DESAFÍO UNIVERSIDAD - EMPRESA

Esta necesidad tecnológica forma parte del Concurso de Proyectos de I+D+i y/o consultoría en colaboración Universidad – Empresa "Desafío Universidad Empresa" 2026 organizado por la Fundación Universidades y Enseñanzas Superiores de Castilla y León.

TÍTULO DE LA DEMANDA TECNOLÓGICA A RESOLVER

Referencia:

NT52

Título de la demanda tecnológica propuesta

Detección automática y explicable de anomalías en los registros de auditoría (audit trail) de sistemas informatizados en entornos regulados

Acrónimo:

AUDIT-IA.

Áreas de interés de la demanda tecnológica

(Principal) Industria inteligente y fabricación avanzada
Salud y Atención social, Tecnologías y ciberseguridad

Resumen:

Los sistemas informatizados de la industria farmacéutica y otros sectores regulados generan registros de auditoría (audit trail) que la normativa obliga a revisar para garantizar la integridad de los datos. Hoy esa revisión es manual, por muestreo y dependiente del revisor, lo que la hace costosa y poco fiable. Se busca desarrollar un método que normalice registros heterogéneos de distintos sistemas, detecte automáticamente patrones anómalos con relevancia regulatoria, explique cada alerta al revisor y pueda justificarse ante una inspección. Se valorará la generación de datos sintéticos realistas para entrenar y validar los modelos sin usar información confidencial.

PALABRAS CLAVE: integridad de datos, audit trail, detección de anomalías, inteligencia artificial explicable, entornos regulados GxP.

DESCRIPCIÓN DE LA NECESIDAD DEMANDADA

1.- Descripción de la demanda tecnológica.

La normativa aplicable a la fabricación y el control de medicamentos (EU GMP Anexo 11, 21 CFR Part 11 y las guías de integridad de datos de las autoridades sanitarias) exige que los datos generados por sistemas informatizados cumplan los principios ALCOA+ y que sus registros de auditoría se revisen de forma periódica y documentada. Estos registros recogen quién creó, modificó o eliminó cada dato, cuándo y por qué.

En la práctica, un centro de fabricación o un laboratorio de control de calidad opera decenas de sistemas (equipos de laboratorio, sistemas de fabricación y acondicionamiento, gestión de calidad, ERP), cada uno con un formato de registro distinto. La revisión se hace a mano, por muestreo y con criterios que dependen de la experiencia del revisor. Es una tarea costosa, difícil de escalar y con riesgo real de que las anomalías relevantes pasen desapercibidas hasta una auditoría o inspección.

Se demanda un proyecto de investigación que desarrolle y valide un método para:

- NORMALIZAR registros de auditoría heterogéneos procedentes de distintos sistemas y fabricantes en un modelo de datos común.
- DETECTAR automáticamente patrones anómalos con relevancia regulatoria: eliminaciones o modificaciones de datos, cambios fuera de horario o fuera de secuencia, modificaciones repetidas sobre un mismo registro, reprocesados no justificados, accesos con usuarios compartidos o atípicos, entre otros.
- EXPLICAR cada alerta de forma comprensible para el revisor, indicando qué patrón la origina y por qué es relevante, de modo que la decisión final siga siendo humana y quede documentada.
- GENERAR datos sintéticos realistas, con anomalías conocidas introducidas de forma controlada, que permitan entrenar y validar los modelos sin depender de datos confidenciales.
- DEFINIR criterios de validación del propio método, coherentes con un enfoque basado en riesgo (GAMP 5), para que su uso pueda justificarse ante una inspección.

Resultado esperado: un prototipo funcional validado con datos sintéticos y, si es posible, con registros reales anonimizados, junto con métricas de rendimiento (detección, falsos positivos, tiempo de revisión ahorrado) y una propuesta de criterios para su uso en entornos regulados.

El enfoque debe ser trasladable a otros sectores con exigencias de trazabilidad e integridad del dato, como la industria alimentaria, los productos sanitarios o los laboratorios de ensayo.

2.- Antecedentes.

La integridad de los datos es desde hace años una de las principales causas de observaciones en inspecciones de la industria farmacéutica. Las autoridades sanitarias (EMA, FDA, MHRA, PIC/S) han publicado guías específicas que exigen revisar los registros de auditoría de los datos críticos, con una frecuencia y un alcance proporcionales al riesgo.

Algunos fabricantes de software incorporan funciones de revisión dentro de sus propios sistemas, pero se limitan a ese sistema concreto y a filtros predefinidos. No existe, hasta donde conoce la entidad demandante, una solución que permita analizar de forma conjunta registros de sistemas distintos, detectar patrones anómalos más allá de reglas fijas y justificar sus resultados ante una inspección.

La entidad demandante cuenta con más de diez años de experiencia en integridad de datos, validación de sistemas informatizados y revisión de registros de auditoría en plantas de fabricación farmacéutica, incluida la definición de los criterios de revisión de sistemas GxP. Aportará al proyecto el conocimiento regulatorio y del proceso, la definición de los patrones con relevancia real, la validación funcional de los resultados y su orientación a la aplicación práctica en empresas.

La inteligencia artificial y, en particular, los modelos de detección de anomalías en secuencias de eventos y los modelos de lenguaje para interpretar formatos heterogéneos, abren hoy una vía que no era viable hace pocos años.

3.- Posibles enfoques del proyecto de investigación.

Sin carácter limitativo, se consideran de interés los siguientes enfoques, que pueden combinarse:

- Modelado de un esquema común de eventos de auditoría y técnicas automáticas, incluidos modelos de lenguaje, para extraer y mapear registros de formatos distintos a ese esquema.
- Detección de anomalías no supervisada o semisupervisada sobre secuencias de eventos por usuario, dato o sistema, combinada con reglas derivadas de los requisitos regulatorios (enfoque híbrido).
- Técnicas de minería de procesos para comparar la secuencia real de acciones con la secuencia esperada de un proceso.
- Métodos de explicabilidad que traduzcan cada alerta a un lenguaje comprensible para un revisor de calidad.
- Simulación de procesos de laboratorio o fabricación para generar registros sintéticos realistas con anomalías etiquetadas.
- Diseño de un marco de validación del método basado en riesgo, con métricas reproducibles.

Se valorarán especialmente las propuestas interdisciplinares que combinen informática, estadística o ciencia de datos con conocimiento farmacéutico o de calidad, y la participación de más de una universidad.

4.- Enfoques sin interés.

- La implantación o adaptación de un producto comercial existente.
- Soluciones limitadas a un único sistema o fabricante sin posibilidad de extenderse a otros.
- Modelos de caja negra que no expliquen el motivo de cada alerta.
- Herramientas genéricas de ciberseguridad orientadas a detectar intrusiones, sin foco en la integridad de los datos de proceso.
- Sistemas que sustituyan la decisión del revisor en lugar de apoyarla.

- Estudios exclusivamente teóricos o documentales sin prototipo validado.
- Enfoques que dependan desde el inicio de acceder a datos reales confidenciales de empresas.

Si desea remitir una propuesta de solución tecnológica *(proyecto de investigación y/o consultoría)* ***deberá enviar el formulario de participación*** *(ANEXO II), descargable en* ***www.redtcue.es/desafio*** ***a una de las direcciones de correo electrónico que se indican en las bases del concurso antes del 18/12/2026***

Acceso a información general del concurso

FUESCYL
FUNDACIÓN UNIVERSIDADES Y ENSEÑANZAS SUPERIORES DE CASTILLA Y LEÓN

Junta de Castilla y León

t-cue TRANSFERENCIA DE CONOCIMIENTO Universidad Empresa

RIS3 CASTILLAYLEÓN 2014-2020

desafío universidad-empresa