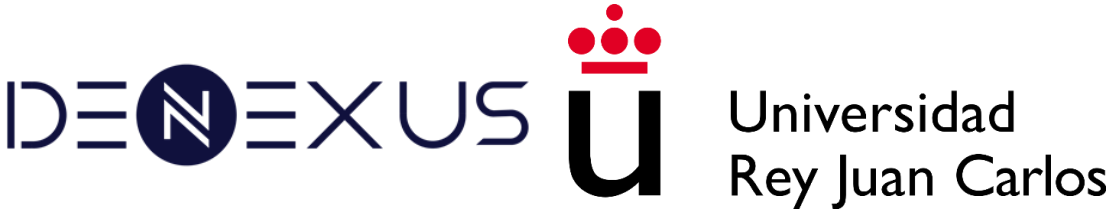
Convocatoria Colaboración Público-Privada CPP2021-009025

# DICYME:

## Dynamic Industrial CYberrisk Modelling based on Evidence

Iniciativa Conjunta de:




## ENTREGABLE 4.4:

## Versión final del sistema completo (integrado con DeRISK) e informe final de resultados de prueba.

*Coordinadores:*

Romy R. Ravines (DeNexus Tech)
Isaac Martín de Diego (Universidad Rey Juan Carlos)
Alberto Fernández Isabel (Universidad Rey Juan Carlos)


## Contenido

# 1 Introducción

Este entregable presenta la **versión final del sistema DICYME**, una solución integral para la **cuantificación dinámica del ciber riesgo industrial**, que consolida y amplía las capacidades desarrolladas en el entregable anterior (E4.3). El objetivo es ofrecer una herramienta robusta y escalable que permita a **operadores industriales, aseguradoras y otros actores clave** gestionar el ciberriesgo de forma más precisa, transparente y basada en evidencia.

En esta versión se han incorporado mejoras sustanciales que convierten a DICYME en un sistema completo, preparado para su integración en entornos reales y alineado con los objetivos del proyecto:

- **Implementación final del sistema**: incluye todos los módulos funcionales — recolección automática de evidencias, generación de indicadores, cuantificación del riesgo (CRQ) y visualización avanzada— integrados en una plataforma web interactiva.
- **Visualización y soporte a la decisión**: se han desarrollado interfaces gráficas intuitivas que permiten explorar datos, analizar indicadores y comprender el impacto financiero y operativo del ciberriesgo, facilitando la toma de decisiones estratégicas.
- **Explicabilidad mediante IA**: el sistema incorpora mecanismos de interpretabilidad que explican los resultados y recomendaciones, aumentando la transparencia y la confianza en los modelos.
- **Generación de informes personalizados**: se habilita la creación automática de reportes adaptados a diferentes perfiles de usuario (industrial, asegurador), con métricas clave y recomendaciones priorizadas.
- **Integración con DeRISK**: DICYME se conecta con el sistema de modelos de CRQ industrial de DeNexus, DeRISK, aportando indicadores avanzados (Attractiveness, THRACT, CVE2TTs) y capacidades dinámicas que enriquecen la arquitectura probabilística de DeRISK.
- **Validación funcional y escalabilidad**: se han realizado pruebas con datos reales en entornos controlados, confirmando la coherencia de los cálculos y la capacidad para operar en escenarios industriales.

En conjunto, este entregable marca la culminación técnica del proyecto, ofreciendo una solución que cuantifica el ciberriesgo con mejor información accionable para la gestión proactiva en sectores críticos.

# 2 Flujo de trabajo de DICYME

DICYME sigue un flujo de trabajo de varias fases para producir el riesgo cibernético industrial final y el apoyo a la toma de decisiones:

1. Recolectar automáticamente evidencias de ciber riesgo,

2. Crear datos para indicadores en repositorios,
3. Transformar datos en indicadores,
4. Cuantificar el riesgo cibernético (CRQ, por sus siglas en inglés),
5. Visualizar y recomendar para el apoyo a la toma de decisiones.

Los indicadores de DICYME pueden introducirse en un sistema CRQ industrial, como el sistema DeRISK. A su vez, el sistema de CRQ industrial se puede complementar con el sistema de recomendación de DICYME (ver Ilustración 1).

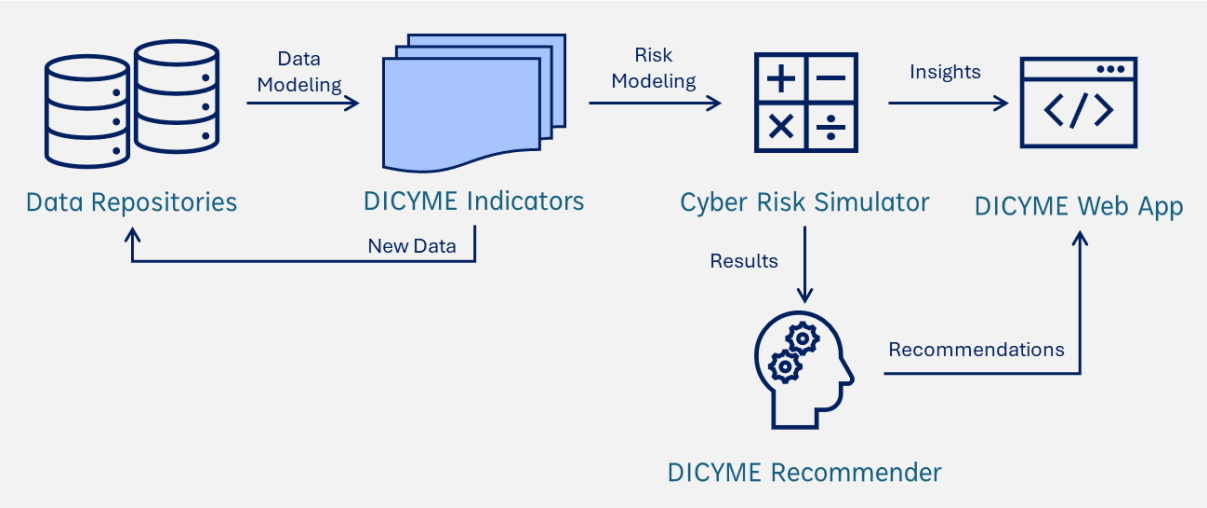



*Ilustración 1: Sistema DICYME completo.*

## 2.1 Recolección automática de evidencias

El primer paso del flujo de trabajo de DICYME se centra en la extracción automatizada de evidencias de riesgo cibernético desde fuentes internas y externas.

- **Fuentes internas:** Plataformas de protección de sistemas ciberfísicos (CPS) y soluciones de visibilidad de red. Ejemplo: IDS de una instalación industrial.
- **Fuentes externas:** Herramientas de inteligencia de código abierto (OSINT) y datos de inteligencia de amenazas. Ejemplos son información sobre ciber incidentes publicados o información sobre ciber actores.

## 2.2 Repositorios de Datos

La evidencia del paso anterior se procesa para obtener datos curados asegurando un nivel de calidad del dato exigido para mantener la calidad de los indicadores de DICYME que se calcularán después. Los conjuntos de datos son:

- Incidentes cibernéticos.
- Perfil de la víctima.
- Sistemas de detección de intrusiones.
- Vulnerabilidades y técnicas asociadas.

## 2.3 VISCA

**VISCA** (Victim Identification and Structured Company Attributes) es un componente clave dentro del sistema DICYME, diseñado para **automatizar la recolección y enriquecimiento de evidencias** relacionadas con víctimas de ciber incidentes. Su objetivo principal es transformar descripciones textuales de incidentes en perfiles firmográficos completos y fiables, que sirvan como base para la cuantificación del ciberriesgo y la correlación con indicadores dinámicos.

VISCA participa en el flujo de trabajo del sistema DICYME de la siguiente manera. El flujo de trabajo multi agente de VISCA colabora en la recolección de evidencias y creación de datos de DICYME mediante los siguientes pasos:

- **Automatización** de la identificación de víctimas a partir de descripciones de incidentes.
- **Enriquecimiento** de los perfiles de entidades con atributos como ingresos, tamaño, código NAICS y estado de cotización pública.
- **Provisión de niveles de confianza** en los datos recopilados.

## 2.4 DICYME indicators

Los datos de los pasos anteriores, basados en evidencias, son usados para calcular los indicadores de ciber riesgo de DICYME:

- Atractivo
- THRACT
- CVE2TTs

## 2.5 Cyber Risk Simulator

Los indicadores se usan en los modelos de cuantificación del ciberriesgo industrial (CQR industrial, por sus siglas en inglés) para estimar:

- **Frecuencia de eventos de pérdida:** Se consideran factores como las capacidades de los actores de amenaza, la visibilidad de la infraestructura y los perfiles firmográficos.
- **Magnitud de la pérdida:** Se modelan los impactos técnicos y financieros utilizando nuevamente datos firmográficos, tendencias históricas e inteligencia de amenazas.

El sistema CRQ industrial de DICYME es una estructura jerárquica basada en la metodología **FAIR (Factor Analysis of Information Risk)** que es un marco cuantitativo estándar para comprender, analizar y cuantificar el ciberriesgo en términos financieros. En este marco, se emplean técnicas estadísticas, como **simulaciones Monte Carlo**, para generar las métricas de riesgo cibernético dinámicas, por ejemplo, pérdida esperada o valor en riesgo (VaR, por sus siglas en inglés). Estos son los resultados finales del ciberriesgo de una potencial víctima que se usarán en el sistema de visualización y apoyo a la toma de decisiones, junto posiblemente otros datos.

## 2.6 DICYME web app: Visualización y apoyo a la toma de decisiones

Los pasos anteriores se integran en una aplicación web para ofrecer todo el flujo de trabajo y dar soporte a múltiples casos de uso que abarcan desde la exploración de datos hasta la evaluación del riesgo basada en indicadores y, finalmente, la cuantificación integral del ciber riesgo, permitiendo una comprensión exhaustiva del riesgo.

Esta aplicación web pone a disposición del usuario todos los modelos y fuentes de datos para medir el ciber riesgo de un conjunto específico de entidades, para las cuales el proyecto ha recopilado datos públicos durante la fase de recolección de evidencias, o bien para entidades ficticias que deben definirse mediante los parámetros requeridos por el modelo.

Además, la inteligencia artificial (IA) facilita al usuario comprender los resultados al proporcionar explicaciones sobre diagramas y resultados. Por otro lado, el sistema incorporará un módulo de recomendaciones, que aprovechará modelos de aprendizaje automático para identificar las variables clave que pueden ajustarse para lograr el mayor retorno en mitigación del riesgo.

# 3 Datos para los indicadores DICYME

En esta sección se explica en más detalle los datos para los indicadore de ciberriesgo de DICYME. Como se explica en el flujo de trabajo de DICYME, los datos para estos indicadores se crean a partir de los datos en crudo de las evidencias recolectadas automáticamente. Las evidencias se seleccionan, limpian, normalizan y organizan para asegurar el valor y la calidad necesarias de los datos para asegurar la calidad de los indicadores. Los datos para los indicadores de ciberriesgo de DICYME se organizan en conjuntos que se presentan en las siguientes subsecciones.

## 3.1 Incidentes cibernéticos

Actualmente, existen múltiples bases de datos públicas que recopilan incidentes cibernéticos, pero suelen estar limitadas a áreas geográficas o sectores industriales específicos.

No hay un repositorio único y autorizado que consolide todos los incidentes a nivel mundial, principalmente porque los datos son altamente sensibles y el acceso está fragmentado entre diferentes actores con distintos niveles de visibilidad y obligaciones regulatorias.

Además, las empresas tienen pocos incentivos para divulgar incidentes salvo que lo exijan las regulaciones, y, aun así, el reporte suele ser mínimo.

Para abordar estos desafíos, DICYME ha desarrollado scripts de automatización que agregan datos de varias bases, tanto específicas de OT como repositorios generales de ciberseguridad. El sistema VISCA de DICYME permite mejorar la calidad y riqueza de los datos identificando automáticamente empresas víctima y datos firmográficos.

La consolidación utiliza técnicas de **Procesamiento de Lenguaje Natural (NLP)** para:

- Fusionar casos duplicados
- Identificar entidades
- Estructurar la información para análisis eficiente.

## 3.2 Perfil de la víctima

Basado en el dataset de incidentes, se han seleccionado entidades impactadas, asegurando que cada objetivo corresponda a una entidad única.

El propósito es construir un dataset que defina entidades mediante características fundamentales:

- Ubicación
- Tamaño
- Atributos financieros
- Reputación

Este dataset sirve como base para modelos que evalúan la probabilidad de que una entidad sea objetivo de actores de amenaza, por ejemplo el indicador de atractivo. El conjunto de datos incluye variables como:

- Localización de sede
- Sector industrial
- Indicadores financieros (ingresos, beneficios, rentabilidad)
- Atributos organizativos (cotización en bolsa, número de empleados)

La base de datos de atractivo también incorpora las siguientes variables dinámicas para algunos de los caos:

- Reputación online (analizada con fórmula temporal en redes sociales)
- Métricas de victimización (incidentes en dispositivos y brechas críticas)

Para comparaciones robustas, se incluyen entidades similares del mismo sector que no han reportado incidentes.

Las fuentes de donde se extrae estos datos se agrupan en **gratuitas** (CompaniesMarketCap, Shodan) y **de suscripción** (financiadas por DeNexus).

## 3.3 Sistemas de detección de intrusiones (IDS)

Además de fuentes externas, DICYME propone indicadores basados en telemetría de sistemas internos de monitorización, agregando datos anonimizados de IDS. Proveedores: **Nozomi OT Security Platform**, **Forescout Technologies**, **Claroty**. El conjunto de datos construido captura métricas clave como:

- Descubrimiento de dispositivos y activos

- Contextualiza vulnerabilidades
- Tiempos de remediación de vulnerabilidades

El dataset está estructurado por:

- Sitio
- Red
- Niveles del Modelo Purdue
- Severidad (CVSS)
- Activo cibernético (por ejemplo, dispositivo conectado a la red)

Estos datos proporcionan información muy sensible de las potenciales víctimas por lo que se enfatiza el uso de información anonimizada para preservar su privacidad sin perder valor analítico.

# 4 Indicadores DICYME

## 4.1 Atractivo

El atractivo es la **probabilidad de que una entidad sea objetivo de ciber amenazas**. El indicador de atractivo de DICYME se basa en tres factores:

- Basal Attractiveness: Analiza características inherentes que influyen en la exposición, usando minería de reglas (FP-Growth) y árboles de decisión para asignar una probabilidad entre [0, 1].
- Online Reputation: Evalúa visibilidad en redes sociales mediante una fórmula ponderada por sentimiento, con valores entre [-0.5, 1].
- Potential Victimization: Clasifica el riesgo en tres niveles (0, 1, 2) según presencia en dispositivos visibles y filtraciones de ransomware.

Estos atributos se integran mediante regresión logística para generar un indicador final normalizado [0, 1]. Validaciones empíricas confirman correlación significativa con incidentes reales.

## 4.2 THRACT

Un indicador diseñado para **clasificar y analizar actores de amenazas** en el tiempo.

- Basado en datos públicos (p.ej., ETDA Threat Actor Encyclopedia).
- Compuesto por tres puntuaciones:
    - Activity Score: Actividad reciente del actor.
    - Capacity Score: Capacidades técnicas.
    - Target Score: Nivel de amenaza según sector e ubicación.

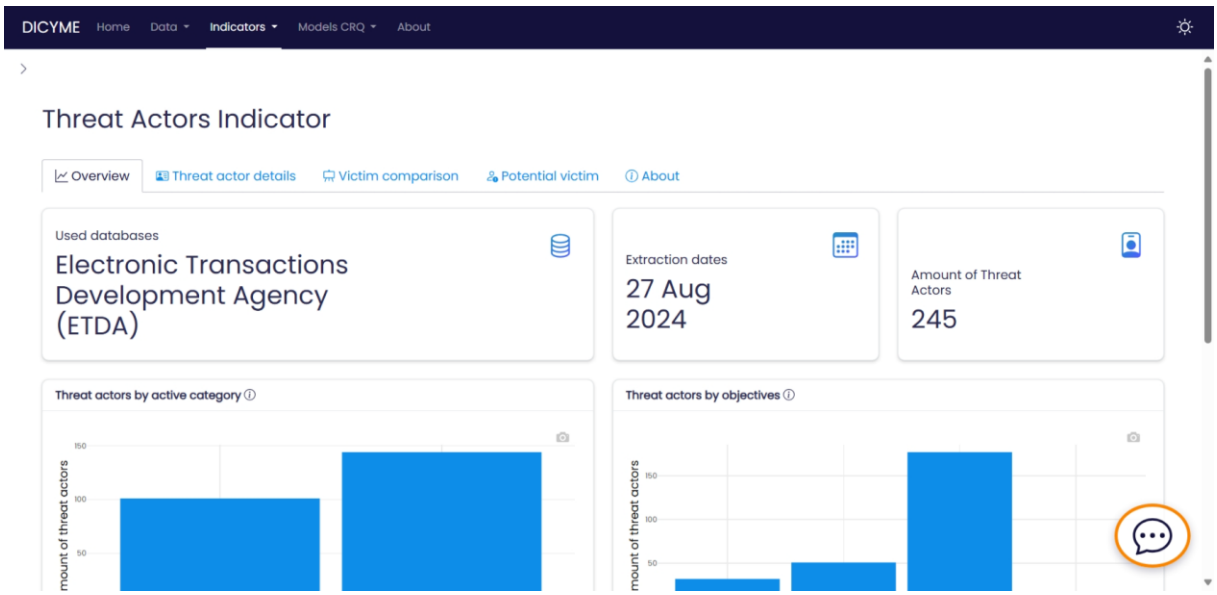



*Ilustración 2: Captura de pantalla de la aplicación web DICYME mostrando la página dedicada al indicador THRACT (Threat Actors Indicator en la imagen).*

Es dinámico, flexible y accesible, permitiendo personalización por industria y región. En la Ilustración 2 se muestra una captura de pantalla de la aplicación web de DICYME para trabajar con este indicador.

### 4.3 CVE2TTs

Aborda la evaluación de **vulnerabilidades conectadas con el ciclo de ataque**.

- Usa el marco MITRE ATT&CK para mapear tácticas y técnicas.
- Propone un modelo de procesamiento de lenguaje natural y usando una red neuronal para clasificación que mapea CVEs a técnicas ATT&CK.
- Actualiza dinámicamente los mapeos para priorizar vulnerabilidades con mayor potencial de explotación real.

Este enfoque mejora la gestión del riesgo al enfocarse en vulnerabilidades más críticas teniendo en cuenta su contexto.

## 5 Cuantificación del ciberriesgo industrial

El marco de cuantificación del ciberriesgo industrial (CQR industrial) en **DICYME** integra múltiples componentes basados en datos para estimar y modelar el impacto financiero y operativo de las amenazas cibernéticas, aprovechando los conjuntos de datos e indicadores desarrollados a lo largo del proyecto. Además, se utilizan datos e inteligencia de última generación del **DeNexus Knowledge Center (DKC)** como referencias y parámetros de entrada para las simulaciones.

Siguiendo la taxonomía **FAIR**, el ciberriesgo industrial se cuantifica como el producto de la **Frecuencia de Eventos de Pérdida** y la **Magnitud de la Pérdida**, garantizando un enfoque estructurado y completo para la evaluación del riesgo. En la Ilustración 3 se muestra la selección de una víctima potencial para CRQ industrial.

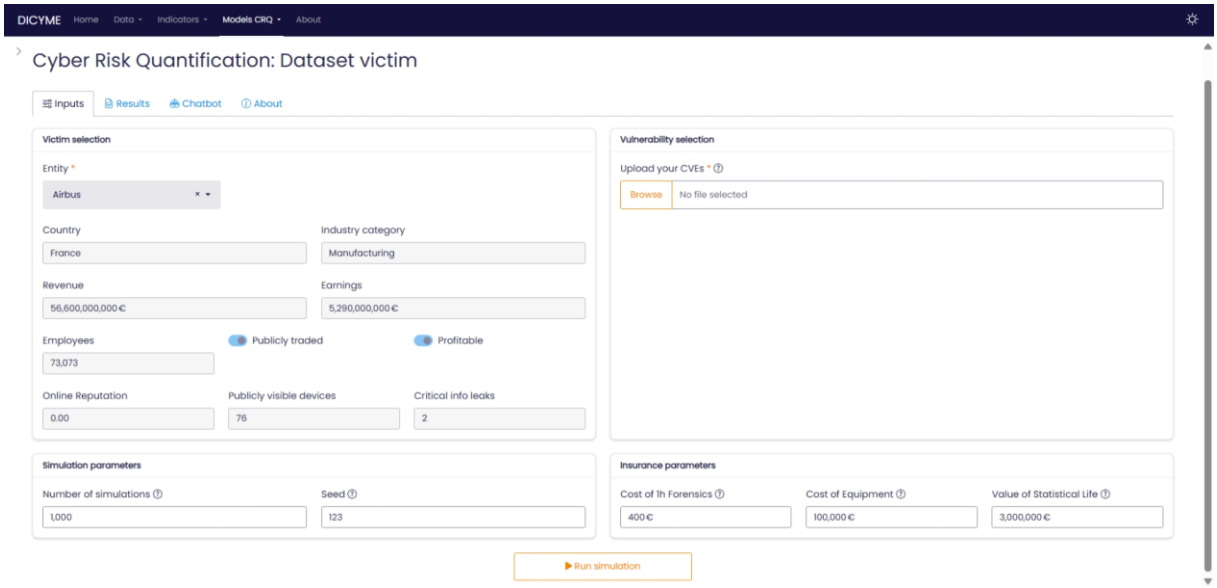



*Ilustración 3: Selección de la víctima en el CRQ industrial de DICYME web app.*

## 5.1 Frecuencia de eventos de pérdida

Representa el número esperado de ciberataques que una entidad puede experimentar en un periodo determinado, comúnmente un año. Se calcula como el producto de la **Frecuencia de Eventos de Amenaza** y la **Vulnerabilidad**.

1. La primera se modela como un proceso distribuido de Poisson, donde la media se calcula a partir de:
    - Una tasa base de ataques (derivada de informes externos específicos del sector).
    - Tendencias de incidentes (estimadas a partir del conjunto de datos de Ciberincidentes).
    - El indicador de **Atractivo**.
2. La probabilidad de un ataque exitoso se calcula utilizando el indicador **THRACT**, la exposición a técnicas MITRE ATT&CK y perfiles de seguridad predefinidos, empleando datos del DKC.

## 5.2 Magnitud de la pérdida

Cuantifica el impacto financiero y operativo de los ciberataques exitosos, definido por pérdidas primarias y secundarias.

Se calcula mediante simulaciones probabilísticas, como distribuciones Gamma y Beta, combinadas con parámetros de ciberseguros (por ejemplo, coste por hora de investigación forense o coste por interrupción de servicio).

También se consideran variables específicas de la entidad, como ingresos o coste por hora de inactividad.

# 6 Visualización y apoyo a la toma de decisiones

La fase de visualización y soporte a la toma de decisiones constituye un componente esencial del sistema DICYME, ya que permite transformar los resultados de los modelos y simulaciones en información comprensible y accionable para los distintos perfiles de usuario.

El sistema de visualización es la aplicación web de DICYME que ofrece una experiencia completa al permitir:

- **Explorar datos y evidencias** recopiladas, facilitando la comprensión del contexto de riesgo.
- **Analizar indicadores clave** (como *Atractivo*, *THRACT* y *CVE2TTs*) y su impacto en la exposición al riesgo.
- **Calcular el ciber riesgo industrial** mediante el motor de cálculo CRQ industrial.
- **Visualizar resultados de cuantificación** mediante gráficos dinámicos y reportes personalizados que presentan métricas financieras y operativas asociadas al ciber riesgo.

Además, el sistema incorpora **mecanismos de explicabilidad basados en IA**, que proporcionan interpretaciones automáticas de los resultados y diagramas, mejorando la transparencia y la confianza en el proceso de análisis.

Finalmente, se incluyen **recomendaciones** que identificará las variables más relevantes para reducir el riesgo y propondrá acciones priorizadas según su impacto y viabilidad. Este enfoque convierte a DICYME en una herramienta no solo de análisis, sino también de **apoyo estratégico a la toma de decisiones** en entornos industriales.

# 7 Integración con DeRISK

DeRISK es un sistema de cuantificación de ciber riesgo industrial basado en modelos probabilísticos, estructurado en tres componentes principales:

- **NoA (Number of Attack Attempts)**: estima la frecuencia esperada de intentos de ataque por vector de acceso inicial.
- **APA (Attack Propagation Algorithm)**: calcula la probabilidad de éxito de un ataque mediante simulaciones probabilísticas sobre grafos de propagación.
- **LEI (Loss Event Impact)**: cuantifica la severidad económica de un ataque exitoso en términos monetarios.

## 7.1 Indicadores integrados en DeRISK

En el marco del proyecto DICYME, se han desarrollado indicadores clave que complementan y enriquecen la arquitectura de DeRISK. Los indicadores desarrollados en DICYME se han adaptado para integrarse con DeRISK, enriqueciendo su arquitectura probabilística:

- **Attractiveness**: actualmente en el plan de desarrollo de DeNexus, este indicador se basa en la base de datos de incidentes cibernéticos y se amplía con información contextual, como el nivel de desarrollo tecnológico del país, para estimar la probabilidad de ser objetivo de un ataque.
- **THRACT**: incorpora datos de actores de amenaza procedentes de ETDA y fuentes privadas como CrowdStrike. Se han ajustado los parámetros y optimizado el código para cumplir con los requisitos de eficiencia y escalabilidad del sistema DeRISK.
- **CVE2TTs**: este componente, basado en técnicas de aprendizaje profundo, se ha adaptado para integrarse con DeRISK y operar en tiempo real, procesando diariamente las vulnerabilidades publicadas y mapeándolas a técnicas MITRE ATT&CK.

La integración de estos indicadores en DeRISK no solo ha requerido ajustes técnicos y metodológicos, sino también la alineación con los procesos internos del sistema, garantizando la coherencia con los modelos probabilísticos existentes. Aunque el trabajo sigue en evolución, representa un avance significativo hacia una cuantificación más dinámica y basada en evidencia, reforzando la capacidad de DeRISK para ofrecer estimaciones robustas y accionables en entornos industriales.

## 7.2 Resultados y validación

Para garantizar la robustez del sistema, los indicadores desarrollados en DICYME que se han integrado en el sistema DeRISK han sido probados con datos reales de clientes en entornos controlados. Estas pruebas han permitido validar:

- La correcta interacción entre los módulos.
- La coherencia de los cálculos.
- La escalabilidad del sistema en escenarios industriales.

Los resultados preliminares confirman que la integración mejora la capacidad predictiva y la actualización dinámica de los modelos, alineándose con los objetivos del proyecto.

## 7.3 Impacto en DeRISK

La integración de los indicadores desarrollados en DICYME en el sistema DeRISK supone un avance significativo en la cuantificación del ciber riesgo industrial, tanto desde el punto de vista técnico como estratégico. Entre los principales impactos destacan:

- **Mejora en la precisión y granularidad**: el uso de indicadores como *Attractiveness*, *THRACT* y *CVE2TTs* permite modelar la probabilidad y el impacto de incidentes considerando más factores dinámicos y contextuales que antes.
- **Actualización continua y escalabilidad**: la capacidad de procesar vulnerabilidades diariamente y mapearlas automáticamente a técnicas MITRE ATT&CK reduce la dependencia de procesos manuales.
- **Impacto estratégico para DeNexus**: esta integración refuerza la posición de DeRISK como plataforma líder en CRQ industrial, diferenciándola por su enfoque dinámico y basado en evidencia.

En conjunto, estos avances consolidan la transferencia tecnológica entre el ámbito académico y la industria, y abren la puerta a nuevas líneas de innovación orientadas a la gestión proactiva del ciber riesgo.

# 8 Conclusiones y próximos pasos

El sistema DICYME ha alcanzado un nivel de madurez que permite su integración (con las correspondientes adaptaciones) con la plataforma de DeNexus, aportando indicadores avanzados y capacidades dinámicas para la cuantificación del ciber riesgo industrial. La validación funcional confirma la coherencia y escalabilidad del sistema, reforzando su potencial como herramienta de referencia para operadores industriales y aseguradoras.

El sistema desarrollado representa un avance significativo en el campo de la cuantificación del ciber riesgo, demostrando el valor de la colaboración entre el ámbito académico y la industria. La asociación entre la Universidad Rey Juan Carlos y DeNexus ha fomentado una transferencia de conocimiento productiva, no solo mejorando las capacidades tecnológicas del modelado de ciber riesgo, sino también asegurando que estos desarrollos sean accesibles para una audiencia más amplia.

DICYME podría seguir desarrollándose con los siguientes pasos:

- Enriquecer la implementación del módulo de recomendaciones basado en IA.
- Ampliar las pruebas con diferentes perfiles industriales y escenarios simulados.
- Incorporar visualizaciones avanzadas y explicaciones automáticas en la interfaz.
- Incorporación de fuentes de datos adicionales y construcción de nuevos indicadores dinámicos.
- Análisis más detallado de la postura de seguridad observada desde Internet