CSN CONSEJO DE SEGURIDAD NUCLEAR

Proyecto CERO

PROYECTO 0

Inteligencia Artificial para la extracción de
Conocimiento de la **E**xpe**R**iencia **O**perativa

11 de junio de 2026

kampal
Complex Data Intelligent Solutions

Instituto Universitario de Investigación
Biocomputación y Física de Sistemas Complejos
Universidad Zaragoza

POLITÉCNICA
UNIVERSIDAD POLITÉCNICA DE MADRID

El Equipo

El éxito del Proyecto CERO se fundamenta en la **colaboración multidisciplinar** de cuatro entidades clave, aunando visión regulatoria, excelencia académica y capacidad tecnológica

Consejo de Seguridad Nuclear: Responsable de la dirección estratégica y supervisión del proyecto.

BIFI (Unizar) y UPM: Aportan el conocimiento experto en ingeniería nuclear y termohidráulica para la interpretación de sucesos, y la experiencia en física de sistemas complejos.

- Dr. César Queral (Catedrático UPM)
- Dr. Kevin Fernández-Cosials (Profesor UPM)
- Dr. Alfonso Tarancón (Catedrático Unizar)
- Dr. David Yllanes (Investigador Senior BIFI)

Kampal Data Solutions: Como socio tecnológico, ha liderado la ejecución técnica, desarrollando los modelos de Inteligencia Artificial, Procesamiento de Lenguaje Natural (NLP) y la arquitectura software de la plataforma.

- Dr. David Íñiguez (Investigador Principal)
- Dr. Juan Luis Durán (Científico de Datos)
- Alberto García (Ingeniero de Aprendizaje Automático)

El Reto de la Experiencia Operativa: Sobrecarga de Datos

En el ámbito de la seguridad nuclear, la experiencia operativa es una fuente crítica de aprendizaje. Sin embargo, la industria se enfrenta a un desafío creciente: el **volumen** de información acumulada durante décadas es inmenso y heterogéneo.

Gran parte de este conocimiento se encuentra 'encerrado' en **formatos no estructurados**, como informes escaneados (PDFs) y narrativas de texto libre. Esto hace que la revisión manual sea una **tarea intensiva** en recursos y propensa a errores, dificultando la identificación ágil de patrones sutiles o tendencias emergentes que podrían ser precursores de futuros incidentes.

Un Ecosistema de Datos Fragmentado

La complejidad se ve agravada por la dispersión de la información en **múltiples repositorios** internacionales, cada uno con sus propios estándares y formatos.

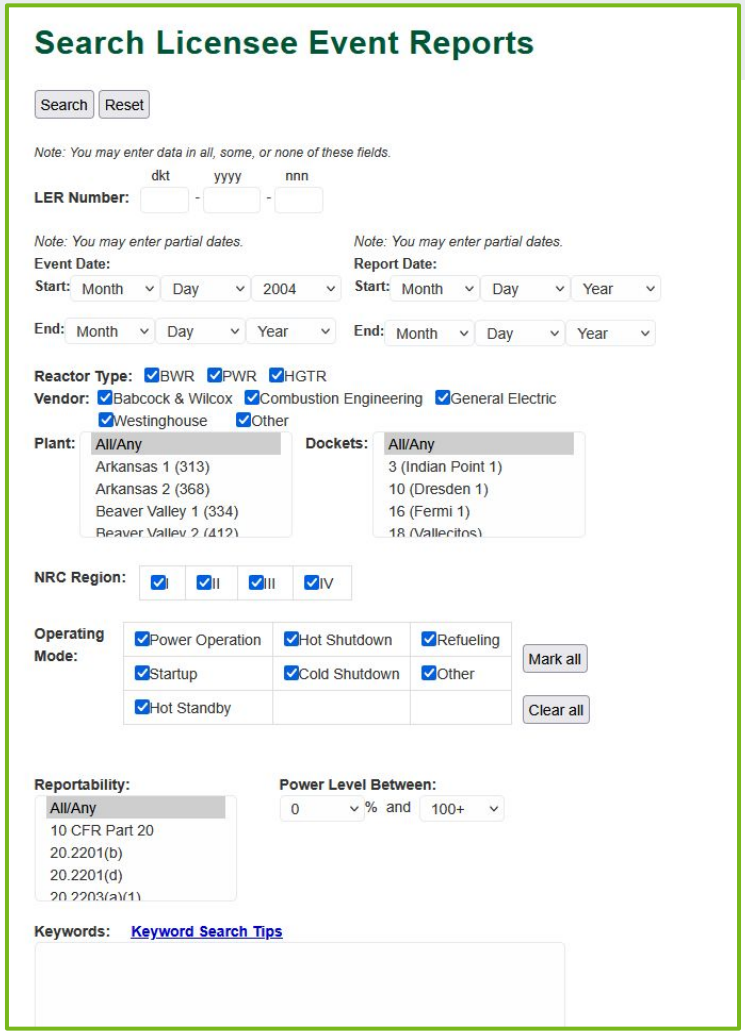
- **ADAMS** (Agencywide Documents Access and Management System)
  - **U.S. NRC**
- **IRS** (Incident Reporting System)
  - **IAEA / OIEA**
- **FIRE** (Fire Incidents Records Exchange)
  - **NEA**
- **CODAP** (Component Operational Experience, Degradation and Ageing Programme)
  - **NEA**
- **NUCIA** (Nuclear Information Archives)
  - **Jpn NRA**
- **FIO** (Ficheros de Incidentes Operativos)
  - **CSN**

La falta de una herramienta unificada que permita interrogar semánticamente estas fuentes crea silos de información que limitan la visión global necesaria para una seguridad proactiva.

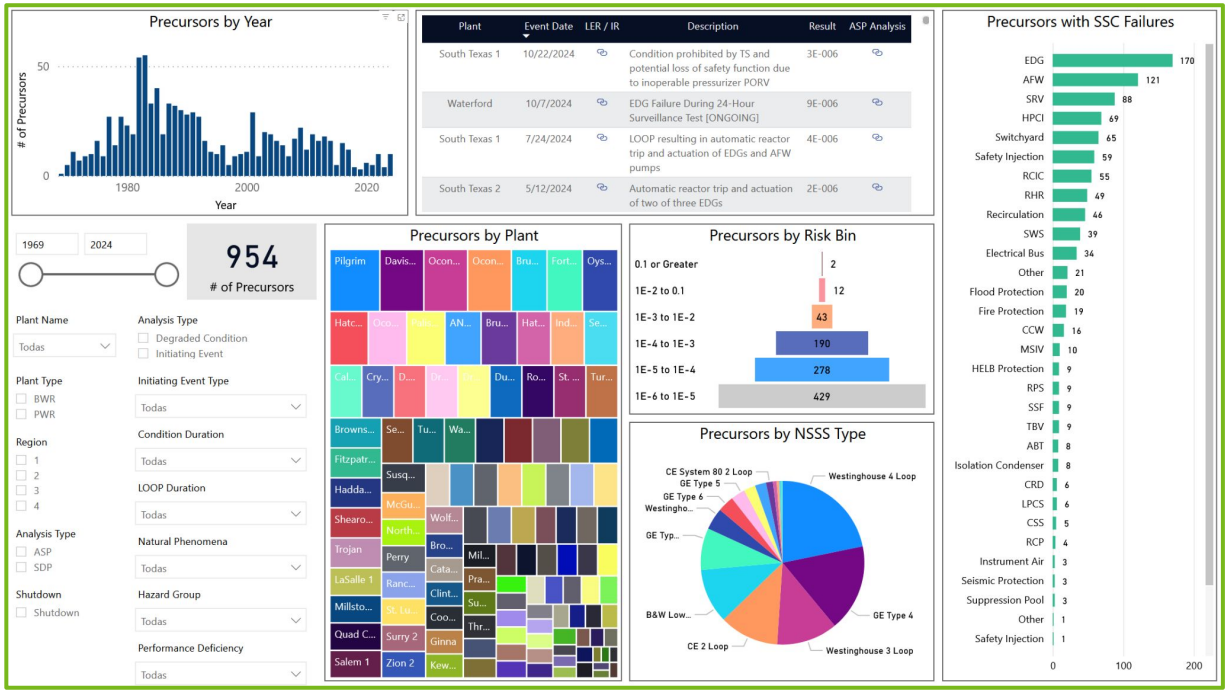
Herramientas Actuales: Buscador de LER

Buscador literal: Depende de la coincidencia exacta de palabras clave. Son herramientas frágiles que no gestionan bien los sinónimos, las variaciones terminológicas o los errores tipográficos derivados del OCR, lo que lleva a omitir información relevante.

Otra debilidad es que **no ordena por relevancia**, sino por fecha o código. Esto obliga a los ingenieros a revisar manualmente listas extensas de informes para encontrar los resultados realmente pertinentes, aumentando el tiempo de análisis y el riesgo de error humano



Herramientas Actuales: ASP Dashboard



Excelente para la **estadística**, pero requieren que los incidentes hayan sido **etiquetados manualmente** en categorías predefinidas.

No permiten descubrir nuevas relaciones no anticipadas ni navegar libremente por el contenido narrativo de los informes

Proyecto CERO: Un Enfoque Dual para la Experiencia Operativa

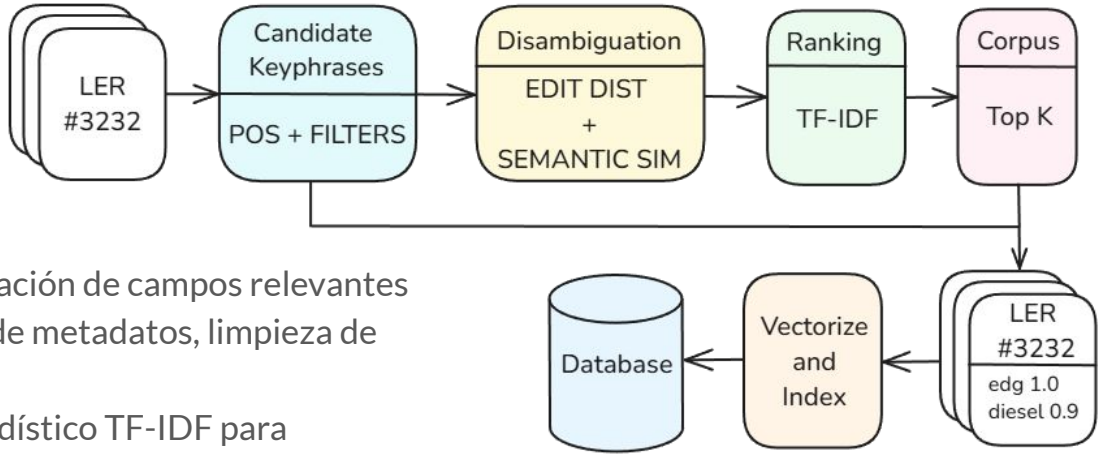
Para superar estas barreras, el Proyecto CERO ha desarrollado una **plataforma basada en Inteligencia Artificial** que no requiere clasificación previa manual. La solución se articula en dos herramientas complementarias:

- **Buscador Inteligente (Reactivo)**: Diseñado para encontrar sucesos similares a partir de una descripción en lenguaje natural. Permite a los operadores comprobar rápidamente si han ocurrido incidentes parecidos al que están analizando, superando la barrera de la literalidad.

- **Herramienta de Clusterización (Preventivo)**: Una visualización basada en redes complejas que agrupa automáticamente los sucesos. Su función es apoyar a los técnicos en la detección de patrones sistémicos y tendencias históricas, facilitando la definición de estrategias de protección a largo plazo. Adicionalmente, este objetivo contempla la construcción y visualización de una red de centrales nucleares interconectadas por su historial de incidencias.

Transformando Documentos en Datos

- **Fuente de Datos:** LER (ADAMS)
- **Ingesta Masiva**: +5.200 informes escaneados (2004-2024)



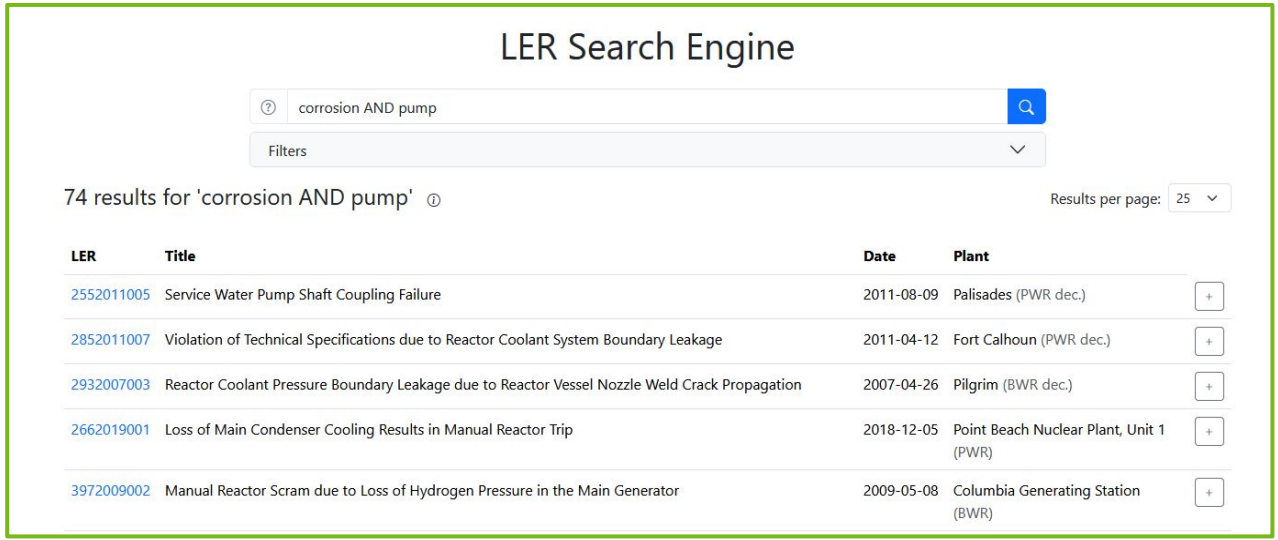
- **Preprocesado**: Lectura PDFs, Identificación de campos relevantes (título, resumen, narrativa,), descarga de metadatos, limpieza de errores del OCR.
- **Extracción de conceptos:** modelo estadístico TF-IDF para identificar los términos técnicos o **palabras clave** (por ejemplo, corrosión o edg) más relevantes, con una puntuación asociada.
- **Vectorización (Embeddings)**: Transformación de textos a vectores numéricos con IA para capturar el **significado semántico** y contextual. Permite al sistema calcular matemáticamente la similitud entre textos.
- **Resultado:** Una base de conocimiento unificada y estandarizada.

Buscador Inteligente de Sucesos

Posee un motor de búsqueda de última generación capaz de interpretar consultas complejas en **lenguaje natural**, interpretando la intención semántica de la consulta y superando las limitaciones de la búsqueda literal exacta.

Nuestro motor clasifica los resultados por **relevancia**, asegurando que los precedentes más críticos para la consulta aparezcan en las primeras posiciones.

Incorpora **filtros híbridos** que permiten cruzar la búsqueda semántica con metadatos estructurados

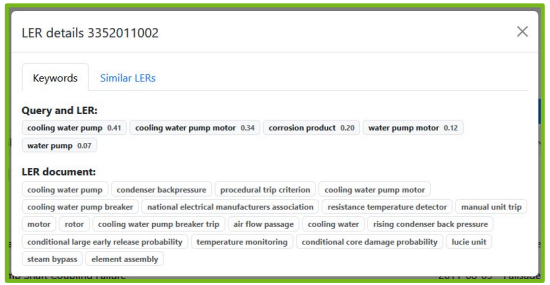


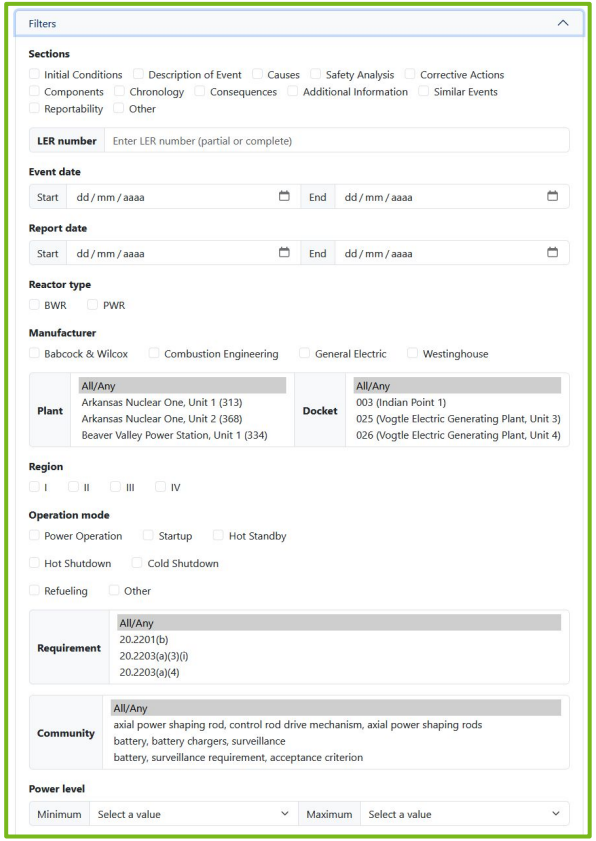
Buscador Inteligente de Sucesos
Filtros y Explicabilidad

La plataforma complementa la potencia de la IA con herramientas de **precisión** y **transparencia** para el analista.

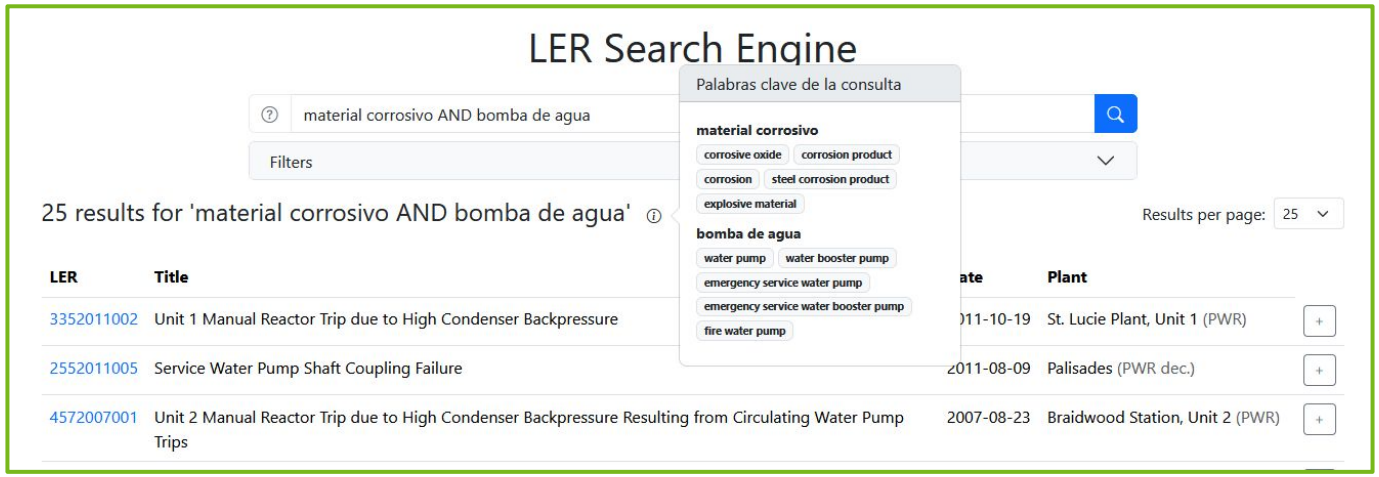
El sistema de **filtros híbridos** permite acotar la búsqueda semántica cruzándola con metadatos estructurados rigurosos.

Por otro lado, hemos integrado mecanismos de **explicabilidad**: el usuario puede visualizar instantáneamente qué términos exactos han provocado la coincidencia.





Buscador Inteligente de Sucesos

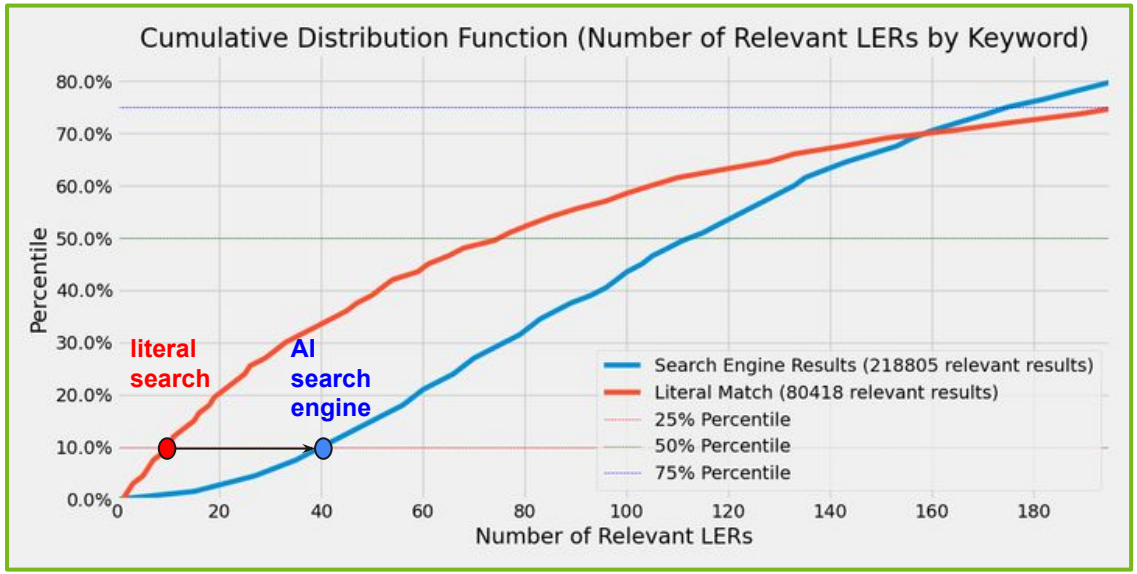


Resultados: Buscador Inteligente

El Buscador Inteligente recupera más de 40 resultados relevantes para el 90% de las consultas seleccionadas, lo que representa una **mejora de 4 veces** respecto a la búsqueda literal.

Un avance fundamental respecto a las herramientas actuales reside en el orden de los resultados. Nuestro motor los clasifica por **relevancia**, garantizando que los precedentes más críticos y similares aparezcan siempre en las primeras posiciones.

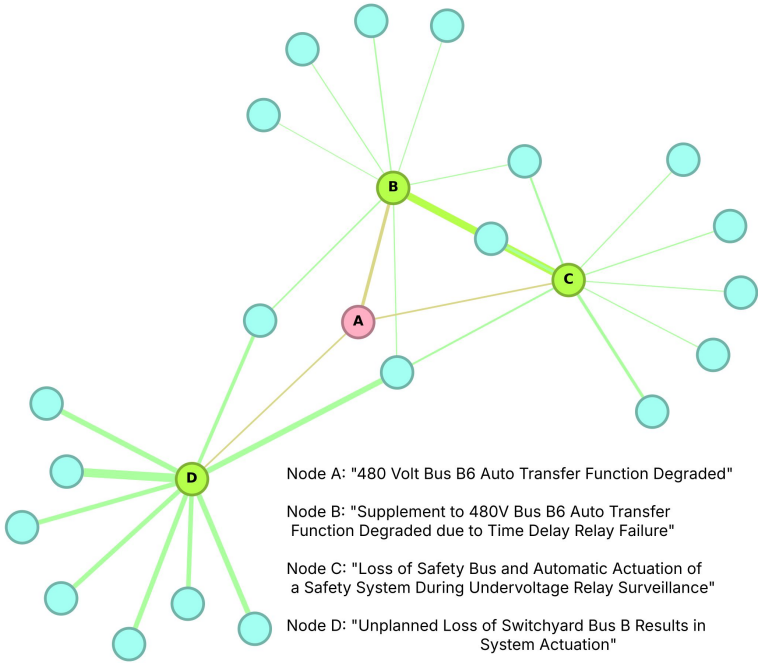
De esta manera, se asegura que el operador localice más incidentes críticos de manera más **eficiente**.



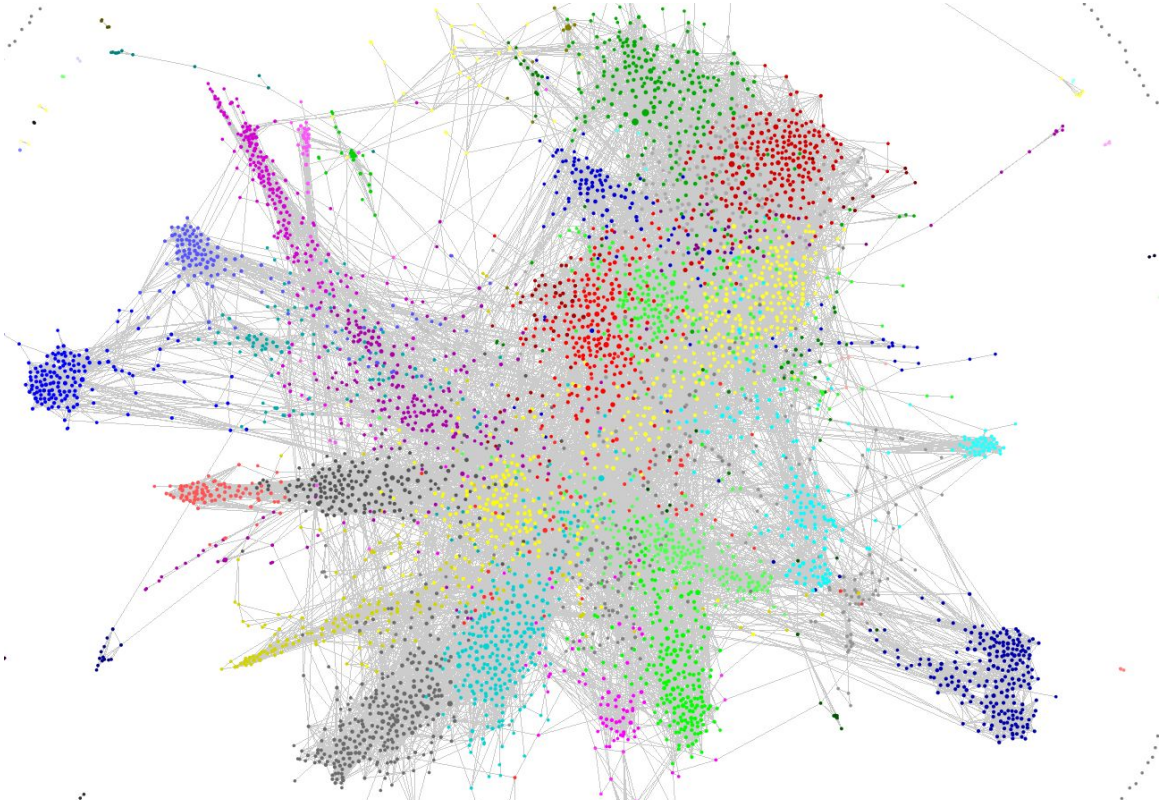
Visualización Global: Incidentes Similares

Para proporcionar una **visión global** de la experiencia operativa, hemos generado una Red de Incidentes donde cada nodo representa un informe y los enlaces indican una fuerte **similitud semántica**, es decir, textos similares (Causas, Consecuencias, Acciones Correctivas...).

Esta visualización transforma una base de datos plana en un **mapa de conocimiento** interconectado.



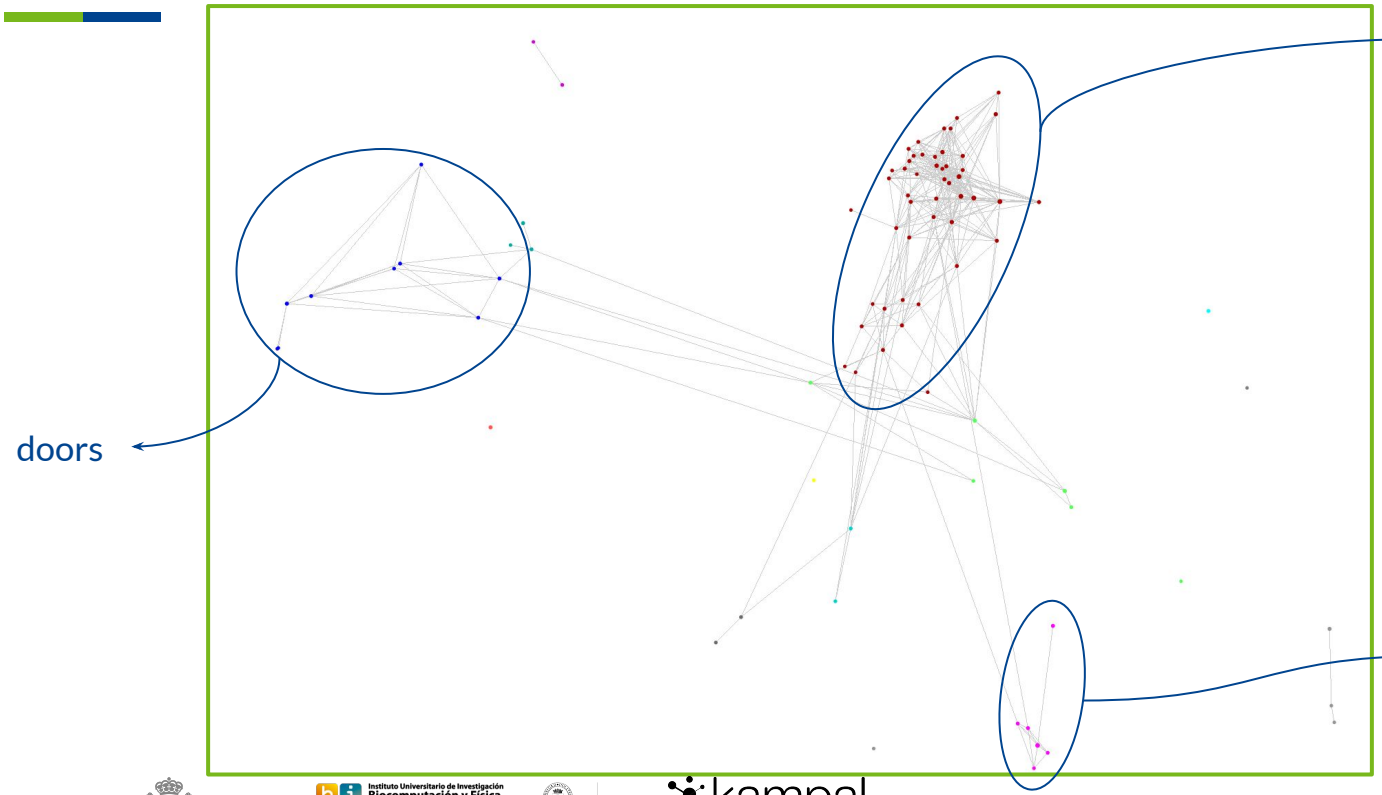
Visualización Global: La Red de Incidentes



Al aplicar algoritmos de aprendizaje no supervisado (detección de comunidades), el sistema ha identificado automáticamente **clústeres temáticos** bien diferenciados, agrupando sin intervención humana incidentes relacionados con problemas específicos como fallos en generadores diésel, fugas de refrigerante o degradación de válvulas.

Gracias a esta visualización interactiva, los analistas pueden navegar por las conexiones entre sucesos, facilitando la identificación de **patrones sistémicos** y problemas recurrentes que quedarían ocultos en una revisión documental tradicional.

Red de Incidentes. Filtrado por Central

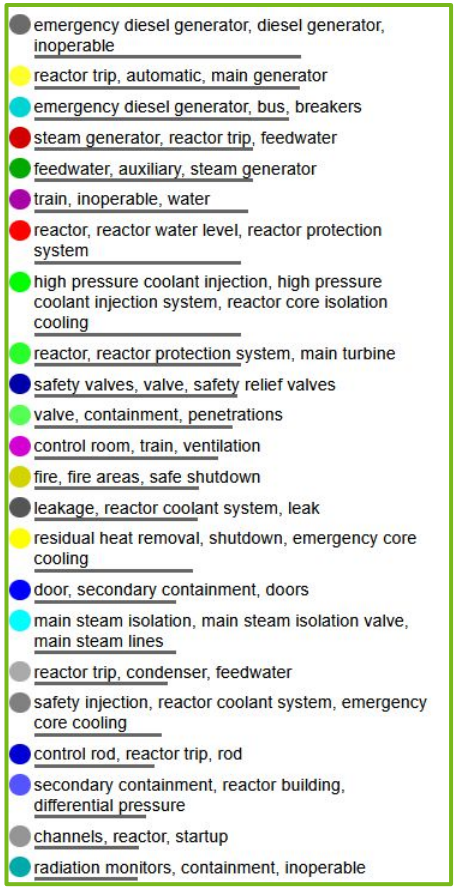


Resultados: Red de Incidentes

El análisis de la red confirma su potencial como herramienta de descubrimiento:

- **Detección Automática:** El sistema ha identificado **75 clústeres temáticos** bien definidos, agrupando sin intervención humana incidentes relacionados con problemas específicos (ej. fugas, generadores).

- **Capacidad de Recuperación:** Las pruebas demuestran que, partiendo de un incidente y explorando solo hasta sus vecinos de segundo orden (vecinos de los vecinos), es posible recuperar el **70% de los incidentes relevantes** conocidos.



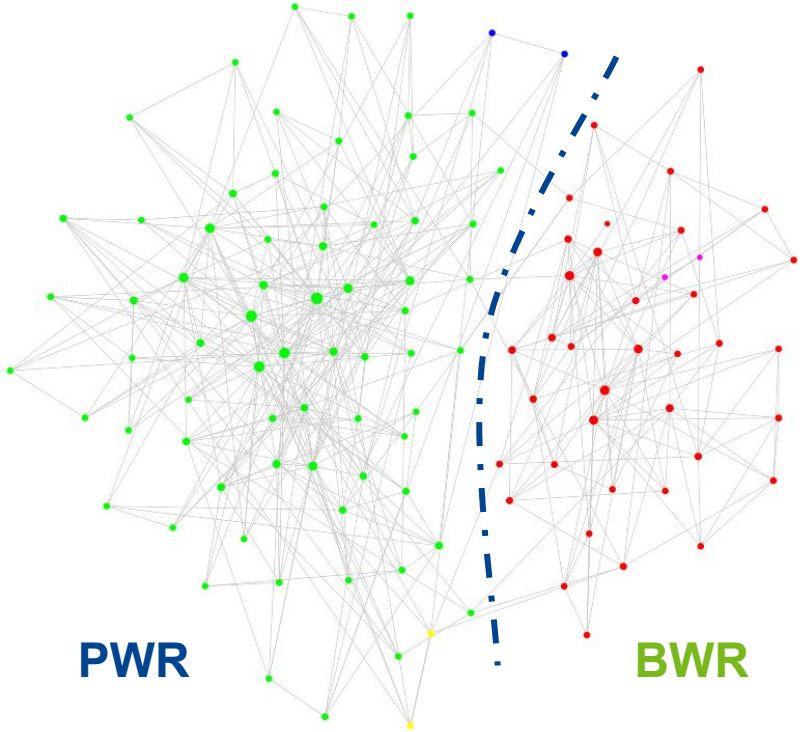
Visualización Global: La Red de Centrales

Elevando el nivel de análisis, hemos construido una **Red de Centrales Nucleares** agregando las similitudes de sus incidentes históricos. En este grafo, dos centrales están conectadas si comparten un historial de problemas operativos similares.

El resultado valida la capacidad del modelo para capturar la realidad física de las instalaciones: el algoritmo agrupa automáticamente las 112 centrales en dos grandes comunidades que corresponden a una **Separación Tecnológica Automática**, con una **precisión del 97%**:

Por un lado, las tecnologías PWR (Agua a Presión) y, por otro, las BWR (Agua en Ebullición).

Esto demuestra que la IA ha 'entendido' las diferencias de diseño intrínsecas a través de la lectura de sus incidentes.



Impacto, Difusión y Transferencia

La calidad innovadora de la metodología ha sido validada y avalada por la comunidad científica internacional en distintas conferencias. Además, aseguramos una transferencia tecnológica efectiva al CSN con una aplicación disponible durante un año.

Presencia Internacional (Congresos y Proceedings)

- ANS Annual Conference 2025 (American Nuclear Society)
- NENE 2025 (Nuclear Energy for New Europe)

Impacto Nacional y Divulgación

- Revista Alfa, N.º 59 (CSN)
- Aceptado en RASNE (Reunión Anual de la Sociedad Nuclear Española)

Producción Científica en Curso

- Paper para revista indexada de alto impacto

Aplicación para usuarios del CSN durante 1 año

Conclusiones

En conclusión, el Proyecto CERO ha transformado un archivo histórico en una capacidad predictiva y preventiva.

Mayor Eficiencia

multiplicando por cuatro la recuperación de información relevante

Visión Global

revelando patrones sistémicos ocultos

Resultados Validados

por métricas, revisión interna de expertos y aval en conferencias

Modelo de Referencia

metodología probada y escalable, lista para ser aplicada a otras bases de datos

Agradecimientos

Agradecemos el soporte técnico y financiero proporcionado en el marco del proyecto 'Inteligencia Artificial para la extracción de Conocimiento de Experiencia Operativa (Proyecto CERO)' (Ref: SUBV-108-2023), financiado por el Consejo de Seguridad Nuclear.

También agradecemos a la U.S. NRC por la disponibilidad de los datos públicos.

Muchas Gracias

