

## 12. CONCLUSIONES

La implementación de IssueSpotlight IA marca un hito en la evolución del ecosistema de Open Journal Systems (OJS), al constituirse como una de las primeras soluciones de integración nativa de Grandes Modelos de Lenguaje (LLM) dentro de la arquitectura base de la plataforma. Este desarrollo valida técnicamente la viabilidad de incorporar capacidades de razonamiento semántico avanzado (Gemini) en entornos PHP tradicionales sin comprometer la estabilidad, seguridad ni mantenibilidad del núcleo, gracias al respeto estricto de los estándares de desarrollo de plugins de PKP.

El valor diferencial de la herramienta radica en su diseño centrado en el editor. Las funcionalidades clave no surgen de la especulación tecnológica, sino del co-diseño activo con equipos editoriales de la Universitat Politècnica de Catalunya (UPC).

Casos de uso como el Mapeo Geográfico Institucional surgen directamente de una necesidad real: automatizar procesos manuales y costosos que los editores de revistas internacionales ya realizaban "a mano" en herramientas externas (como Google My Maps) para visualizar el alcance global de su autoría.

Simultáneamente, la Evaluación de Impacto ODS responde a una demanda estratégica identificada en reuniones con la comunidad editorial: la urgencia de visibilizar la alineación de la investigación con la Agenda 2030, un requisito cada vez más valorado por agencias financiadoras y evaluadores de impacto social.

En definitiva, IssueSpotlight IA no solo optimiza la gestión de metadatos, sino que impulsa la transición de las revistas científicas hacia un modelo de "Plataforma de Conocimiento Interactivo". Al democratizar el acceso a métricas avanzadas de innovación y sostenibilidad sin costes de licencia prohibitivos, el plugin empodera a revistas de cualquier tamaño para competir en visibilidad y transparencia en la era de la Ciencia Abierta.

## 13. REFERENCIAS

- PKP. (2023). *Open Journal Systems 3.3 Reference Guide*. Public Knowledge Project. <https://docs.pkp.sfu.ca/>
- Google. (2024). *Gemini API Documentation: Google AI SDK for PHP*. Google AI Studio. <https://ai.google.dev/docs>
- United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. UN Department of Economic and Social Affairs. <https://sdgs.un.org/2030agenda>
- Highcharts API Reference. (2024). *Packed Bubble Series Options and Physics*. Highcharts.com. <https://api.highcharts.com/highcharts/series.packedbubble>
- PKP. (2022). *Plugin Development Guide for OJS and OMP*. PKP Docs. <https://docs.pkp.sfu.ca/dev/plugin-guide/es/>
- Willinsky, J. (2006). *The Access Principle: The Case for Open Access to Research and Scholarship*. MIT Press.
- PKP. (2021). *OAI-PMH and Metadata Interoperability in PKP Systems*. PKP Support.
- Highsoft. (2024). *Visualizing Complex Data Sets with Highcharts*. Whitepaper.